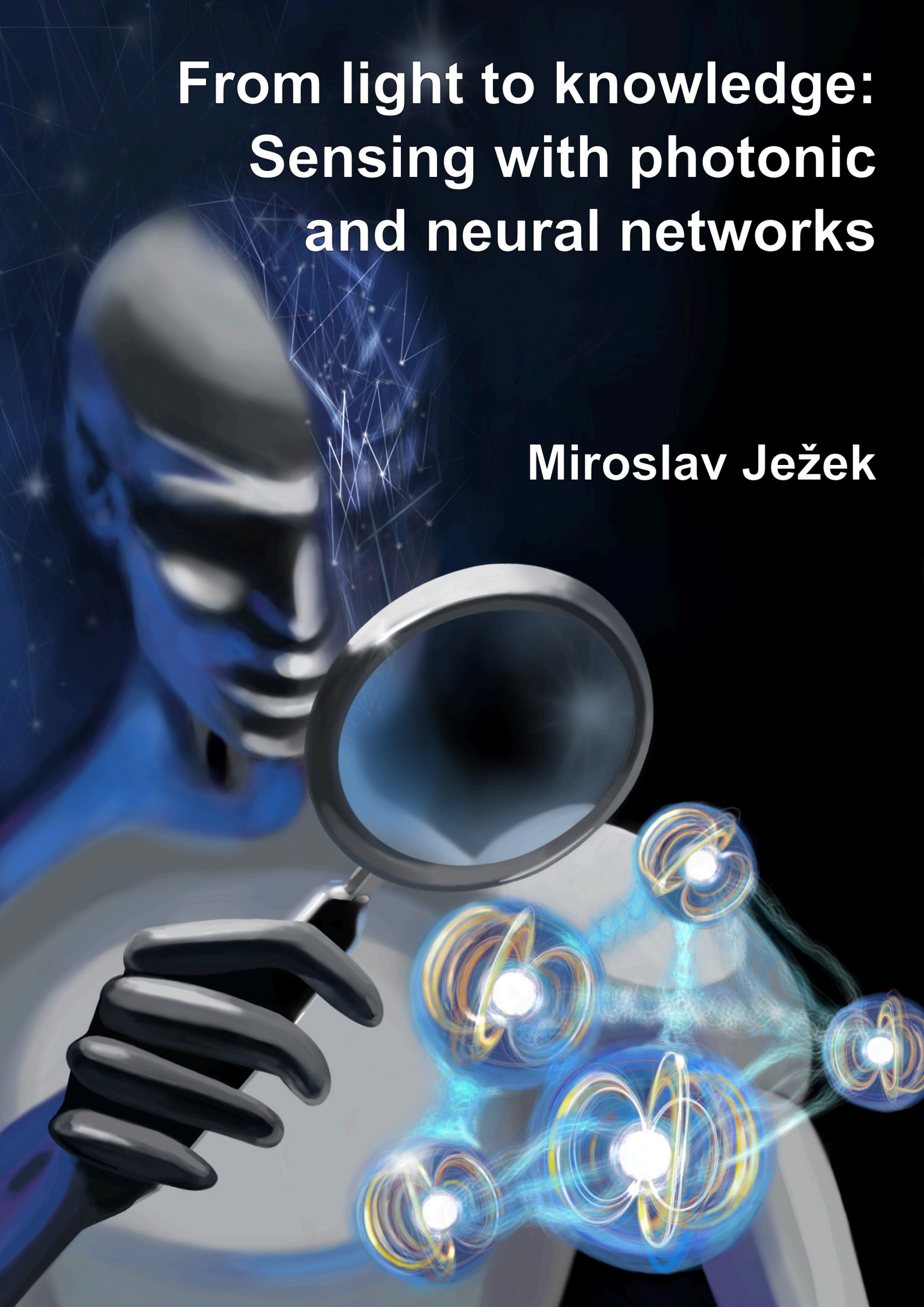
From light to knowledge: Sensing with photonic and neural networks

Miroslav Ježek



From light to knowledge: Sensing with photonic and neural networks

Miroslav Ježek

Habilitation thesis
From light to knowledge: Sensing with photonic and neural networks

Table of contents

1. Introduction **1**

1.1. The importance of light for sensing 1
1.2. Vision: Maximal accuracy and minimal calibration 2
1.3. Structure of the thesis 4

2. Overview of core detection and sensing techniques, the state-of-the-art **5**

2.1. Light routing and scattering 6
2.2. Photonic detectors: Toward resolving individual photons 12
2.3. Statistical properties of light 17
2.4. Detector response and calibration 17
2.5. Vector properties of light 18
2.6. Nonclassical correlations and entanglement 21
2.7. Single emitters and nanosensors 23
2.8. Artificial intelligence: Optimal control and readout 26

3. Single-photon detector non-linearity **33**

4. Photon statistics detection, nonclassicality sensing, and emitter counting **39**

5. Optimal preparation, control, and sensing of polarization states of light **47**

6. Quantum correlation sensing **55**

7. Device-agnostic super-resolution imaging **63**

8. Summary, outlook, and open questions **71**

8.1. What we learned 71
8.2. Outlook: toward programmable, multiparameter, and self-calibrating sensors 72
8.3. Open questions and challenges 73

References **75**

1. Introduction

1.1. The importance of light for sensing

Light is our most precise ruler and at the same time the fastest carrier of information. It plays a central role in sensing and metrology, offering a versatile, precise, and non-invasive probe across scales, from quantum systems to celestial bodies. Many modern sensors start by detecting light (or, more widely, electromagnetic radiation) and then convert that optical reading into the required physical, chemical or biological quantity, whether the instrument is used in a laboratory, a materials facility, a clinic, an environmental station, or on a space telescope.

At the quantum level, measurements of photon statistics and correlations provide direct probes of quantum signatures, namely quantum coherence and entanglement, which are foundational to quantum physics. They let us witness the quantum character of physical systems and their nonclassical dynamics [NobelPrize2022], expose quantum phase transitions [Baumann2010, Aspuru-Guzik2012], and even explore whether similar effects play a role in living organisms [Panitchayangkoon2010, Lambert2013, Xu2021, LiQ2023]. From an application perspective, the detection and characterization of quantum light, such as individual or entangled photons and squeezed light, is essential for quantum communication [Wehner2018, Pirandola2020], quantum computing [Flamini2019, Slussarenko2019, Wang2025], quantum metrology [Giovannetti2004, Giovannetti2011, Polino2020, LiuZ2025], and quantum sensing [Degen2017, Aslam2023].

In metrology, the phase, amplitude, frequency, and polarization of light are used to measure distance (e.g., laser interferometry), time (optical clocks), and external fields (electric, magnetic, gravitational). The precision of optical measurements underpins standards such as the meter. In 2015, the Laser Interferometer Gravitational-Wave Observatory (LIGO) detected mirror displacements of 10^{-18} m, corresponding to a gravitational-wave strain of 10^{-21}, using high-power coherent light [LIGO2016]. In 2019, the sensitivity was suddenly improved by a factor of two through the use of squeezed light, expanding by eightfold the volume of the observable universe [Tse2019, Acernese2019]. This represents an unprecedented demonstration of the literally cosmic-scale impact of subtle quantum-noise manipulation. Not all metrological protocols require high power or kilometer-scale setups: in magnetometry, for instance, sensitivities down to fT/√Hz are achieved by detecting weak fluorescence signals either in wide-field imaging (with spatial resolution of 250 nm) or with a scanning probe (below 10 nm, but at the cost of reduced sensitivity) [Budker2007, Pelliccione2016, Amawi2024].

Microscopy is another area where light serves as a key information carrier. State-of-the-art microscopy techniques use various parameters of light to visualize structures and processes in a sample and beat the limits of conventional intensity microscopy. Phase (holographic) microscopy measures optical path and dry mass in living cells, without the need for staining or labeling [Huang2024, Běhal2025]. Fluorescence microscopy exploits the specific emission of tagged molecules to reveal functional

and biochemical activity within cells. The resolution of single-molecule localization microscopy goes significantly beyond the diffraction limit down to 10 nm [Rust2006, Betzig2006, Lelek2021]. Polarization microscopy maps molecular orientation and order in tissues [Brasselet2019]. Single-photon detection in confocal and multiphoton systems increases resolution and depth, enabling real-time, in vivo imaging of biological processes. Photon statistics and their correlations are used to decode number of molecules [Ta2015], track single emitters [Israel2017], and formulate an electronic model of quantum emitters [Fishman2023]. These advanced functional microscopy techniques play crucial role in material research, environmental monitoring, and biomedical imaging.

In chemistry and materials science, spectroscopy with light identifies molecular composition and structure. Absorption, emission, and Raman scattering spectra give fingerprints of substances, enabling quantitative analysis of concentration and reaction kinetics [Das2011, Poonia2024]. This is essential for process monitoring, environmental sensing, and molecular engineering.

In biology and medicine, optical techniques such as spectroscopy, optical coherence tomography, fluorescence imaging, photoacoustic sensing, time-gated diffuse optical imaging, and various microscopy modalities are routinely used and enable non-invasive diagnostics. Light reveals cellular dynamics, tissue morphology, and functional biomarkers, aiding in early disease detection and therapy monitoring. Light-aided noninvasive glucose sensing promises to improve life of millions [Arnold2005, ZhangY2025].

In astronomy, detecting and analyzing light from distant sources allows the resolution of individual stars and galaxies [Cava2018, Zhang2024]. Techniques like interferometry and adaptive optics improve angular resolution, while spectroscopic measurements reveal stellar composition, motion, and redshifts. Low-light sensing and photon counting are crucial for observing weak sources and time-correlated measurements.

In all these domains, the ability of light to be generated, manipulated, and detected with high sensitivity and accuracy, along with its various classical and quantum properties, makes it indispensable. It enables sensing with precision limited only by fundamental physical laws, and continues to drive innovation in both fundamental and applied science.

1.2. Vision: Maximal accuracy and minimal calibration

Light carries an abundance of physical information yet each detected photon is costly in power, sample dose, or acquisition time. Modern quantum technologies and nanoscale imaging push experiments into regimes where only a handful of photons are available, so information efficiency rather than raw count rate becomes the critical figure of merit to reach the ultimate accuracy.

The central ambition of the work collected in this thesis is therefore to force each photon to speak as loudly as physics allows. To this end, we use a combination of

measurement projections implemented using optical networks, single-photon detection to reach the ultimate sensitivity, and efficient data post-processing to extract the target information.

The route we follow has five main building blocks, see Fig. 1.1:

i. Design an optical network—sequenced, multiplexed, ordered, or deliberately random—that redistributes the incoming light from a sample over a few (or many) detection channels. This process projects the quantity of interest (carried by the light) onto a detection signature.
ii. Record the raw detection statistics (clicks) with single-photon sensitivity.
iii. Characterize and calibrate the network just enough to make its response reproducible. Determine detection parameters (efficiency, noise, linearity) with metrological protocols that are themselves immune to modeling bias.
iv. Process or invert the detection output with either a statistical estimator or an artificial intelligence modality (typically an artificial neural network) trained on experimental or synthetic data.
v. Report the quantity of interest directly, without reconstructing excess degrees of freedom.

By chaining those five steps we will witness a massive information gain in a series of experiments ranging from counting photons to imaging stars, from sensing quantum entanglement to resolving individual molecules.

A long-term aspiration is to reduce explicit calibration to the bare minimum. We will focus on deep-learning approaches, such as artificial neural networks, trained on synthetic data or using self-supervision. The net result is a progressive shift of labour from painstaking laboratory calibration toward digital processing, self-calibration, or, ideally, measurement-device independence.

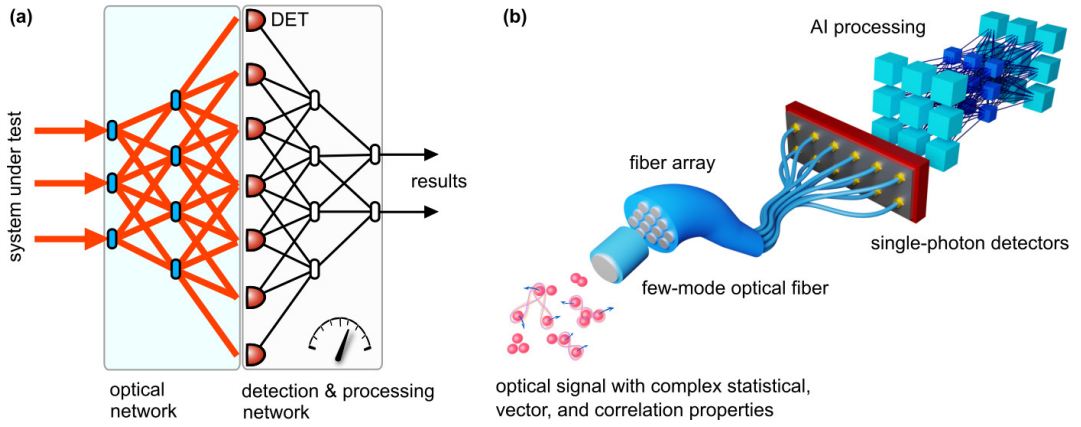


Figure 1.1. (a) Conceptual scheme of a network-based photonic sensor. Light emitted by an observed system is routed in an optical network, detected, and processed by a computational network. The optical part produces a scattering pattern that encodes information carried by the input signal. The subsequent detection (DET) and processing network decodes the information and outputs the target quantities or classifications. (b) An example of the network-based sensor using a short piece of a few-mode optical fiber as a random scattering network, a fiber array as an interface to single-photon detectors, and a deep neural network for decoding the complex properties of the input light [Bielak2025].

1.3. Structure of the thesis

First, we will survey the state-of-the-art in photonic sensing with focus on quantum and quantum-inspired approaches in Chapter 2. Chiefly, we introduce concepts of optical networks as a metrology tool, single-photon detection, key parameters of light used as information carriers, single emitters and nanosensors, and artificial intelligence approaches relevant for photonic sensing. We also briefly review the key experimental and theoretical tools for advanced photonic detection and sensing, highlighting hardware and methods developed in our group under my supervision. In Chapter 3, we will discuss nonlinear response of single-photon detectors and its accurate calibration. Chapter 4 will focus on multiplexed photon-number-resolving detection and sensing of photon statistics and advanced statistical features. In Chapter 5, we will present optimal control and sensing of vector properties of light. Chiefly, we will introduce all-fiber polarization micro-endoscopy as an example of a sensor based on a random scattering network followed by a computational neural network. In Chapter 6, we will study quantum entanglement sensing using deep neural networks. We will focus on processing incomplete measurement data and device-agnostic approaches. Chapter 7 will report on full device-agnostic modeling aiming at universal super-resolution imaging of single emitters. Finally, in Chapter 8, we will summarize the results and discuss open questions and future outlook.

2. Overview of core detection and sensing techniques, the state-of-the-art

In this chapter, we introduce the basic concepts and techniques relevant to advanced photonic sensors, which cover routing, scattering, and detection of photons, key parameters of light used to encode information, interactions with nanoemitters, and data-driven computation.

Optical networks serve as versatile "light scattering machines" for imaging, communications, metrology, and quantum technology. They host quantum walks and, with many photons, boson-sampling experiments that already outperform classical simulations and reach quantum advantage. They also probe particle statistics, reveal entanglement and topological phases of matter, and perform quantum tomography. Multiplexing and scattering also enable accurate spectral and temporal analysis, laser wavelength locking, polarimetry, and single-molecule mapping and imaging. Scaling up now hinges on faster, noise-tolerant ways to characterize and control large, ordered and disordered networks.

Sensitive and accurate detection of light represents another core element for developing cutting-edge optical technology and extending the frontiers of photonic science. The field increasingly targets high-sensitivity detection at quantum limits. Low-light and single-photon detection are vital not only for quantum technology applications, such as quantum cryptography and photonic quantum computing, but also for advanced biomedical imaging and metrology. In-situ and in-vivo characterization at microscopic scales and with limited light intensity naturally demands ultimate detector performance. The state-of-the-art photonic detection includes room-temperature single-photon avalanche diodes and superconducting single-photon detectors based on bolometers or superconductivity disruption.

Precise multi-parameter optical sensing is another key trend in modern photonics: measuring photon statistics, vector properties, and correlations goes beyond simple intensity detection and provides significant information about optical signals and their interaction with the environment. Photon statistics characterization benchmarks quantum sources and enhances advanced fluorescence microscopy. The characterization and control of the vector properties of light – polarization – lie at the heart of materials and surface diagnostics, polarization microscopy, and quantum technologies. Correlations of light uncover properties of distant stars as well as nanoparticle interactions. In the quantum regime, nonclassical correlations are essential for quantum advantage in communication and computing.

Interaction of light with matter at the nanoscale enables sensors that combine the precision of photonic probes with the environmental sensitivity of single quantum emitters. Single atoms, ions, molecules, and solid-state emitters respond to electromagnetic, mechanical, and chemical changes while providing efficient fluorescence readout. Engineered or field-tunable levels are available in many solid-state platforms, while room-temperature and on-chip operation are feasible in both solid-state and some molecular systems. These features yield versatile, scalable

probes for metrology, materials science, and biological sensing, reaching regimes inaccessible to photons alone.

Finally, advanced statistical techniques and artificial intelligence modalities represent efficient tools to control photonic devices, calibrate sensors, and extract information from measured data. Many theoretical breakthroughs in deep learning and a significant increase in computational power in the last decade have led to a paradigm shift in detector and sensor readout and imaging technology.

2.1. Light routing and scattering

Optical networks are crucial to imaging, optical communications, photonic information processing, and optical and quantum metrology. Light propagating between individual nodes of the network is repeatedly split and combined. This process modulates the initial optical signal and produces a scattering pattern, which conveys the information to the output. The pattern is detected and the resulting classical signal is processed by electronic circuitry and numerical methods, often possessing some sort of network architecture, see Fig. 1.1 (a).

There is a long history of using ordered and disordered networks for information processing tasks and, particularly, to demonstrate quantum advantage. Quantum walks represent an archetypal application of network propagation for information processing. A traveling quantum particle moves between the nodes of an ordered network spreading its possible position distribution faster than in a classical random walk. This has been exploited in various applications, including efficient search on graphs [Childs2004] and studies of topological phenomena [Kitagawa2012]. Quantum walks have been implemented by photonic technology using spatially- and time-multiplexed networks [Perets2008, Schreiber2010, Kitagawa2012, Xiao2020]. By introducing disorder in the quantum walk it is possible to simulate Anderson localization and noise-assisted quantum transport [Schreiber2011, Segev2013, Maier2019]. Quantum walks with two or more particles were proposed for universal quantum computing [Childs2013] and conveniently realized using photon pairs in waveguide lattices, time-bin-encoded networks, and multimode fibers [Bromberg2009, Peruzzo2010, Schreiber2012, Defienne2016]. The advancement of experimental quantum walks is severely limited by the necessity of characterizing and controlling large optical networks. Boson sampling, introduced as a restricted model of non-universal quantum computation [Aaronson2013], generalizes quantum walks in two main directions: 1. many particles with bosonic statistics scatter in the network; 2. the network is heavily multimode and disordered, with parameters of the optical nodes chosen randomly across the network. Experimental implementations of boson sampling based on photonic networks [WangH2019, Zhong2020, Madsen2022, Deng2023, Maring2024, Wang2025, LiuH2025] and random-circuit sampling in superconducting quantum circuits [Arute2019, Morvan2024, Gao2025] currently tackle the region of parameters that is not treatable by classical computers. Random networks can be precisely controlled by engineering intermodal interference, e.g., in multimode optical fibers, using spatial modulation and deep learning, giving rise to a new class of quantum devices [Leedumrongwatthanakun2020, Goel2024].

Optical networks have also found applications in quantum detection methods and quantum metrology. Particularly, the sensitivity of the quantum walk of two or more particles to their correlations allows for distinguishing bosonic, fermionic, and anyonic statistics [Sansoni2012, Matthews2013]. In photonic walks, where the statistics is emulated by the entanglement of photons, this can be exploited to distinguish various entanglement classes [Leedumrongwatthanakun2025]. Furthermore, the identification of topological phases of matter was proposed using classical deep learning from scattering patterns formed in quantum walks [Ming2019]. Quantum walks and scattering processes in general can also facilitate detailed characterization of a physical system by means of quantum tomography. Tomographic methods are essential to quantum technology and are routinely used for full characterization of single-photon sources, entanglement sources, and optical networks. Quantum tomography of two- and three-photon signals was reported by measuring all detection events at the output of a fixed optical network [Titchener2018]. The minimum number of output ports of the network scales linearly with the number of photons but the achieved fidelity of the retrieved quantum states decreases with their increasing rank. Also the resilience of the method to excess noise represents an issue [Oren2017]. The major limitation of the network-assisted tomography lies in the necessity of detailed characterization of the network [Fyrillas2024]. Even for a random network, its structure has to be known or experimentally measured prior to the tomography, which is extremely difficult for large networks.

In optical metrology, networks can act as analyzers that translate spectral, temporal, or polarization information into patterns that are read out with a single detector or an array of detectors. Common dispersive-Fourier-transform or photonic time-stretch spectroscopy use a long section of highly dispersive fiber, sometimes combined with in-line Raman or Brillouin amplification, which maps an optical spectrum onto a stretched temporal waveform that is captured in one shot by a fast oscilloscope or time-to-digital converter. The amplified variant overcomes loss and has resolved sub-gigahertz features in single-shot absorption spectra while maintaining megahertz frame rates, an ability that has become invaluable for tracking nonequilibrium laser dynamics and ultrafast chemical reactions [Solli2008, Solli2012]. Replacing a photodiode by a single-photon detector allows for fast spectroscopy at the single-photon level inaccessible to common spectrometers [Avenhaus2009, Sun2024].

Random or pseudo-random networks analyze wavelength spectra through speckle encoding. Multiple scattering inside a disordered silicon-on-insulator network mesh produces wavelength-dependent intensity fingerprints; after a one-time calibration the input spectrum is reconstructed computationally with 0.75 nm resolution in a 25 µm-radius device, proving that disorder can beat the size–resolution trade-off that limits gratings and arrayed-waveguide spectrometers [Redding2013]. The same principle applies to a multimode fiber: interference between thousands of guided modes creates a unique speckle for every wavelength. A 100 m coil achieves 1 pm resolution at 1.5 µm, outperforming large bench-top echelle spectrographs while occupying only a few cubic centimeters. Short (4 cm) fibers deliver broadband 400–750 nm coverage with 1 nm resolution, making coilable computational spectrometers attractive for field sensing [Redding2014]. Free-space scattering in random media can serve the same purpose: its wavelength-specific speckle pattern enables a broadband wavemeter with sub-femtometer resolution , and, embedded in a feedback

loop, it locks a laser to below-megahertz linewidths [Metzger2017, Gupta2020]. Programmable interferometric networks close the gap between random and deterministic analyzers. A recent six-stage Mach–Zehnder network reaches <10 pm resolution over a 200 nm span—an order-of-magnitude improvement in the bandwidth–resolution product of integrated spectrometers. Because the network is reconfigurable it can trade resolution for throughput in real time, opening prospects for adaptive sensing [Yao2023].

Various optical networks have found applications in polarimetry and imaging. Scattering in random media enabled single-shot measurement of light polarization [Juhl2019, Facchin2020] and analysis of optical vector fields [Pierangeli2023]. Replacing a free-space scattering to intermodal interference in a short piece of a few-mode fiber and a few-pixel detector sampling allowed for in-situ polarization measurement with kHz repetition rate and single-photon sensitivity [Bielak2025]. Lensless imaging [Freund1990, Bertolotti2012, Katz2014, Edrei2016, Bertolotti2022] builds upon advances in complex media characterization and control [Rotter2017]. Random scattering in multimode fibers was demonstrated to support visual information transmission [Čižmár2012, Borhani2018, Caramazza2019] and unlocked breakthroughs in endoscopy imaging [Wen2023, Stibůrek2023].

The experimental implementation of photonic networks is typically realized in the spatial or temporal domain. Spatially multiplexed networks have been built in free space using discrete elements (typically in 2D) or via optical-contact bonding in 3D [WangH2019, Zhong2020, Deng2023], and as waveguide circuits on optical chips. Free-space networks are typically constructed from bulk elements such as wave plates, phase plates, beam splitters, polarizing beam splitters [Paul1996, Hloušek2019, Hloušek2024a, O'Brien2003, Lanyon2009], and birefringent beam displacers [Kitagawa2012, Xiao2017, Stárek2018a, Xiao2020], see Fig. 2.1 (a). They offer only a limited number of spatial modes but allow full control over internal structure, straightforward reconfigurability, and high efficiency. In our lab, we have developed various free-space, spatially multiplexed networks for photon-number-resolving detection [Hloušek2023, Večerka2025] (see Chapter 4) and for quantum information processing [Stárek2022] (see Fig. 2.2). Controlling a larger number of spatial modes enables structured light with many applications in imaging, sensing, communication, and optical neural networks [ZhangZ2025].

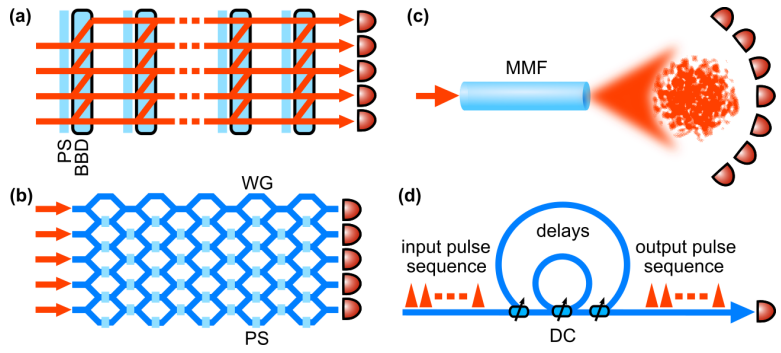


Figure 2.1. Examples of optical networks based on spatial (a,b,c) and time (d) multiplexing. Legend: PS – phase plate, BBD – birefringent beam displacer, MMF – multimode fiber, WG – waveguide, DC – dynamically controlled coupler.

To improve scalability, waveguide circuits integrated on an optical chip have been massively used for information processing and metrology [Perets2008, Bromberg2009, Peruzzo2010, Sansoni2012, Matthews2013, Titchener2018, Maring2024, Wang2020], see Fig. 2.1 (b). There is a plethora of waveguide manufacturing techniques [Bogaerts2020, Wang2020] used to prepare on-chip circuits, such as lithography, ion diffusion, and direct laser writing. Material platforms include silicon-on-insulator (Si, SOI) [Silverstone2014, Qiang2018, PsiQuantum2025], silica-on-silicon (SiO2) [Politi2008, Matthews2009, Carolan2015], laser-written silica [Smith2009, Sansoni2010, Crespi2011, Meany2015], silicon nitride (Si_3N_4) [Schuck2016, Lu2019], lithium niobate ($LiNbO_3$, Ti:$LiNbO_3$) [Jin2014, Luo2019], gallium arsenide (GaAs) [Wang2014], indium phosphide (InP), and many others. Si waveguides can tightly confine light, allowing direct single-photon generation in the waveguides and high-density integration [Qiang2018, PsiQuantum2025]. Laser-written silica circuits provide possibilities for investigating complex physical systems through three-dimensional circuit geometries [Meany2015]. Si_3N_4 provides ultralow propagation losses and can create single photons within a broad transparent window [Lu2019]. While the above are generally passive materials, $LiNbO_3$ [Luo2019], GaAs [Wang2014], and InP exhibit strong electro-optic properties that allow for fast manipulations of single photons. Despite recent progress of integrated photonic networks, their accurate characterization remains a significant roadblock [Fyrillas2024].

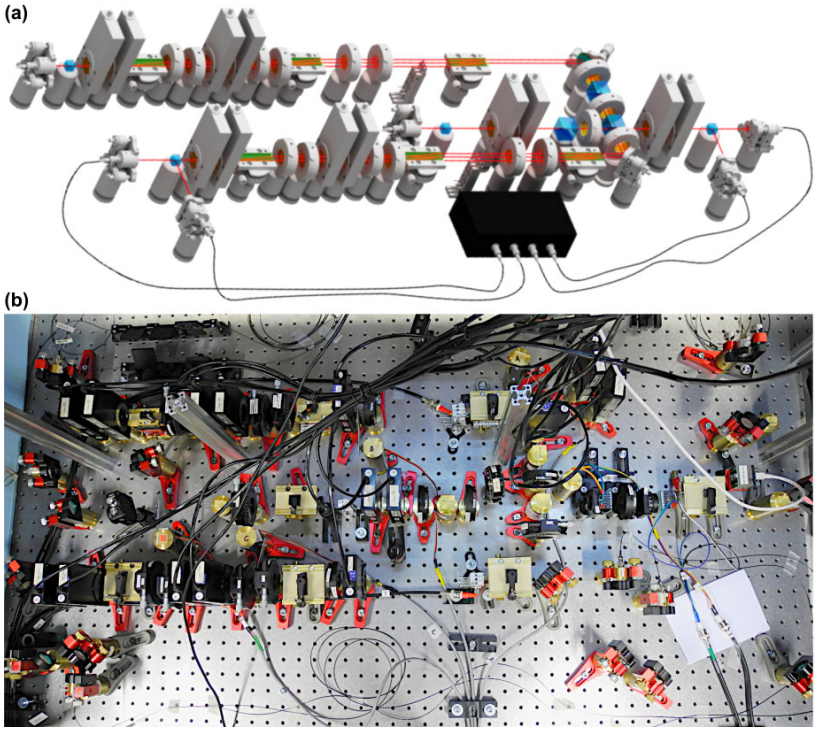


Figure 2.2. (a) Model and (b) photo of a free-space discrete photonic network based on inherently stable multi-path interferometers using birefringent beam displacers [Stárek2018a]. Photo from our lab (QOLO) by R. Stárek, see [Stárek2022] for details.

Alternatively, common scattering media such as optical diffusers and multimode fibers have been implemented as random networks [Mosk2012, Rotter2017] for various applications in photonic information processing [Huisman2014, Defienne2014, Huisman2015, Defienne2016, Wolterink2016,

Leedumrongwatthanakun2020, Goel2024], imaging [Čižmár2012, Borhani2018, Caramazza2019, Wen2023, Stibůrek2023], and sensing [Metzger2017, Juhl2019, Gupta2020, Facchin2020, Pierangeli2023, Bielak2025, Leedumrongwatthanakun2025], see Fig. 2.1 (c). Wavefront shaping is used in such cases to encode a particular transmission function [Popoff2010, Kim2012, Vellekoop2015, Cao2023]. However, full characterization of the scattering medium is extremely challenging [Popoff2010, Carpenter2016], and the stability of the transmission function of a multimode fiber is quite limited [Caravaca-Aguirre2013, Wen2023, Bielak2025]. Our recent result on using a few-mode fiber for high-stability polarization sensing is discussed in detail in Chapter 5 [Bielak2025].

Wavefront shaping uses a spatial light modulator (SLM) to program the input field so that, after the mixing in a scattering medium, the desired pattern (e.g., a diffraction-limited focus) appears at the output. Typically, coherent laser light is modulated on an SLM and coupled into the medium, e.g., proximal facet of a multimode fiber. A camera at the distal facet records a random interference pattern (speckle) that serves as feedback for either iterative optimization or model-based control. The same hardware also supports transmission-matrix methods: a sequence of SLM-generated probe fields is sent while the output interferes with a phase reference (external arm or co-propagating through the fiber), and phase-stepping ($\geq$3 steps over 2π) reconstructs the complex field at every camera pixel, effectively measuring one column of the transmission-matrix per shot. Intensity-only schemes (phase retrieval) and digital phase-conjugation variants are also used. Choice of the SLM is a trade-off between pixel count, modulation depth, diffraction efficiency, frame rate, and upload latency. Liquid-crystal SLMs provide direct phase control with high diffraction efficiency (off-axis, up to ~90%) and fine spatial resolution, but are typically limited to a few hundred hertz; ferroelectric LC devices reach kHz rates at the expense of binary $0/\pi$ phase and low first-order efficiency (~10%). Micro-electromechanical system options, especially digital micromirror devices, offer tens-of-kHz refresh and operate as binary amplitude modulators. Complex-field control is achieved in the first diffraction order using Lee holograms, though off-axis efficiency is only a few percent. Efficiency and bandwidth can be improved with double-pass geometries that realize binary phase modulation and correct spatial dispersion; emerging MEMS such as grating-light-valve modulators push toward hundreds of kHz. In practice, liquid-crystal SLMs are favored for power-efficient, high-fidelity control, whereas digital micromirror devices enable high-speed wavefront shaping for dynamic scenes and fast endoscopy.

Time-multiplexed networks have been used for random walks [Schreiber2010, Schreiber2012], boson sampling [He2017, Madsen2022], and quantum information processing [Yu2024, Singh2025], namely preparation of large cluster states and entanglement synthesis [Asavanant2019, Takeda2019, Larsen2019]. Time-bin multiplexing has also enabled more efficient single-photon sources [Kaneda2019] and photon-number-resolving detectors [Banaszek2003, Achilles2003, Fitch2003, Řeháček2003, Mičuda2008, Tiedau2019, Švarc2019]. Together with spatial multiplexing it holds the current record in reaching quantum computational advantage in Gaussian boson sampling with 1024 high-efficiency squeezed states injected into a hybrid spatial-temporal encoded, 8176-mode, programmable photonic quantum processor, Jiuzhang 4.0 [LiuH2025]. Temporal modes are encoded as time bins (or pulses) and the networks typically use loop based architecture, where the

optical signal repeatedly passes through a loop fed by a beam splitter with a switchable splitting ratio, see Fig. 2.1 (d). Such switch can be implemented either as an interferometer with an electro-optic phase modulator [Švarc2019] or using polarization modulation in a Pockels cell [He2017, Takeda2019]. The switching requires precise timing and fast electro-optic control; this can be done using optical fibers and integrated modulators (at the expense of higher losses) [Švarc2019] or in free-space with bulk modulators (complicated high-voltage driving and slower operation) [He2017, Takeda2019]. The main issue with the latter is a severely limited number of switching levels (corresponding to driving voltages). Previous works demonstrated up to eight fast-switchable levels with moderate losses [He2017] and four consecutive fast-switchable levels with low losses [Takeda2019]. Losses and other imperfections of the employed modulators and delay loops affect the maximum achievable number of time bins. The ultimate goal is to reach a high number of switching levels and time bins simultaneously. Multi-level switching and low-latency control based on field programmable gate arrays (FPGA) allow for driving the time-multiplexed network in different regimes, such as balanced operation, low-loss regime, random network etc. Reaching a high number of channels requires the sufficiently low losses and high extinction ratio, which is experimentally demanding. Time response and high efficiency can be boosted by using superconducting nanowire single-photon detectors.

Very recently, we developed a low-loss, multi-level, free-space electro-optical network for single-photon applications [Švarc2024, Mazin2025] that combines bulk optics with a FPGA-controlled high-voltage driver to realize polarization switching with 25 programmable levels (0–500 V), 7% optical loss per loop cycle, 23 dB extinction ratio, and 16 ns switching time. We employed a two-crystal KD*P Pockels cell inside a 60 ns optical delay loop (~18 m) driven by a 300 MHz-clocked Cyclone-V FPGA (3.3 ns timing granularity and 31 ns electronics latency) and a custom high-voltage switch that provides 25 distinct transmittance settings via level combinations. We reached stable generation of up to 16 balanced time-bin patterns with near-uniform probabilities (channel relative entropy of 0.9991) and demonstrated arbitrary control of time-bin intensities (e.g., a linear decay). Overall, our platform significantly improves time multiplexing, retaining very low loss relative to prior free-space electro-optic modulation approaches, see Fig. 2.3. We also developed custom timing electronics for universal pulse control [Hošák2018] and delay generation [Mazin2021].

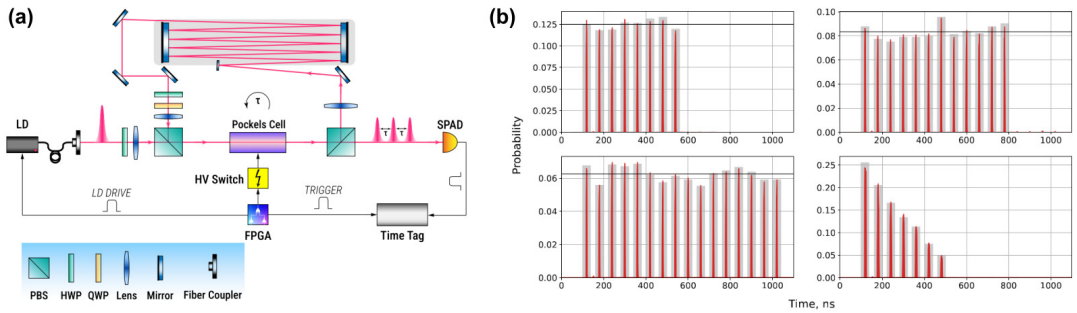


Figure 2.3. (a) Scheme of the loop-based time-switching network and (b) results of various time-bin configurations [Švarc2024, Mazin2025]. Legend: laser diode (LD), single-photon avalanche diode (SPAD), polarizing beam splitter (PBS), half-wave plate (HWP), quarter-wave plate (QWP). Image by G. Mazin [Mazin2025].

2.2. Photonic detectors: Toward resolving individual photons

Single-photon detectors are fundamental for ultra-sensitive sensors and photonic detection in general [Hadfield2009, Buller2010, Eisaman2011, Migdal2013]. They are able to register the arrival of individual photons with high efficiency (detection probability) while keeping false detections low. Many single-photon detectors follow the same high-level bistability operational principle: The device is held in a metastable state, which stores energy that can be released when the state is disturbed by absorbing a photon. The microscopic excitation pushes the biased detector over its threshold. A positive-feedback mechanism turns that single-photon event into a macroscopic change in current or voltage. External electronics senses the electrical pulse, shapes it to a logic-level "click", and then brings the device back to its initial state so it can detect the next photon.

There are two main types of single-photon detectors frequently used nowadays: single-photon avalanche diodes and superconducting nanowire single-photon detectors. Single-photon avalanche diodes (SPADs) [Cova1996, Cova2004, Migdal2013] are based on the use of a semiconductor avalanche diode that is reverse-biased to create an avalanche breakdown (Geiger mode), see Fig. 2.4. The avalanche triggers the output detection event and, simultaneously, the quenching process to reset the diode before another detection process. SPADs provide very good performance at room temperature, especially for visible and near-infrared light (silicon technology), but mediocre for infrared light, such as the telecom region (InGaAs/InP technology). They reach efficiency as high as 75% for the visible and near-infrared regions (30% for telecom wavelengths) and false detection rate down to 10 counts per second in the visible and near-infrared regions (10^4 cps for telecom). Free-running silicon SPADs typically have a dead time of dozens of nanoseconds and temporal jitter of hundreds of picoseconds, with a possible improvement down to dozens of picoseconds when sacrificing the efficiency. SPADs can be easily integrated in larger arrays due to the use of a common semiconductor technology, with vast applications in imaging [Zappa2007, Bruschini2019, Hadfield2023, Elmalem2025, Ronceray2025].

Superconducting nanowire/nanostrip single-photon detectors (SNSPDs) [Natarajan2012, Hadfield2016, Zadeh2021, Lita2022, Venza2025] are based on the use of a superconducting wire that is cooled to the subcritical temperature, see Fig. 2.5. An absorbed photon breaks Cooper pairs, forming a resistive hotspot where current is diverted, giving a measurable voltage. The detector resets once the hotspot cools. The first SNSPD was presented by Gol'tsman et al. [Gol'tsman2001] and has been massively commercialized in the last decade. SNSPDs can detect single photons with system efficiencies exceeding 95% [Reddy2020] and with noise levels as low as 10^{-5} false counts per second [Shibata2014, Shibata2017, Chiles2022], see Fig. 2.6. They also reach lower temporal jitter down to a few picoseconds (especially when equipped with cryogenic amplifiers) [Korzh2020] and higher repetition rates than SPADs, at the expense of significantly more complicated operation conditions. A significant challenge remains in reaching these performances simultaneously and in the long term without aging-induced decay. Scaling of SNSPD arrays is also an issue [Hadfield2023]; only recently a 500x800 pixel SNSPD camera was developed [Oripov2023].

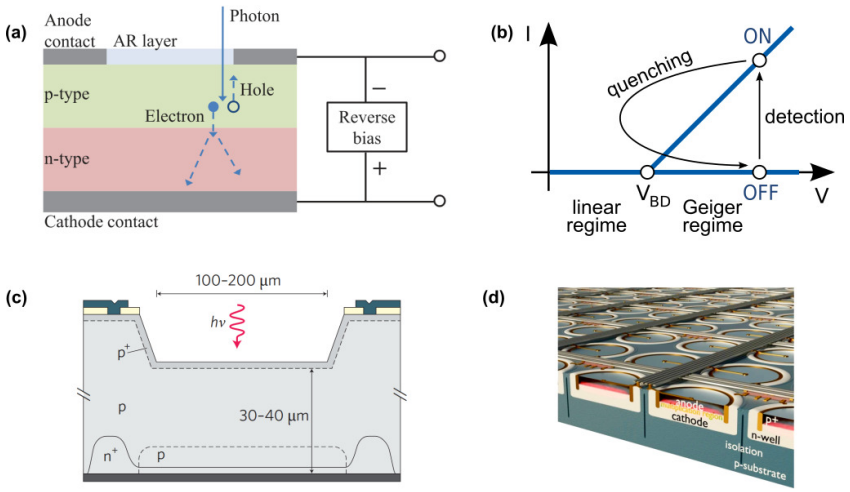


Figure 2.4. (a) A single-photon avalanche diode (SPAD) is shown with distinct regions for the photo-absorption and carrier multiplication processes [Eisaman2011]. The reverse voltage is applied to accelerate the electrons toward the multiplication region, where the avalanche evolves. A front-illuminated geometry with an antireflection (AR) coating to improve efficiency is illustrated, but back-illuminated designs are also used. (b) Bistability of photocurrent is shown when SPAD is operated above the breakdown voltage (V_{BD}), i.e., in Geiger mode. During a quenching process, the voltage is reduced below V_{BD} for a short period of time, quenching the avalanche and resetting the SPAD. (c) A device structure of a thick-junction Si SPAD optimized for high detection efficiency and low dark counts [Cova2004, Hadfield2009]. (d) Artist's impression of a SPAD array with an example of the corresponding cross-section for a substrate isolated SPAD in a conventional CMOS process, depicting some of the key components (diode anode/cathode and corresponding p-n junction, multiplication region in which the avalanche is triggered, the substrate, and isolation between individual SPADs) [Bruschini2019].

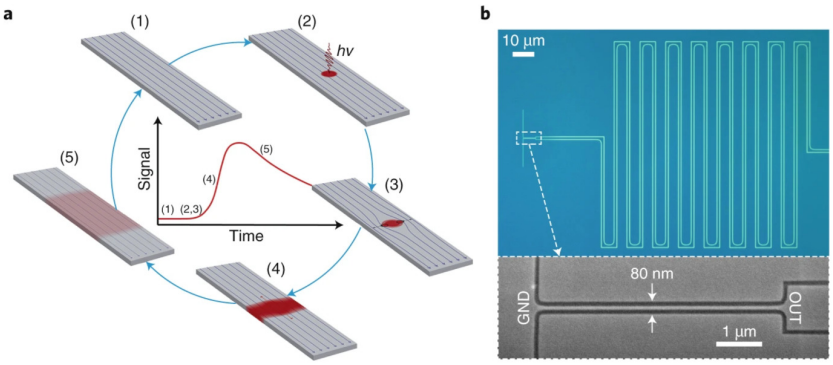


Figure 2.5. (a) Schematic showing the generation of a transient hotspot in a nanowire following the absorption of a single photon: (1) supercurrent flows in the nanowire; (2) a photon is absorbed by a single electron that relaxes, creating a hotspot of excited quasiparticles; (3) vortex motion breaks superconductivity across the nanowire; (4) the normal domain grows along the length of the nanowire leading to a large resistance; (5) as current is diverted, the nanowire relaxes back into the superconducting state. (b) Microscope images of a fast SNSPD [Korzh2020], showing the nanowire path and geometry. The nanowire is embedded in a coplanar waveguide structure; GND is the electrical ground and OUT is the bias and output connection to the nanowire. Image courtesy of R. H. Hadfield, Nat. Photonics 14, 201 (2020).

The main disadvantage of single-photon detectors is inability to distinguish the number of photons in a given light pulse. Despite the widespread effort in reaching at least partial photon-number resolution with SPADs [Kardynał2008, Thomas2010, Thomas2012] and SNSPDs [Cahall2017, Los2024, Schapeler2024, Endo2025], single-photon detectors are still considered as binary (on-off) detectors due to their bistable operation. In the context of this thesis, the work Schapeler et al. [Schapeler2024] stands out; it used a standard SNSPD hardware complemented with optimal detector readout based on dimensionality reduction via principal component analysis. While reaching discrimination only between one- and two-photon events and very poor resolution up to five photons, it represents an example of advanced data processing (unsupervised learning in this case) to gain more information from raw detection signals.

Photon-number-resolving detectors target the number of photons in an optical signal. There are several different types of photon-number-resolving detectors, including cryogenic transition-edge sensors (TES), multiplexed single-photon detectors, and multi-pixel silicon photomultipliers. TES detectors [Irwin2005, Hadfield2016, Lita2022] are based on the use of a thermally-isolated superconducting film that is cooled to a very low temperature ≤100 mK. The film, biased on the steep edge of resistance curve, is heated by absorption of photons so the resistance rises sharply, converting the absorbed energy to a μA-level current in a superconducting quantum interference device (SQUID) readout. TESs reach efficiency > 95% for a broad wavelength range. Their dynamic range of photon-number resolution is limited to approximately 30 photons by the finite inherent energy resolution (overlap between energy detection channels). Also, they are very slow (a few 100s of kHz) with significant time jitter (dozens of nanoseconds), which negatively impacts their application in communications and quantum technology. Recently, TES operation rate was extended to 800 kHz using machine learning approaches, namely supervised and unsupervised classification and clustering [Li2025].

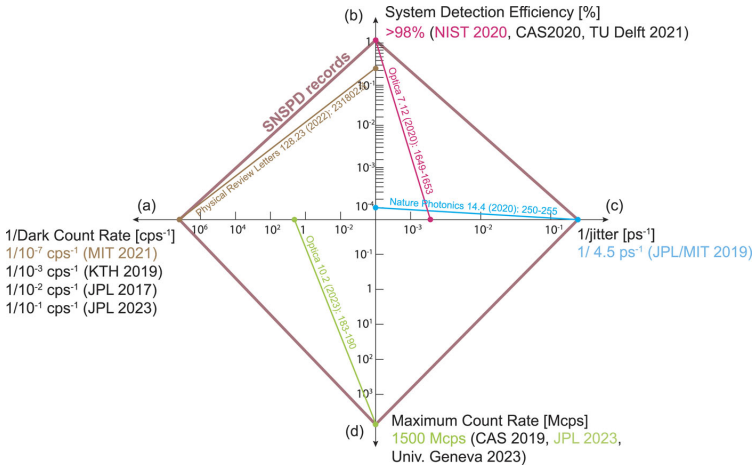


Figure 2.6. Performance records from experimental SNSPD demonstrations on four characteristic metrics: dark count rate (a), system detection efficiency (b), jitter (c), and maximum count rate (d). Axes (a) and (c) display the reciprocal of the metric described following the way the metrics appear in the figure of merit introduced by Hadfield [Hadfield2009]. The inner lines trace the trade-off between two metrics. Achieving a record on one metric typically implies unbalancing a trade-off, leading to a degraded coupled metric. The text on the inner lines indicates the manuscript where the experimental record is presented. Image and caption by Venza and Colangelo [Venza2025].

Multiplexed photon-number-resolving detectors operate by distributing an incident multiphoton optical signal across a sufficiently large number of single-photon detectors. The multiplexing ensures that each detector receives only a small fraction of the total signal, such that the probability of multiple photons impinging on the same single-photon detector is minimized. By registering simultaneous detection events (coincidences) among different detectors, the scheme infers the presence and number of multiple photons in the original optical signal. The multiplexing approach to PNRD was for the first time proposed by Paul, et al. [Paul1996, Sperling2012]. Implementations include temporal [Banaszek2003, Achilles2003, Fitch2003, Řeháček2003, Tiedau2019] and spatial [Hloušek2019, Hloušek2024a] multiplexing with bulk on–off detectors, integrated on–off detector arrays [Divochiy2008, Allevi2010a, Kalashnikov2011, Mattioli2016, Kröger2017, Zhu2018, Cheng2023] (Fig. 2.7), or configurations combining a few detectors with inherent photon-number resolution [Calkins2013, Harder2016, Eaton2023]. Multiplexed PNRDs with bulk detectors excel in low crosstalk and offer limited scalability to higher number of high-efficient detection channels [Hloušek2019, Hloušek2022, Hloušek2024a, Večerka2025]; see also Chapter 4. Temporal multiplexing can reduce the number of physical detectors needed, but it increases the number of effective detection channels at the cost of reduced measurement speed. Minimizing losses and achieving balanced performance in temporal multiplexers require careful optimization [Mičuda2008] or even active signal switching [Tiedau2019, Švarc2019, Mazin2025]. Using a single on-off detector for optical signals significantly longer than the dead time of the detector represents an extreme case typically applicable to continuous-wave signals or proof-of-concept detectors [Straka2018b, Banner2024]. Alternatively, instead of temporal multiplexing, a single on-off detector can be "multiplexed" or modulated in detection efficiency providing a larger number of effective detection outputs, allowing for photon statistics retrieval [Mogilevtsev1998, Rossia2005, Zambra2005].

Integrated multi-pixel PNRDs, also termed silicon photomultipliers (SiPM) [Acerbi2019, Gundacker2020], consist of an array of silicon SPADs, similar to imaging SPAD arrays but without individual pixel readout. They can easily have thousands of detection channels but suffer from a significant crosstalk between individual pixels, relatively low fill-factor, and limited efficiency, particularly in near-infrared and infrared regions. Moreover, the integrated multi-pixel PNRDs have limited or no reconfigurability, making channel balancing virtually impossible. Consequently, these detectors require detailed characterization and sophisticated numerical postprocessing to correct for the imperfections [Kalashnikov2011, Kröger2017]. A promising alternative is the on-chip integration of independent on–off cryogenic detectors like SNSPDs [Mattioli2016, Zhu2018, Cheng2023], though their performance for diverse classical and especially nonclassical light sources remains to be demonstrated. Recently, a PNRD consisting of integrated 100 SNSPDs was presented by [Cheng2023], however, showing significant discrepancy between measured and expected data even for classical weak coherent and thermal states, see Fig. 2.7.

The various photon-number-resolving detection techniques can be combined, for example by multiplexing TES detectors or using both time and spatial multiplexing, to improve their overall dynamic range [Calkins2013, Harder2016, Eaton2023]. Photon-number-resolved measurements currently reach about 100 photons, mainly

for classical states [Eaton2023, Cheng2023, Hloušek2024a]. Further research is needed to go beyond this limit and approach the performance of true linear detector with photon-number resolution.

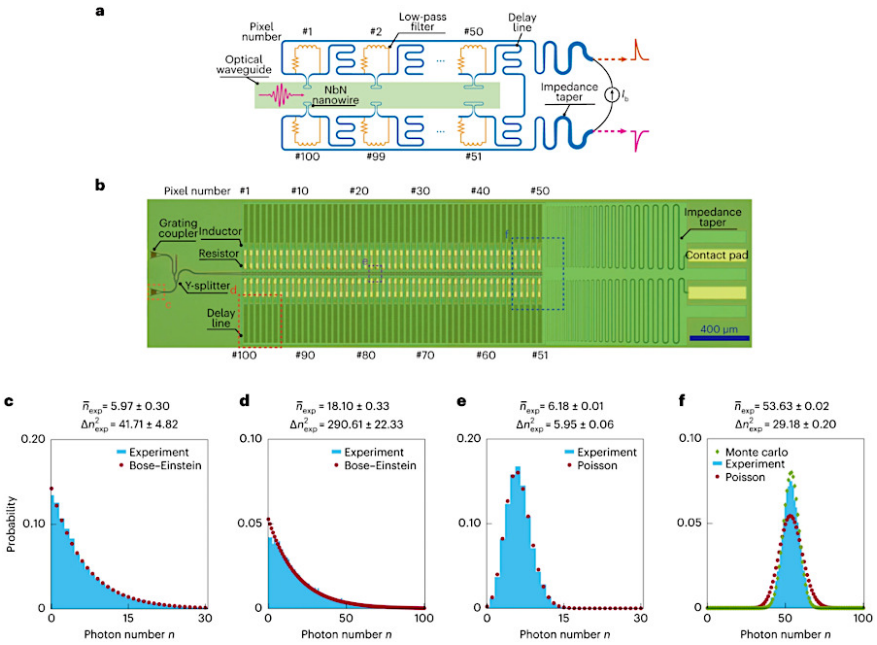


Figure 2.7. (a) Schematic illustration of the 100-pixel quasi-PNR detector structure based on the spatiotemporal-multiplexed SNSPD array [Cheng2023]. The high-frequency detector signals are read out through the bus line consisting of the series-connected nanowire detectors and delay-line sections (blue color), whereas the bias currents of the fired detector pixels are recovered via the low-frequency reset loop formed by the on-chip inductor and resistor (orange color). (b) Overview optical micrograph image of the device. (c,d) Photon number probability distribution for thermal light. (e,f) Photon number probability distribution for coherent light. The Bose–Einstein and Poisson fits deviate from the experimental data, significantly in (f), due to the detector saturation effect. Monte Carlo simulation taking into account this effect (green diamond markers) decreases the deviation partially. Image and caption adapted from Chang, et al. [Cheng2023].

In addition to single-photon and photon-number-resolving detectors, there are also coherent detection techniques that are used to measure continuous variables, such as the quadrature components of a light field. These techniques include homodyne detection, in which the light field is mixed with a local oscillator and the resulting strong fields are measured with two semiconductor photodiodes. The difference between their photocurrents is proportional to the quadrature component of the measured optical signal. Homodyne detection is a key technique for utilization of squeezed states of light [Breitenbach1997, Furusawa1998, Taylor2013, Li2018, Tse2019, Acernese2019]. For fully time-resolved detection of a pulsed signal, the homodyne detector readout is done individually for each pulse [Smithey1993, Hansen2001, Haderka2009, Kumar2012]. This method is particularly useful for the complete characterization of highly nonclassical quantum states, such as those generated by nonlinear processes excited by ultrafast lasers [Lvovsky2001, Zavatta2004, Ourjoumtsev2006, Lvovsky2009, Lvovsky2020, Fadrný2024]. The major challenge in quantum homodyne detection is achieving sufficiently high signal-to-noise ratio and precise balancing. These requirements are hard to fulfill especially for large bandwidths [Kawasaki2025]. Furthermore, the local oscillator is

often not available for light emitted from single-emitter sources, such as individual atoms, molecules, or quantum dots, and the only options are self-homodyning or iSCAT techniques [Wrigge2008, Špačková2022].

2.3. Statistical properties of light

Detection statistics of the output detection channels of a photon-number-resolving detector does not coincide exactly with photon statistics of the incident light. The main reason is a finite number of detection channels, i.e., limited dynamic range. Other causes include less-than-unity detection efficiency, noise, and channel crosstalk. With sufficiently large detection statistics and accurate calibration of the detector, the actual photon statistics can be retrieved to some degree of confidence [Fiurášek2025]. There is a whole research subfield targeting the efficient photon-statistics retrieval that covers direct inverse methods [Tikhonov1995, Starkov2009] as well as statistical methods [Dempster1977, Vardi1993, Banaszek1998, Mogilevtsev1998, Řeháček2003, Rossia2005, Zambra2005, Marsili2009, Hloušek2019]. As a state-of-the-art method, we can mention expectation-maximization algorithm with weak entropic regularization developed by our team [Hloušek2019], allowing for highly accurate photon-statistics readout from spatially multiplexed [Hloušek2019, Hloušek2024a] and time-multiplexed [Banner2024] detectors, see Chapter 5.

Despite this progress, retrieving the full photon statistics is an inherently ill-posed inverse problem, involving open questions on optimal regularization and bias-variance trade-off. Instead, only partial information can be targeted, such as global statistical parameters. The statistical feature extraction includes moments, nonclassical and quantum non-Gaussian characteristics, and classification of various classes of light. For example, we experimentally verified new methods of nonclassical multi-photon state classification [Straka2018a, Lachman2019]. These methods represent initial steps toward accurate sensing of multiple single-photon emitters, a task currently dominated by multiparameter estimation methods [Li2023]. However, the resilience of the reported approaches to losses, background noise, and different emission probabilities of the emitters, has to be improved. Recent developments in machine learning have enabled classification between two statistics [You2020], but the full-scale applications of artificial intelligence methods for photon statistics retrieval and statistical feature extraction have yet to be demonstrated.

2.4. Detector response and calibration

Another major challenge is the complexity of photonic detector operation. Detector responses are nonlinear and lack precise theoretical modeling. The performance of single-photon detectors is affected by many low-level processes participating in the conversion of a single-photon signal to a macroscopic electronic output. The particular detection imperfections and their modeling are specific to SPADs and their quenching circuits [Stipčević2013, Kornilov2014, Wang2016a, Straka2020] or SNSPDs [Marsili2016, Akhlaghi2009, Akhlaghi2011] due to their different operational principles. These imperfections cause a nonlinear relation between the

incident photon rate and observed detection rate. Our team recently developed a direct single-source measurement of the absolute nonlinearity of single-photon detectors with unprecedented accuracy [Hloušek2023]. We discovered a surprising supralinear behavior of SPADs and showed that it cannot be explained using known theoretical models. We also characterized sub- and supralinear operation regimes of SNSPDs and uncovered the supralinearity under faint continuous illumination. The results identify new detector anomalies that supersede existing knowledge of nonlinear effects at the single-photon level, which underlines the need to characterize the response of single-photon detectors by direct measurement, see Chapter 4. Another level of complexity comes from the fact that single-photon and photon-number-resolving detectors are nonclassical, i.e., they perform projections onto strongly nonclassical quantum states. The nonclassicality characterization is usually performed using an indirect, tedious, and poorly scalable process of quantum detector tomography [Lundeen2009, Zhang2012, Schapeler2025]. Instead, we developed a direct method for detector nonclassicality certification requiring the minimum resources [Hloušek2021, Grygar2022].

There is a rapidly increasing demand for accurate photonic detection and, particularly, photon counting in communication, metrology, and imaging – from benchmarking of optical and quantum sources to optical receivers, from counting and imaging of single emitters to computed tomography. In short, the field urgently requires detectors with large dynamic range, their accurate calibration, and robust retrieval techniques resistant to detector imperfections.

2.5. Vector properties of light

Polarization of light plays a central role across many scientific disciplines, such as imaging [Jameson2010, Qi2017, Rimoli2022], sensing [Zhan2021], ellipsometry [Liu2021, Wen2024], optical communications [Wang2016b, Chen2016], topological photonic devices [Abbaszadeh2021], and quantum information processing [Flamini2019, Slussarenko2019]. In particular, the polarization state of light emitted by anisotropic specimens encodes valuable information about material structure and optical properties. This principle has driven a broad range of biomedical applications, where polarization provides structural and morphological insights into biological tissues and enables label-free detection of pathological changes [Qi2017, Brasselet2019]. Polarization-resolved endoscopy and imaging techniques have become particularly useful for detecting, classifying, and staging cancerous lesions in organs such as the stomach, colon, prostate, and larynx [Doradla2013, Qi2023].

In the quantum regime, polarization represents an important information carrier for quantum communications and photonic quantum processors and simulators. Photonic quantum bits (qubits) are commonly encoded in polarization states of a single photon [Flamini2019, Slussarenko2019]. Nonclassical correlations, i.e., quantum entanglement, between polarization states of individual photons represent an essential resource for quantum information processing and quantum metrology, see Section 2.6. Almost all seminal works in quantum information processing utilized polarization of light for carrying or processing information: quantum teleportation [Bouwmeester1997], entanglement swapping [Pan1998], quantum key distribution

[Ursin2007, Liao2017, Yin2017], quantum logic gates [O'Brien2003], one-way quantum computing [Walther2005], and quantum simulators [Lu2009, Schreiber2011, Aspuru-Guzik2012, Cardano2016, Pitsios2017, WangK2019]. This broad range of applications, from medicine to quantum science and technology, collectively motivates accurate preparation and rapid, reliable measurement of polarization across power levels and operating environments.

Polarization can be manipulated by mechanical retarders, electro-optic modulators (free-space or integrated), or liquid crystal devices. Mechanical birefringent components achieve high accuracy but are slow and vibration-prone. Free-space Pockels cells provide high-speed operation and are widely used in switching, routing, and quantum applications [Spagnolo2008, He2017, Takeda2019], but their large size and requirement of high-voltage drivers hinder integration. Integrated photonic polarization controllers enable ultrafast reconfiguration [Altepeter2011], though at the cost of insertion losses and challenging calibration due to waveguide coupling.

Liquid crystals represent a practical compromise for polarization manipulation: they allow low-voltage operation, stability without moving parts, and compatibility with compact free-space and imaging setups. Parallel-aligned nematic liquid crystal cells are commonly used as voltage-tunable retarders for polarization preparation [Adamson2007, Peters2005], modulation [Zhuang1999, Defienne2021, Lohrmann2019, Villar2020], measurement [Bueno2000, DeMartino2003, Peinado2010, Adamson2010], and implementation of quantum channels [Orieux2013]. However, accurate modeling of liquid crystal response remains an open problem due to alignment imperfections, depolarization, and nonlinear voltage dependence [Yariv1984, Coy1996, Marquez2000, Marquez2001, Yamauchi2005]. The modeling challenge is particularly severe for twisted nematic liquid crystals (TNLCs), widely used in display technology, whose complex action combines retardation and rotation. Recently, we tackled this challenge and developed two approaches for precise calibration of TNLCs for high-fidelity polarization (and polarization qubit) preparation, manipulation, and measurement. The core experimental technique is based on a sequence of twisted nematic liquid crystals calibrated and controlled by an evolutionary (genetic) optimization [Bielak2021] or bidirectional learning based on artificial neural networks [Vašinka2022]. These methods are now routinely used for controlling the polarization state of light in many other research projects, see Chapter 7.

Optical multiplexers and networks can be used for polarization measurement, similar to photon statistics detection. Several ordered polarization projections are performed, spatially multiplexed or in a time sequence, using birefringent networks [Ling2006], TNLCs [Bielak2021], or metasurfaces [WangK2018, Rubin2022]. The individual projections need to be carefully calibrated. Alternatively, polarimetry with disordered networks has recently been demonstrated, see below. Complete characterization of quantum states and devices, including those utilizing polarization of light, relies on complex detection protocols, such as quantum state tomography [Vogel1989, Ježek2003, Hradil2004, Paris2004]. The tomography consists of many individual projections or detection channels (quorum), which have to be accurately known, from which the full quantum state is reconstructed. If the experimentally implemented measurement projections deviate from the expected ones, which is always the case, the accuracy of the tomographic characterization is compromised. The only solution

suggested so far to calibrate the constituent measurements is quantum measurement tomography. Unfortunately, the measurement tomography requires accurately known probe states, which is also not feasible in real-world applications. This proverbial chicken-or-egg scenario in quantum tomography has remained unsolved until our recent result on measurement-device-agnostic quantum tomography [Stárek2025].

The number of required projections and computational complexity of quantum tomography scale exponentially with a number of qubits in the characterized system [Häffner2005, Barreiro2005, Monz2011, Song2017, WangX2018]. Because of this unfavorable scaling, the largest system characterized experimentally by full quantum state tomography to date contains only ten qubits [Song2017]. Several methods have been developed to make this scaling more favorable, by imposing an a priori structure or symmetry to the system [Gross2010, Cramer2010, Tóth2010, Lanyon2017, Torlai2018, Tiunov2020, Danaci2021]. Furthermore, optimal tomographic approaches often require impractical entangled measurements. The detailed accuracy analysis of local measurement schemes is missing, as are their experimental demonstrations for a large number of detection channels. The ultimate accuracy limit has not yet been reached. Poor scalability of the full quantum tomography may be addressed by directly detecting polarization and correlation features like depolarization, purity, coherence, and various entanglement quantifiers.

While polarization measurements can be relatively straightforward in controlled laboratory conditions, accurate and reliable characterization under realistic, complex scenarios remains highly challenging. Low light intensities represent the first major difficulty: reduced photon numbers lead to poor signal-to-noise ratios, making it difficult to extract meaningful polarization information from noisy data. A second challenge arises in high-speed measurements, where short acquisition times reduce the number of collected photons. Increasing the illumination power could mitigate this, but it is constrained by delivery limitations and, critically, by the photodamage threshold of sensitive samples. Incorporating single-photon detectors improves accuracy in polarization-state reconstruction protocols, yet their relatively large size and lack of in-situ or even in-vivo integrability pose significant limitations. A third challenge emerges in constrained spaces, where compact setups are essential but alignment and placement of optical components are severely restricted. These combined conditions present a serious obstacle, especially for invasive medical procedures such as microendoscopy [Turtaev2018, Wen2023], environmental monitoring [Valentino2022, Běhal2022], and internal or in-situ materials inspection [Deng2021, Coppola2024].

To overcome these limitations, one promising direction is to exploit light propagation in disordered media for polarization sensing. When coherent light traverses a strongly scattering medium, multimode interference generates a granular speckle pattern. Although seemingly degrading, this speckle pattern encodes detailed information about the input field [Rotter2017, Bertolotti2022, Cao2022, Vynck2023, Cao2023]. By analyzing speckle structures with suitable post-processing techniques, one can retrieve information about the polarization state of the incident light [Juhl2020, Facchin2020]. Although considerable progress has been made in advancing disordered-media-based optical methods [Pierangeli2023, Stibůrek2023, Goel2024], the development of a compact polarimeter capable of fast, reliable, and single-photon-sensitive operation remained an open challenge until recently [Bielak2025].

We employed a short piece of a few-mode fiber that maps the polarization state to a random scattering pattern, which is sparsely sampled with a fiber array to leverage fast single-photon detection. This all-fiber design greatly enhances detection speed, sensitivity, and data throughput. Crucially, it can enable simultaneous detection of multiple optical parameters, potentially including quantum correlations and entanglement.

2.6. Nonclassical correlations and entanglement

Quantum entanglement manifests as nonclassical correlations (stronger than any classical ones) between measurements performed on individual subsystems of a larger physical system [Einstein1935, Bell1964]. These effects are central to fundamental tests of quantum mechanics, including direct experimental verification of quantum nonlocality [Hensen2015, Shalm2015, Giustina2015]. The critical role of entanglement has also been demonstrated at the opposite end of the complexity spectrum, in macroscopic phase transitions [Osterloh2002, Osborne2002, Amico2008, Bohnet2016, Khemani2017, Valdez2017, Kokail2021, Joshi2023, Karamlou2024]. Beyond foundational aspects, entanglement is an essential resource for quantum information processing and enables quantum advantage [Asavanant2019, Larsen2019, Arute2019, Larsen2021, Madsen2022, Evered2023, Google2023]. Modern quantum communication networks also rely crucially on entanglement sources [Wengerowsky2018, Yin2020, Guccione2020, BassoBasset2021]. Consequently, characterizing entanglement is paramount for both fundamental research and quantum technology applications [Horodecki2009, Erhard2020, Eisert2020].

Reliable entanglement quantification remains an open problem in quantum physics. Direct measurements of entanglement can be realized by exploiting quantum interference of two (or more) identical copies of a system, see Fig. 2.8 (a) [Horodecki2003, Fiurášek2004, Walborn2006, Islam2015, Kaufman2016]. This nonlocal multi-copy strategy is rooted in measuring nonlinear functionals of quantum states [Filip2002, Ekert2002]. However, such measurements are experimentally very demanding, which has motivated single-copy approaches based solely on local measurements, such as quantum tomography.

Quantum tomography provides a full description of a quantum state, including its entanglement, see Fig. 2.8 (b) [Paris2004]. However, the number of measurements required for tomography grows exponentially with the number of qubits or degrees of freedom, making the approach inherently unscalable [Häffner2005, Barreiro2005, Monz2011, Song2017, WangX2018]. When a few-parameter state model is assumed, quantum estimation can be used to optimally infer entanglement [Brida2010a, Brida2011, Benedetti2013]. Another route to emulate quantum correlations with fewer resources uses neural-network quantum states [Carleo2017, Deng2017, Torlai2018, Carleo2018, Hartmann2019, Reh2021]. This approach, however, suffers from the sign problem, whose resolution requires additional assumptions [Szabó2020, Westerhout2020]. Neural-network quantum-state methods have also been applied to tomography under nonideal experimental conditions [Torlai2018, Torlai2019, Tiunov2020, Palmieri2020, Danaci2021]. Still, the amount of

information needed to represent a generic quantum state to a given accuracy remains an open question [Rocchetto2019, Gebhart2023, Du2025].

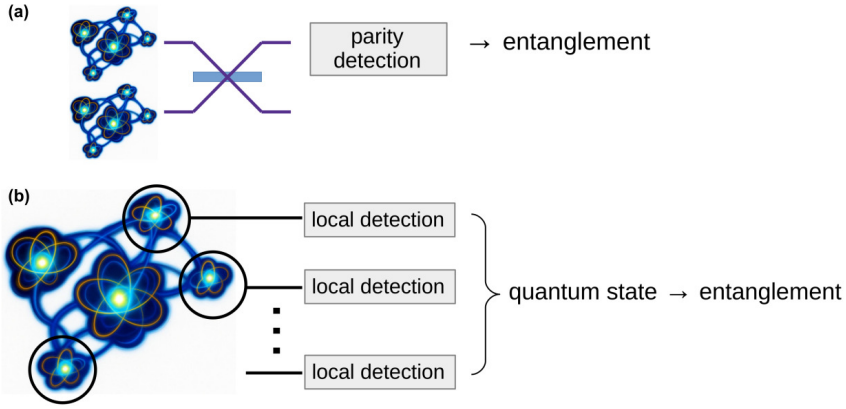


Figure 2.8. (a) Principle of a nonlocal two-copy strategy to directly measure quantum entanglement. (b) Indirect characterization of entanglement via quantum state tomography based on local measurements.

Instead of characterizing the entire system, one can target only mean values of selected observables, substantially reducing the required number of measurements. This strategy, known as shadow tomography [Aaronson2018], can also be used to estimate the entanglement entropy of a small subsystem by effectively reconstructing its reduced state [Huang2020, Struchalin2021]. An alternative approach estimates the second-order Rényi entropy of a subsystem from random measurements [vanEnk2012, Elben2018, Brydges2019, Elben2023]. However, these methods do not quantify entanglement distributed across the entire system.

Entanglement witnessing is a viable alternative to tomographic methods when the goal is only to distinguish entangled from nonentangled states (or between entanglement classes) without quantifying the degree or structure of entanglement. Nevertheless, witnessing may still require full knowledge of the underlying quantum state, as in the positive partial transpose criterion [Horodecki2009]. A witness cannot be measured directly, but it can be approximated via a completely positive map [Horodecki2002], which is essentially equivalent to full quantum state tomography [Fiurášek2002, Carmeli2016]. Other witnessing strategies use minimal local decompositions [Guhne2002, Barbieri2003], semidefinite programming [Audenaert2006, Jungnitsch2011], entanglement polytopes [Walter2013], or correlations from randomized measurements [Ketterer2019, Knips2020]. Entanglement witnessing can also be assisted by neural-network classifiers [Gao2018, Harney2020, Roik2021, Du2025]. Despite these advances, witnessing typically provides only witnesses or lower bounds and often relies on prior information about the state.

In summary, the relation between the entanglement present in a physical system and the correlations accessible through measurements on its subsystems is highly nontrivial [Kaszlikowski2008, Mičuda2019]. Full entanglement characterization using single-copy local measurements appears achievable only via complete quantum state tomography and therefore entails an exponential scaling of the required measurements [Fiurášek2002, Carmeli2016, Lu2016, Yu2020].

Our team recently demonstrated deep-learning-based direct quantification of polarization correlations and entanglement from incomplete local measurements [Koutný2023]. We focused on all pairwise correlations quantified by concurrence [Horodecki2009] and the adjacency matrix of quantum mutual information [Valdez2017]. We successfully tested this approach for photonic sources based on parametric generators and quantum dots, outperforming state-of-the-art methods in accuracy and speed. See Chapter 6 for details. However, the scaling to systems with a high number of qubits, genuine multipartite entanglement, measurement cost of entanglement quantification, and measurement-device independence are still missing. In summary, major open challenges remain for polarization and correlation sensing in demanding scenarios and quantum-limit conditions, with quantum entanglement detection as the ultimate case.

2.7. Single emitters and nanosensors

Although single-photon detection, optical networks, and photonic correlations can already boost precision on their own, coupling photons to single quantum emitters, such as single atoms, ions, or solid-state "artificial atoms", adds matter-based degrees of freedom that can sense external fields or store quantum states, with their own fluorescence offering a low-noise readout for metrology, sensing, and imaging. Neutral atoms and charged ions confined in optical or electromagnetic traps are exemplary single-photon emitters whose optical transitions reach transform-limited linewidths and long coherence times. Realizing this performance demands ultra-high-vacuum chambers, trapping fields, many lasers for cooling and control, and often also cryogenic cooling. Although unrivaled for time metrology [Marshall2025] and quantum information [vanLeent2022, Mose2023, Guo2024, Main2025], the trapped-ion implementation complexity limits their use in sensing and imaging, particularly in materials and live systems.

Single emitters embedded in solid-state media, such as color centers in crystals, quantum dots, or organic molecules, are promising building blocks in modern photonics. These systems have a wide range of applications, from single-photon sources for secure communications [Eisaman2011, BassoBasset2021, Toninelli2021] to markers in biomedical imaging and sensing [Sahl2017, Barry2020, Lelek2021, Amawi2024, Roesel2025]. Solid-state single emitters offer scalability, vacuum-free operation, and even room-temperature operation. Their energy level separation and other parameters can be controlled by the manufacturing process and external stimuli, such as light, electric and magnetic fields, mechanical strain, or the presence of other nanoparticles.

Organic dye molecules represent an important example of solid-state single emitters, which are bright single-photon sources and possess non-trivial interaction with light and external stimuli [Toninelli2021, Roesel2025]. They can be integrated into photonic chips [Ren2022, Rattenbacher2023] and prepared even without clean-room facilities. However, the fine-structure spectral features of organic molecules are hidden at room temperature due to vibrational state transitions and coupling to the environment. As a result, spectrally sensitive measurement schemes are limited to

cryogenic conditions, which are experimentally demanding and restrictive in the context of sensors, especially for sensing in living organisms. Using single molecules as nanosensors operating at room temperature represents a novel and challenging research direction with broad application potential. The measured photoemission rate of a single molecule was shown to witness the presence of oxygen [Naito2006], determine local acoustic vibrations in nanomechanical devices [Xie2022], and the transport effects of a single electron in a protein using the method of single-molecule fluorescence (Förster) resonance energy transfer (FRET) [Lerner2018], see Fig. 2.9.

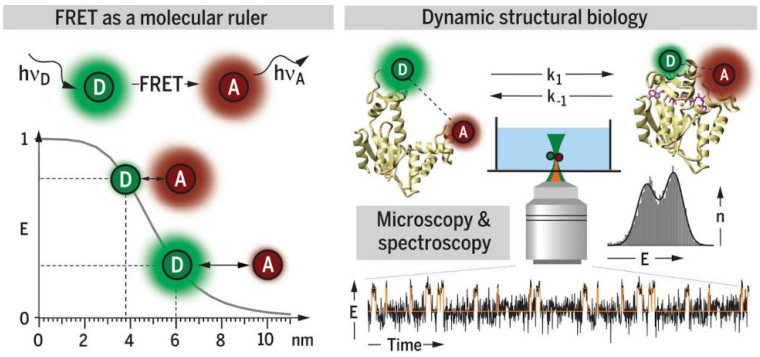


Figure 2.9. Dynamic structural biology using single-molecule FRET. Left: Principle of FRET as a molecular ruler. In a system with a pair of dyes, after the donor dye (D) is excited, it transfers the excitation energy to a nearby acceptor dye (A; top) with an efficiency (E) that depends on the sixth power of the distance between the dyes (bottom). Right: Use of FRET to study structural dynamics at the single-macromolecule level. The experimental setup (top), a combination of single-molecule fluorescence microscopy and spectroscopy, can be used to determine conformational states or dynamics in solution or on immobilized molecules. Here E is calculated per each single-molecule burst of photons, and bursts (n) are accumulated in E histograms (middle) or for different time bins to form a single-molecule E trajectory (bottom). Figure and caption by E. Lerner, et al. [Lerner2018].

The most advanced application of organic molecules currently lies in imaging. Single-molecule localization microscopy is a workhorse of today's super-resolution microscopy [Sahl2017, Lelek2021]. In stochastic optical reconstruction microscopy (STORM) [Rust2006], photoactivated localization microscopy (PALM) [Betzig2006], and MINFLUX microscopy [Balzarotti2017], fluorophores are driven to blink stochastically so that, in any given camera frame, only a sparse subset is active. PALM uses photoactivatable fluorescent proteins that are switched on in small batches, imaged, and then irreversibly bleached. STORM uses reversibly photoswitchable organic dyes and chemical buffers to keep most emitters dark while only a few are temporarily switched on. Each visible emitter produces a diffraction-limited spot whose center can be fitted (typically with a Gaussian point-spread function (PSF) model using maximum likelihood or least squares) to estimate its position with precision that scales roughly as $\sigma/\sqrt{N}$ (σ: PSF width, N: detected photons, with additional terms from pixelation and background). Repeating this activation–imaging–localization sequence over many frames builds a dense set of molecular coordinates; rendering those coordinates yields a super-resolved image that surpasses the diffraction limit because emitters are separated in time rather than in space, see Fig. 2.10. Stochastic/photoactivated switching can be combined with other super-resolving techniques, such as light-sheet microscopy, to improve the resolution even further [Chen2022, Bodén2024]. Instead of emitter fitting, machine

learning has been successfully used for emitter reconstruction and localization [Nehme2018, Speiser2021] with various pros and cons [Belthangady2019].

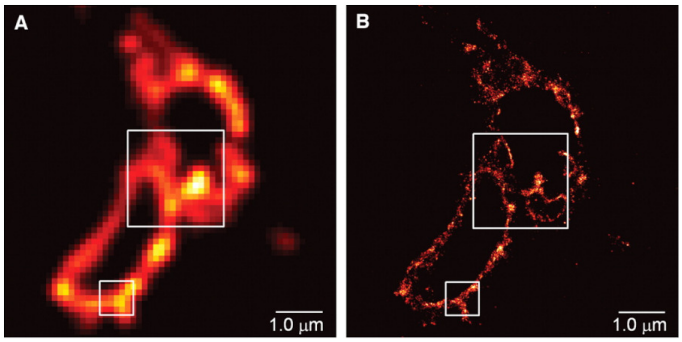


Figure 2.10. Comparative summed-molecule TIRF (A) and PALM (B) images of the same region within a cryo-prepared thin section from a COS-7 cell expressing the lysosomal transmembrane protein CD63 tagged with the PA-FP Kaede. Figure and caption by E. Betzig, et al. [Betzig2006]. Structures well below 100 nm can be clearly resolved using PALM.

However, these methods have practical limits. They are slow because hundreds, if not thousands, of individual frames are required to collect enough non-overlapping localizations, which hinders live-cell and high-throughput studies. High labeling density increases spot overlap and degrades localization unless activation duty cycles are kept extremely low, further increasing acquisition time. Accurate, unbiased localization also depends on careful calibration of the optical system and detector: PSF width and shape (including aberrations and 3D astigmatism), pixel size and magnification, camera gain and read noise, and even priors about the sample such as emitter density all affect estimates and can introduce bias if misspecified [Belthangady2019]. It would be desirable to achieve comparable resolution from a single frame even for dense samples, and to have a device-agnostic super-resolution approach that does not rely on extensive, microscope-specific calibration. We have recently solved these issues and developed a single-frame device-agnostic super-resolution imaging aided by deep learning [Vašinka2025a, Dostálová2025a, Vašinka2025b], see Chapter 7.

The common detection strategies in single-molecule sensing rely on on–off counting using a single-photon detector [Lerner2018, Xu2019]. Photon-counting histogram and number-and-brightness methods employ the same detection hardware but analyze the long-term evolution of the emission rate and its two lowest moments [Digman2008]. The spectroscopy of fluorescence lifetimes and photon-correlation measurements can reveal dynamics at a single protein level and lay a basis for modern microscopy methods [Datta2020]. Higher sensitivity and dynamic range are expected from using advanced multi-detector schemes and evaluating photon-statistics features, as showcased in emitter mapping, localization, and counting [Ta2015, Israel2017, Straka2018a, Qi2018, Fishman2023, Li2023, Vasquez-Lozano2024, LiS2024, Dostálová2025b]. However, very little has been done experimentally in this direction.

Our team uses single terrylene and pentacene molecules self-aligned in thin crystals of para-terphenyl or anthracene, depending on the preferred molecule dipole orientation. The density of the molecules can be set by their concentration in the

solvent prior to spin-coating. Alternatively, other organic dye molecules (DBT, DBATT), colloidal quantum dots, and hexagonal boron nitride defects are considered as potential sensors at room temperature. Especially DBT represents a promising alternative to terrylene as its emission lies close to 0.8 μm, where the absorption of biological tissues is minimal, ultrafast titan-sapphire lasers perform best, and single-photon detectors reach the highest efficiency. Currently used terrylene and pentacene are both compatible with our custom inverted fluorescence microscope setup with a 532 nm excitation laser. The microscope allows for wide field and confocal imaging, and the possibility of total-internal-reflection illumination. It is equipped with effective collection (numerical aperture of 1.4) and common detection techniques, such as spectral analysis, scientific cameras, and Hanbury–Brown-Twiss autocorrelation measurement. The microscope is also augmented with a 8-channel photon-number-resolving detector to retrieve the full photon statistics even from large molecular clusters on various time scales. Photoinduced oxidative bleaching of organic molecules might represent a problem for long acquisition times, but can be mitigated by introducing additional polymer layers or actively exploited for sensing. The key statistical features and the full photon statistics are imaged across a sample by means of its scanning.

2.8. Artificial intelligence: Optimal control and readout

Machine learning, a subfield of artificial intelligence (AI), has emerged as a powerful tool for data analysis and modeling with a wide range of applications. There is a long history of transferring ideas between machine learning and physics. In 2024, the Nobel Prize in Physics was jointly awarded to John J. Hopfield and Geoffrey E. Hinton for foundational discoveries and inventions that enable machine learning with artificial neural networks. Hopfield introduced a model of associative memory as a network of binary “neurons” with symmetric coupling weights [Hopfield1982]. Dynamics converge to energy minima of a global energy function, analogous to magnetic spin systems (Ising models). Hinton extended the Hopfield-type ideas into probabilistic models, notably the Boltzmann machine, where neurons are stochastic (e.g., logistic activation) and weights define an energy function [Ackley1985]. Learning in Boltzmann machines is often phrased in terms of minimizing Kullback–Leibler divergence or free energy (in restricted Boltzmann machines [Smolensky1986, Hinton2002]) between model and data distributions, using ideas analogous to thermodynamics, statistical ensembles, and partition functions.

In recent years, machine learning has made remarkable progress in natural language and image processing, complex systems control, and many other applications [LeCun2015, Bommasani2021]. Machine learning tasks are often grouped into regression and classification. Regression predicts a numeric value from given input features, such as a 0–10 quality score of wine based on chemistry measurements [Cortez2009] or penguin body mass using flipper length, bill depth, etc. [Gorman2014]. Classification assigns each input example to a category (class), such as cats vs. dogs classification from images [Elson2007] or mosquito species classification from wingbeat sounds recorded on smartphones to support disease control [Mukundarajan2017]. Both tasks use labeled examples during training, and inputs and outputs can have many components. A related third family is generative

modeling, which learns the structure of the data to produce new, realistic samples similar to those it was trained on.

Figure 2.11. A schematic contrasting traditional algorithmic programming with the data-driven, experience-based approach central to machine learning. Instead of hand-coded rules (program), machine learning models learn abstractions by extracting shared patterns from data. Image by A. Dawid, et al. [Dawid2025].

Data availability shapes the type of learning. In supervised learning, a model (machine learning program) is trained on labeled examples (input–output pairs) to predict outputs or assign categories [LeCun2015], see Fig. 2.11. It is powerful and widely used but often limited by the cost and difficulty of obtaining accurate labels, and it typically needs large, well-curated datasets. In unsupervised learning, models work with unlabeled data to discover structure, such as clusters or lower-dimensional representations [Kramer1991, Hinton2006], see Fig. 2.12 (d); generative methods can even expand the feature space or synthesize new samples. This can reveal patterns when labels are scarce or unknown. Similar to unsupervised approaches, self-supervised learning uses unlabeled data to create supervisory signals via pretext tasks, such as agreement under augmentations [Zbontar2021]. The learned representations are then used for downstream tasks, such as classification, regression, or anomaly detection. In reinforcement learning, there is no fixed dataset at the start. An agent interacts with an environment, receives feedback about success or failure, and uses that feedback to improve its decisions over time, see Fig. 2.12 (e). Other schemes include active learning, which selects the most informative examples to reduce labeling effort, and semi-supervised learning, which combines many unlabeled samples with a small labeled set to boost performance.

To encode a data transformation (from input features to output targets) or learned algorithm (agent behavior), deep neural networks are often used, see Fig. 2.12 (a, b, c). They consist of multiple layers of interconnected artificial neurons and act as highly nonlinear transformations parametrized by a large number of trainable parameters. Such networks can asymptotically approximate an arbitrary continuous function on a compact set [Cybenko1989, Hornik1991, Leshno1993, Barron1993, Pinkus1999, Lu2017]. An artificial neuron [McCulloch1943, Rosenblatt1958] takes several input values, multiplies each by a weight, adds a bias term, and then passes the result through a nonlinear activation function to produce an output, see Fig. 2.12 (a). The activation (e.g., sigmoid, tanh, ReLU) is what lets networks model nonlinear relationships; without it, stacked neurons would collapse to a single linear mapping. In a layer, many neurons operate in parallel on the same inputs to extract different features; stacked layers let the network build up from simple patterns to more complex ones, see Fig. 2.12 (b). The network parameters (weights and biases) are efficiently learned using (stochastic) gradient descent [Robbins1951, Rosenblatt1958, LeCun2015], see Fig. 2.12 (c). The parameters are adjusted to reduce a loss computed on the training data, i.e., information the model can see during learning. The loss is task-dependent and somewhat flexible: for example, mean

squared error or mean absolute error are common in regression, while cross-entropy is typical for classification. The loss can also include regularization terms to discourage overly complex models. In stochastic gradient descent [Bottou2010] and its variants, such as Adam and NAdam [Kingma2015, Dozat2016], parameters are updated using the loss and gradient computed on small, randomly sampled mini-batches at each step, with many such steps making up an epoch. The resulting gradient noise makes update directions fluctuate, which can help the optimizer pass through saddle points and escape sharp, narrow minima [Feng2021], though theory shows that even deterministic gradient descent can still avoid strict saddles and converge to minimizers under mild conditions [Lee2016]. Using mini-batches also cuts the cost of each update and makes training on large datasets feasible on limited hardware [Bottou2010]. Deep neural networks can generalize from the learning stage – once trained, they can perform surprisingly well even for inputs unobserved during the learning stage. The basic principles of deep neural network operation are well known, but the full span of their generalization ability is the subject of current research [Belkin2019, Kawaguchi2022].

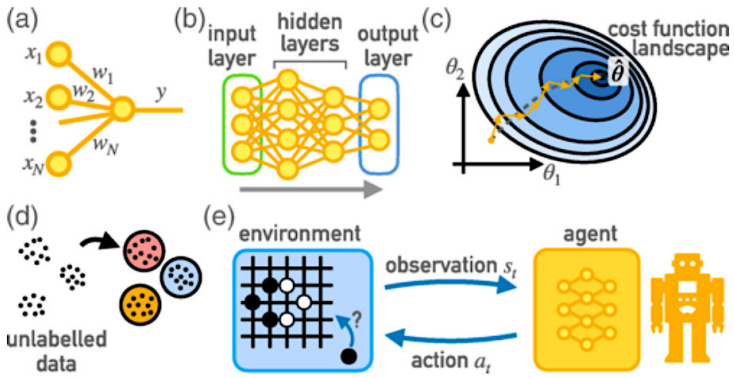


Figure 2.12. Basics of neural networks and machine learning techniques. (a) Operation of a single artificial neuron, where inputs x_i are combined with weights w_i, optionally a bias is added, and the result is passed through an activation function providing the output y. (b) Structure of a neural network with several dense layers illustrating how information flows from input to output through multiple hidden layers. (c) Evolution of the network parameters θ (including weights and biases) across the cost function landscape using stochastic gradient descent. At each update step (orange arrows), the gradient of the average cost with respect to θ (gray dashed line) is approximated by the gradient evaluated on a randomly chosen data batch. The optimized parameters $\hat{\theta}$ correspond to the minimum of the averaged cost function. (d) Example of an unsupervised classification task, where unlabeled input data are assigned to distinct categories. (e) Reinforcement learning scenario, in which an agent interacts with an environment by taking actions, receiving rewards, and improving its policy over time. Image by M. Krenn, et al. [Krenn2023].

In contrast to fully connected networks that do not induce any prior symmetry or structure, there are other deep learning modalities with particular symmetries. Convolutional neural networks [LeCun1998] enforce translation invariance, making them natural for image-like data. Equivariant neural networks [Cohen2016] can encode rotational or permutation invariance, which is crucial in physical systems where such symmetries reduce the effective dimensionality of the learning problem. A particularly important class are graph neural networks [Scarselli2009], which respect permutation invariance of graph nodes. This makes them ideally suited to modeling molecules, crystals, and quantum many-body systems, where the physical

system is naturally represented as a graph with nodes (sites, atoms, spins) and edges (bonds, interactions) [Gilmer2017, Jumper2021, Reiser2022].

By combining two neural networks (encoder and decoder) in sequence or in parallel, various deep learning modalities can be produced with applications in unsupervised or self-supervised learning. A serial arrangement forms an autoencoder [Kramer1991], which learns to compress data into a latent representation and reconstruct it. Variational autoencoders extend this idea by imposing a probabilistic structure on the latent space, enabling smooth sampling and generation of new data. Generative adversarial networks use an adversarial game between a generator and a discriminator to produce highly realistic synthetic data without explicit reconstruction of the input. A pair of similar networks trained in parallel on augmented data to maximize their prediction correlations can reduce information redundancy, such as in SimCLR, BYOL, SimSiam, and Barlow Twins [Chen2020, Grill2020, Chen2021, Zbontar2021]. For applications in natural sciences, however, the augmentation has to comply with the physical properties of the studied systems [Zhu2022, Wang2022, Zimmermann2023].

The combination of AI and photonics has the potential to enable new approaches to imaging and sensing [Rivenson2017, Barbastathis2019, Ballard2021, Qiao2021, Yuan2023, Freire2023]. Machine learning models can extract useful information from complex photonic data and enable new methods for inventing, controlling, and optimizing photonic devices [Ma2021, Rodríguez2024, Krenn2025]. In the opposite direction, photonics can readily implement many machine learning architectures toward neuromorphic hardware [Shastri2021, LiR2024, ZhangZ2025]. Various machine learning modalities have found applications in quantum physics and photonics relevant to the topic of this work [Carleo2019, Kudyshev2021, Freire2023, Krenn2023, Dawid2025, Acampora2025, Du2025]. Restricted Boltzmann machines were exploited for quantum state modeling and reconstruction [Carleo2017, Torlai2018, Tiunov2020, Du2025], and deep fully connected neural networks for correcting systematic experimental errors during the photonic measurements [Palmieri2020]. Extraction and classification of high-level features directly from experimental data can also benefit from machine learning techniques, which outperform state-of-the-art methods in accuracy and speed. For example, an adaptive linear element neuron was used for identification of light sources based on their statistics [You2020], unsupervised machine learning isolation forest algorithm was exploited for anomaly detection of single proteins [Dahmardeh2023], and deep neural networks were used for entanglement quantification [Koutný2023, Du2025]. Convolutional networks found many applications in imaging. They were used to boost super-resolution microscopy based on the localization of single fluorescent molecules, such as Deep-STORM and DECODE [Nehme2018, Speiser2021], and improve optical metrology aided by scattering media [Gupta2020]. Autoencoders were applied to unsupervised learning of quantum phase transitions [Wetzel2017], phase recovery and holographic image reconstruction [Rivenson2018], site-resolved microscopy of neutral atoms in optical lattices [Impertro2023], and with skip connections (U-Net) also for supervised image segmentation, denoising, and reconstruction [Ronneberger2015]. An autoencoder-like pair of deep neural networks was used for optimal bidirectional control of photonic devices [Vašinka2022].

Despite the notable progress, there are still open questions and challenges in the integration of machine learning and physics, such as efficiency and data sufficiency, generalization ability versus "dreaming", incorporating physical constraints, transferability (and device independence), predicting uncertainties, explainability or even interpretability. The major challenge is the need for efficient data acquisition and processing. Many machine learning models are data-driven and require large amounts of high-quality training data to achieve good performance. However, acquiring such data can be time-consuming or even impossible with current technology, especially in the case of quantum photonics, where the signals of interest may be weak and difficult to measure. When possible, we mitigate this problem by learning from numerically simulated patterns, which brings two main benefits: 1. a large amount of data can be numerically generated much faster than using an experimental setup, 2. we do not learn systematic imperfections and artifacts of the particular experimental setup. Generalization is a model ability to make correct predictions on new but in-distribution data. It comes from learning stable patterns rather than memorizing examples. "Dreaming" (often called hallucination) is what happens when the model freewheels producing fluent but unfounded outputs [Hoffman2021], see Fig. 2.13. It typically happens under distribution shift, weak grounding, or aggressive sampling. Deep neural networks with high capacity and good regularization can generalize well, yet they can still dream confidently when asked to extrapolate without evidence. Practical antidotes are to enforce constraints and checks, detect out-of-distribution inputs, and report uncertainties together with outputs. Another challenge is the need to develop physically meaningful models that take into account the unique features of physical systems (physical priors) [Karniadakis2021]. For example, quantum photonic systems are subject to constraints such as positive semi-definiteness, probability conservation, and symmetry properties, which must be incorporated into machine learning models to ensure their physical validity. Consequently, the transferability of well-developed large models, such as language processing or image generation, to physical systems is severely limited.

Uncertainties in deep learning are estimated by having the model predict a probability distribution over the target rather than a single value [Kendall2017]. The network outputs both the usual prediction (e.g., a mean or a normalized probability map) and one or more uncertainty parameters (e.g., a variance for Gaussian targets [Nix1994, Sluijterman2024] or Dirichlet concentrations for probability maps [Sensoy2018, Malinin2018]). These parameters are learned directly from data by minimizing the appropriate negative log-likelihood, so no explicit "uncertainty labels" are required [Nix1994, Kendall2017, Amini2020]. At inference, total predictive uncertainty can be decomposed into aleatoric and epistemic components, separating noise captured by the predicted likelihood from variability due to the fitted parameters. The first term is available in a single forward pass via the learned variance or concentration [Nix1994, Sensoy2018, Amini2020]. The second can be estimated, without changing the training target, using techniques that expose parameter variability (e.g., keeping stochastic regularizers active at test time, training independent ensembles, or using local posterior approximations around the fitted weights) [Kendall2017]. This unified view—choose a likelihood, have the network output its parameters, and account for parameter variability when needed—covers most practical recipes for uncertainty quantification in current applications of deep learning [Liu2025].

Finally, there is a need to develop machine learning models with a higher degree of interpretability. Many machine learning models are black boxes that can be difficult to interpret for humans and may not provide insights into the mechanisms underlying their generalization ability and performance. Recently, symbolic regression and dimensionality reduction were proposed for interpretable modeling in physics [Cranmer2020]. An analogous strategy for photonic detectors and sensors is yet to be developed. If full interpretability is not within reach, particularly for deep neural networks, we should aim at explainability. It refers to methods that provide post-hoc explanations of models that are not intrinsically interpretable. An illustrative example of explainability approach is a method termed SHapley Additive exPlanations (SHAP) [Lundberg2017, Lundberg2020]. It provides a principled, game-theoretic framework that decomposes each model prediction into additive contributions from individual input features relative to a baseline prediction. To obtain a global measure of relevance, SHAP feature importance values are used, defined as the average absolute SHAP values across all data points. These values quantify how strongly a particular feature, on average, shifts the model output in its natural prediction scale away from the baseline [PonceBobadilla2024].

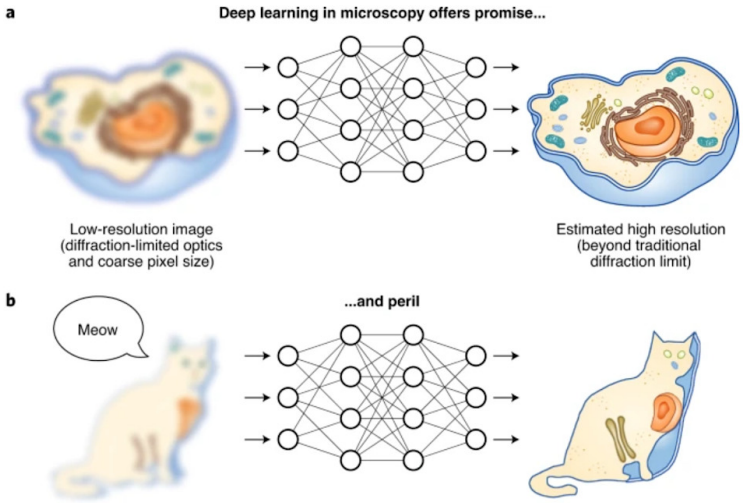
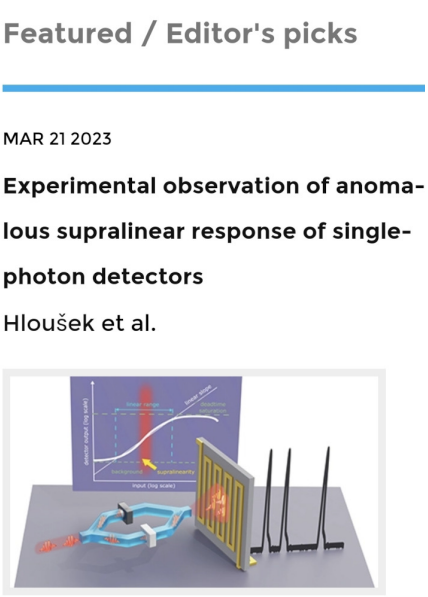


Figure 2.13. (a) Under the right conditions a properly trained deep neural network can predict a super-resolved micrograph from a blurry diffraction-limited (or lower resolution) image. (b) However, if the neural network encounters unknown specimens, or known specimens imaged with unknown microscopes, it can produce nonsensical results. Image and caption by D. P. Hoffman, et al. [Hoffman2021].

Most machine learning models in physics and particularly photonics are tailored for a specific measurement or imaging apparatus, such as a combination of optical response, signal-to-noise ratio, etc. The model has to be retrained if any specification of the apparatus changes [Belthangady2019], i.e., it lacks a device-independence property. Only very recently, our team presented a first step toward completely device-agnostic super-resolution imaging approach [Vašinka2025a, Dostálová2025a, Vašinka2025b]. Its generalization to multiparametric sensing is the subject of further work. Also, a systematic study of required detection channels and model complexity for target accuracy is lacking. In summary, AI shows strong potential for photonic detection and sensing, yet many open challenges remain.

3. Single-photon detector non-linearity

Single-photon detectors are indispensable tools in quantum technology, sensitive and quantum-enhanced metrology and advanced biomedical imaging, enabling the detection of individual photons. Crucially, the performance of these detectors is often assumed to be linear, meaning that the output detection rate scales proportionally with the incident photon flux. However, as shown in recent works [Akhlaghi2011, Kauten2014, Straka2020, Hloušek2023, Mark2024], this assumption may not always hold. This chapter presents a detailed summary of the direct, high-accuracy experimental study that revealed anomalous nonlinear behavior—particularly supralinearity—in widely used single-photon detectors, specifically silicon single-photon avalanche diodes (SPADs) and superconducting nanowire single-photon detectors (SNSPDs).



The presented results are based on the work:

- J. Hloušek, I. Straka, and M. Ježek, Experimental observation of anomalous supralinear response of single-photon detectors, Appl. Phys. Rev. 10, 011412 (2023).

In conventional optical measurements, linearity of the detector ensures that relative comparisons of light intensity levels are meaningful. A prime example is the measurement of sample transmittance, where intensity readings with and without the sample should differ only by the true attenuation factor, as shown in Fig. 3.1 (a). Any deviation from linearity introduces systematic errors, potentially degrading accuracy of the measurement. While correction models taking into account dark counts, dead time, and recovery dynamics can mitigate this issue, they depend on the validity of underlying theoretical assumptions. Existing models have limited predictive power and are often applicable only in narrow regimes. The importance of linear detection becomes critical in quantum-enhanced applications, such as quantum-enhanced metrology, where nonclassical states of light are employed to surpass classical sensitivity bounds [Giovannetti2004, Giovannetti2006, Giovannetti2011]. Shaping the statistics of light and using nonclassical optical signals as measurement probes allow for increasing the precision of length measurements [Slussarenko2017], imaging and particle tracking [Brida2010b, Ono2013, Israel2017, Moreau2019], and spectrophotometry [Kalashnikov2016, Whittaker2017]. The optical transmittance measurement assisted by correlated photons and single-photon detectors can serve as a prominent example of a quantum-enhanced measurement scheme [Jakeman1986, Sabines-Chesterking2019]. Nonlinear detector responses not only bias flux measurements but also distort photon statistics, impeding the realization of the quantum advantage. Notable experiments susceptible to detection nonlinearity are the tests of Born's rule in quantum mechanics [Sinha2010, Sollner2012, Magana-

Loaiza2016, Kauten2017, Cotter2017]. Precise experimental characterization of detection nonlinearities beyond limited theoretical modeling (model-independent characterization) is of paramount importance for advanced applications of single-photon detectors. In this context, we introduced an approach that resolves nonlinearities as small as 0.1% across seven orders of magnitude in count rate, without requiring calibrated references or time-resolved detection [Hloušek2023].

SPADs and SNSPDs operate in a binary mode, registering discrete voltage or current pulses in response to absorbed photons. Their detection output, a count rate R (number of detection events per second), depends on both the actual photon flux Φ and internal detection mechanisms. Traditional models predict a ∫-shaped nonlinearity curve (sublinearity) due to a combination of background noise at low flux and dead-time saturation at high flux, see Fig. 3.1 (b). However, we found deviations from this shape—most notably, an unexpected supralinear region where the detection rate increased faster than proportionally with the incident flux. This supralinear response had previously been observed only under specific conditions in classical silicon photodiodes [Schaefer1983, Tanabe2016] and SNSPDs [Kerman2013, Marsili2016]. Yet no existing SPAD or SNSPD model could account for such behavior under continuous low-power excitation. These discrepancies highlight the need for experimental model-independent characterization.

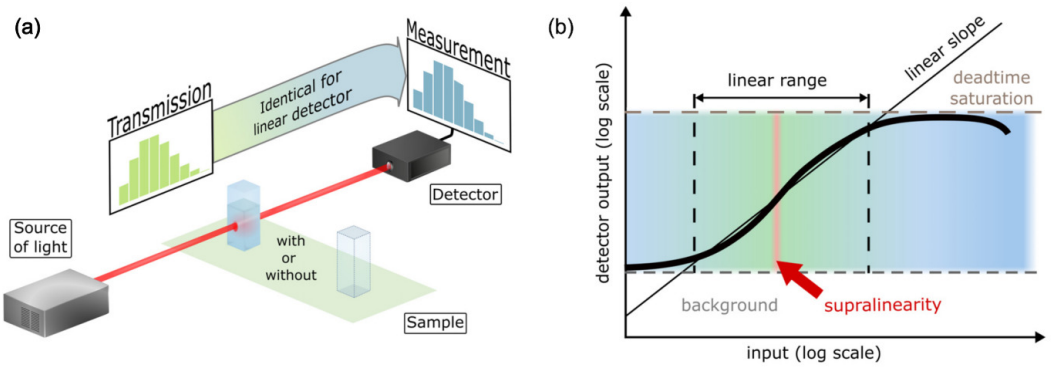


Figure 3.1. (a) Optical transmittance measurement of a sample relies on comparing two levels of intensity captured by a detector. Any deviation from the ideal linear detection then critically affects the measurement accuracy. Nonlinearity is a particularly vexing problem in quantum metrology when nonclassical statistics of light is often required together with single-photon detectors. Their inherently nonlinear response distorts the photon statistics and makes it difficult to reach the quantum advantage. (b) Nonlinear behavior of a single-photon detector. Apart from the background detections and saturation effects, single-photon detectors may exhibit supralinearity.

The core technique used in our study is an absolute (self-referenced), single-source, two-beam superposition method. A light source is split into two incoherent beams, labeled A and B, using a switched optical network. The detection rates R^A, R^B, and R^{AB} are measured for the two individual beams and their superposition, respectively, see Fig. 3.2 (a). Nonlinearity is quantified via the deviation: $\Delta = (R^A + R^B)/R^{AB} - 1$, which follows an approximate V-curve in a log-log scale due to dark counts (left slope) and saturation (right slope). $\Delta > 0$ indicates sublinearity and $\Delta < 0$ indicates supralinearity. A 50:50 split of the A and B beams is used for robustness and calibration independence.

The experimental setup depicted in Fig. 3.2 (b) achieves high coherence suppression and power stability by employing a stabilized and isolated superluminescent diode, polarization-maintaining fibers, unbalanced actively switched Mach–Zehnder paths, and high-extinction polarization beam splitters. For SNSPDs, an additional bandpass filter and polarization optimization were used. Extra effort was made to reach 10^{-5} power stability in the setup, including the fiber coupling stages. The acquisition time of each individual rate measurement was set to 20 s. The complete characterization of a single detector typically took 20 hours, with each measurement repeated 30 times for 40 different rate levels. The precision was found to be limited mostly by fundamental Poissonian shot noise. The detailed description of the setup, its stability, acquisition procedures, and statistical analysis can be found in Appendices A–E and G of our work [Hloušek2023].

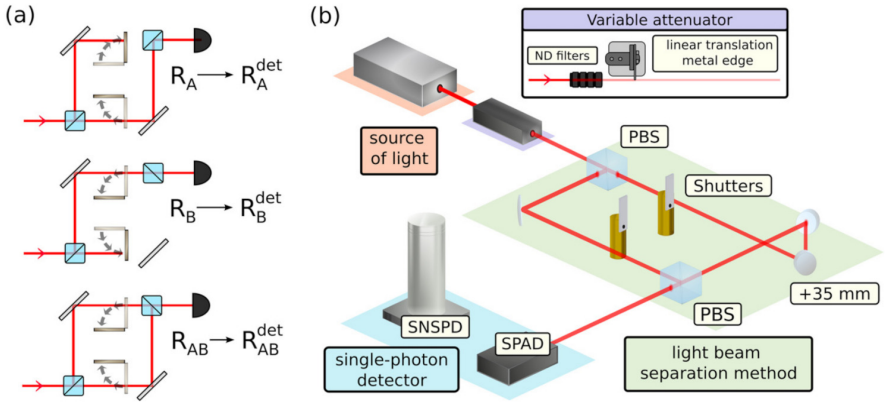


Figure 3.2. The single-source two-beam method for the absolute measurement of nonlinearity consisting of three measurements of the detection rate (a) and the simplified experimental scheme of the method (b).

The three actively quenched SPADs (SPAD-1 to SPAD-3) and one passively quenched SPAD (SPAD-P) were tested. The data revealed statistically significant deviations from the expected V-shaped nonlinearity curves, see Fig. 3.3 (a). Notably, SPAD-3 exhibited strong supralinear response in the mid-range detection rates (10^4 – 10^5 cps), inconsistent with all known models (see Table I and Fig. 11 in [Hloušek2023]). Even after applying dead-time corrections and comparing with independent recovery time measurements, supralinearity persisted. Alternative modeling attempts, including paralyzable/non-paralyzable dead time mixtures and empirical corrections for afterpulsing, failed to account for the effect.

For the SNSPD, the response was measured for varying bias currents. The nonlinearity plot (Fig. 3.3 (b)) exhibited complex sub- and supralinear features depending on both rate and bias. At moderate rates (~10^4 cps) and subcritical bias currents, a consistent supralinear regime emerged. At higher rates (>10^6 cps), latching and saturation effects became dominant. Prior explanations of supralinearity in SNSPDs included AC coupling artifacts at very high rates ≥ 10 Mcps [Kerman2013] and multiphoton absorption under ultrashort pulsed illumination with mean photon numbers ≥ 0.1 [Marsili2016]. However, the observed supralinearity at rates as low as 10^3 cps under continuous-wave excitation lacks precedent and points to intrinsic physical mechanisms, possibly involving complex relaxation dynamics of the detection hotspot.

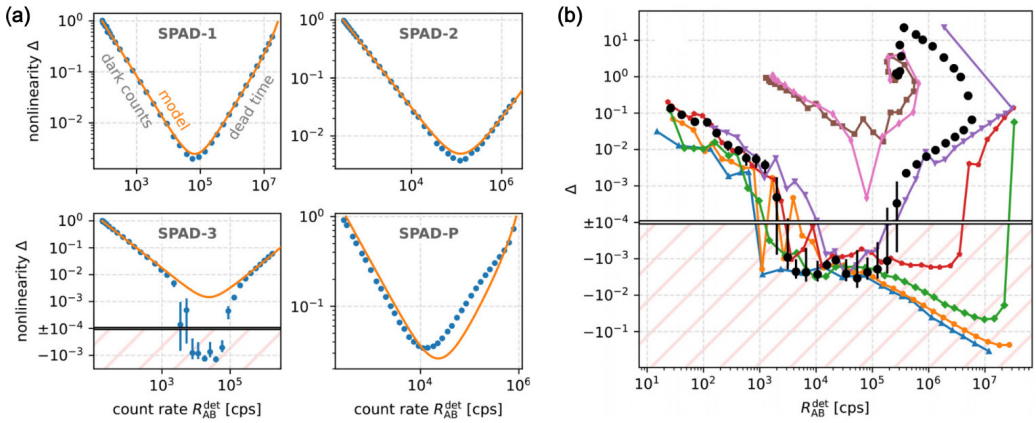


Figure 3.3. (a) Nonlinear response of the tested SPADs showing the supralinear tendency for all actively quenched detectors (SPAD-1 – SPAD-3) with unambiguous supralinear behavior for detector SPAD-3. Each point was measured 30 times, and the error bars show the corresponding standard error of the mean. The solid line represents a theoretical model. (b) The measured nonlinearity of the SNSPD for various values of bias current. For an operational current $I_{bias} \approx 25$ µA (black markers), the response transitions from sublinear (due to dark counts for detection rates below 10^3) to supralinear (detection rates from 2×10^3 to 2×10^5), and back to sublinear regime (due to saturation for detection rates larger than 2×10^5), ultimately strongly decreasing the detected rate with increasing photon flux (because of latching for rates above 10^7).

Recent work by Mark et al. further investigates the origin of supralinearity in SNSPDs under continuous-wave illumination [Mark2024], using a method based on incoherent beam superposition similar to that of [Hloušek2023]. They test both a standard dead-time model and an extended model including two-photon detection events, finding that only the latter reproduces the experimental data. However, the best fit requires assuming an effective dead time of more than a hundred nanoseconds, far longer than typical hotspot recovery times, indicating that even this extended model does not realistically capture the detector physical response. Consequently, two-photon absorption is a plausible mechanism to explain SNSPD supralinearity, but only if unexpectedly long or complex relaxation dynamics are assumed, beyond what standard SNSPD models predict.

The findings strongly indicate that the established single-photon models are insufficient to capture the full range of detector behaviors under practical conditions. For SPADs, supralinearity challenges assumptions even in well-understood devices. For SNSPDs, the observed behavior demands new modeling approaches that consider memory effects, rate-induced bias changes, and multiphoton absorption effects.

Our results have broad implications. Any optical measurement relying on single-photon detector linearity—be it classical radiometry, quantum metrology, or tests of fundamental physics—must consider direct calibration of the detector nonlinearity. Furthermore, our calibration approach is directly applicable to next-generation detectors based on low-dimensional materials (e.g., graphene, MoS_2), quantum dots, and Josephson junctions [Guo2020, Gu2019, Zhang2021]. Many such systems operate in the counting regime and show evidence of strong and bias-sensitive

nonlinearity. For example, photogating-based detectors have been shown to exhibit sub- and supra-linearities even at very low photon flux [Chen2018].

In summary, in the study [Hloušek2023] we established a new experimental standard for quantifying and understanding nonlinear responses in SPDs. By circumventing reliance on theoretical response models, it reveals previously unknown behaviors that must be accounted for in accurate photon detection in photonics and quantum technologies.

4. Photon statistics detection, nonclassicality sensing, and emitter counting

Advancing photonic detection from binary single-photon detection to true photon counting marks a paradigm shift, enabling applications from quantum technologies to biomedical imaging. Photon-number-resolving detectors (PNRDs) are central to this progress. Beyond the experimental challenges of implementing and characterizing PNRDs, a key difficulty lies in processing detection data to recover photon statistics across the widest range of photonic states—from classical chaotic and coherent light to strongly nonclassical states with reduced amplitude noise (squeezed light) or photon number uncertainty (Fock states). Current state-of-the-art detectors have been demonstrated mainly for states with low mean photon numbers or for larger classical states, typically with modest fidelities. Direct certification of nonclassicality, without reconstructing photon statistics, is also critical for benchmarking photon sources, quantum readout of matter systems, counting quantum emitters, and other advanced sensing and imaging. While such certification has been demonstrated, it remains generally suboptimal, particularly for larger states and under realistic experimental conditions with noise. This chapter presents a detailed summary of experimental achievements developing a large, accurate, and reconfigurable PNRD, cutting-edge photon statistics retrieval method, and direct certification of nonclassicality of multi-photon states.

PHYSICAL REVIEW LETTERS **123,** 153604 (2019)

Accurate Detection of Arbitrary Photon Statistics

Josef Hloušek, Michal Dudka, Ivo Straka, and Miroslav Ježek*
Department of Optics, Palacký University, 17. listopadu 12, 77146 Olomouc, Czechia

The presented results are based on the works:

- J. Hloušek, M. Dudka, I. Straka, and M. Ježek, Accurate detection of arbitrary photon statistics, Phys. Rev. Lett. 123, 153604 (2019);
- J. Hloušek, J. Grygar, M. Dudka, and M. Ježek, High-resolution coincidence counting system for large-scale photonics applications, Phys. Rev. Applied 21, 024023 (2024).

The following works are also briefly summarized:

- I. Straka, L. Lachman, J. Hloušek, M. Miková, M. Mičuda, M. Ježek, and R. Filip, Quantum non-Gaussian multiphoton light, npj Quant. Information 4, 4 (2018);
- L. Lachman, I. Straka, J. Hloušek, M. Ježek, and R. Filip, Faithful hierarchy of genuine n-photon quantum non-Gaussian light, Phys. Rev. Lett. 123, 043601 (2019).

The aim of a PNRD is to access the photon-number distribution $\mathbf{p}=\{p_n\}$ of an incoming optical signal rather than just detect its presence. Since most practical single-photon detectors are binary devices, the photon-number resolution is achieved by splitting the incoming light into multiple spatial or temporal channels, each detected independently with a binary single-photon detector. This multiplexing

distributes photons probabilistically among the channels so that multi-photon arrivals produce multi-detector click events. These coincidence detections are evaluated by an electronic circuit. The measured detection statistics $\mathbf{c}=\{c_m\}$ (probabilities of simultaneous clicks of m detection channels) are then related to p_n via a known linear transformation, $\mathbf{c} = \mathbf{D}\,\mathbf{p}$, where $\mathbf{D}$ is detection matrix dependent on the number of channels, detection efficiency, noise, and other parameters of the PNRD. By inverting this transformation—carefully, to avoid ill-posed inversion artifacts—one can reconstruct the photon-number distribution of the input state or assess the statistical features of the detected light. The whole workflow is shown in Fig. 4.1 (a).

The multiplexing of our PNRDs is implemented using a discrete spatially-multiplexed optical network constructed from cascaded tunable beam splitters. Each tunable beam splitter consists of a half-wave plate (HWP) and a polarizing beam splitter (PBS). Rotating the HWP changes the polarization and thus the splitting ratio at the PBS, allowing continuous and precise adjustment of power distribution between output modes. By cascading these tunable splitters, the incoming optical mode can be divided into M spatial channels in a balanced (absolute error below ~0.3%) or arbitrary ratio configuration. The PNRD is fully reconfigurable, meaning both the number of output channels and their splitting ratios can be precisely tuned without replacing optics. This allows adaptation to different source characteristics, balancing of channel efficiencies, and even specialized configurations with fewer channels for direct evaluation of certain correlation or nonclassicality measures without full statistics retrieval.

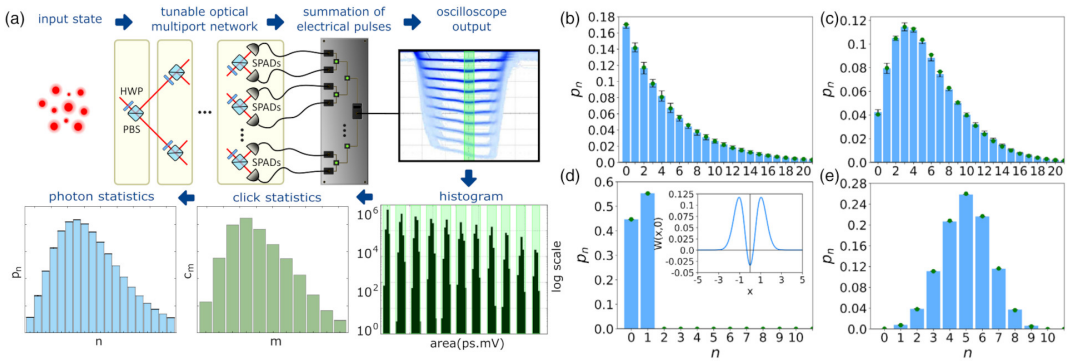


Figure 4.1. (a) Experimental setup of the PNRD based on a discrete optical network with full reconfigurability and continuous tunability of splitting ratios, pulse-height spectrum of the analog output of the detector, and scheme of photon-statistics retrieval [Hloušek2019]. Measured (blue bars) and the corresponding theoretical photon statistics (green dots) for (b) thermal state mean photon number of approximately 5, (c) two-photon-subtracted thermal state, (d) single-photon, and (e) heralded nine-photon state that emulates emission from a cluster of single-photon emitters. (Inset) Wigner function evaluated from the measured statistics. Note that the data agree with theory even beyond the number of channels of the PNRD (M=10).

Light in each output channel of the optical network is delivered to a single-photon avalanche diode (SPAD) in free space or coupled into anti-reflection-coated optical fibers (depending on the experiment). Typical parameters of SPADs used in the PNRD: detection efficiency 55–70% at 0.8 μm, timing jitter 200–300 ps, and dead time 20–30 ns. The overall system efficiency of the ten-channel PNRD was around 50(1)%. Ultra-low-loss optics and superconducting nanowire detectors (instead of

SPADs) can be used to boost the system efficiency above 85%. The electronic circuitry processes the binary detection signals from all detection channels and determines the number and, optionally, identity of detectors firing in each optical pulse window. The implementation evolved from analog fan-in circuitry to fully digital coincidence logic.

In the first implementation [Hloušek2019, Hloušek2022], the SPAD outputs were standardized and synchronized by 300 MHz discriminators and passive delay lines, and fed into a splitter/merger that summed electrical pulses from all channels while preserving their relative timing. The circuit produced a multilevel analog output where each voltage level corresponded to a specific number of simultaneous clicks (0 to M). The output was visualized and processed by an oscilloscope. Alternatively, the passive merger was replaced by an array of active fan-in/out units with bandwidth of 250 MHz. The circuitry was designed to be crosstalk-free, synchronized across channels, and with latency < 30 ns (including SPAD response), enabling use in feedback and quantum communication tasks. The setup operated in a gated pulse regime (≤ 5 MHz repetition) to suppress dark counts and afterpulsing.

Later we introduced a coincidence counting unit based on fast positive emitter-coupled logic and programmable delay lines [Hloušek2024a]. This system accepts up to 16 channels (+ gate input), shapes each SPAD signal into a well-defined digital pulse, and aligns them temporally with <10 ps timing resolution. The coincidence window is tunable from sub-100 ps to 20 ns. Two modes are supported: Full histogram mode records a complete 2^M-element histogram of all possible click patterns with full channel resolution (up to 65 536 combinations for M=16). Reduced histogram mode records total coincidence numbers only (M+1 numbers), without channel identities, enabling higher processing rates. For balanced PNRD, the normalized reduced histogram gives the detection statistics c_m.

The coincidence unit can handle input rates up to 1.5 GHz (limited by delay line speed), with an overall event processing rate of 2–3 million events/s depending on mode. It achieves ultralow coincidence error probability ($< 10^{-7}$) and allow real-time feed-forward triggering to other experimental hardware [Švarc2020, Hloušek2024b]. When even higher temporal resolution and data throughput are required, and the feed-forward capability can be sacrificed, a time-to-digital converter (time tagger) can be used instead of the coincidence unit.

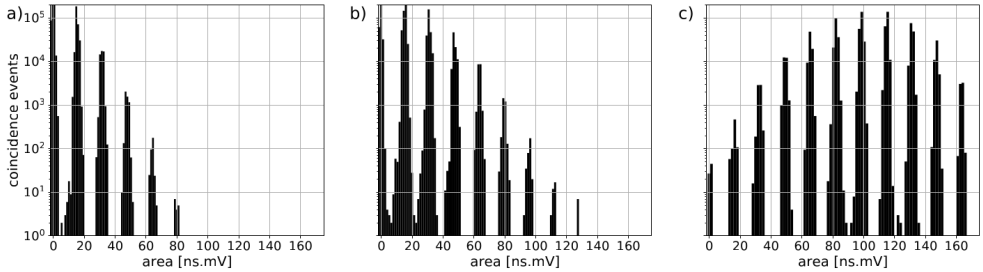


Figure 4.2. Pulse-height spectrum of the analog output of the PNRD reported in [Hloušek2019] for coherent states of various mean photon number ⟨n⟩ = 1 (a), 5 (b), and 20 (c). The spectra are plotted in log scale to emphasize the perfect energy resolution and the absence of any crosstalk effects or background noise. The relative magnitude of these adverse effects is below 10^{-5} and has been further reduced in more recent experiments.

The use of independent detectors combined with a well-balanced merging/coincidence circuit completely eliminates crosstalk between histogram channels, ensuring accurate resolution of discrete photon numbers up to the total number of channels, as shown in Fig. 4.2. In addition, operating the detector in a pulsed regime with a repetition rate below approximately 5 MHz effectively suppresses dark counts and afterpulses. The interval between measurement runs can be reduced to slightly more than the recovery time of the single-photon detectors, provided that afterpulsing is negligible or decays rapidly, as in superconducting nanowire single-photon detectors. As a result, the presented PNRD measurements are inherently free from systematic errors such as channel crosstalk and temporal correlations, and do not require challenging detector tomography [Lundeen2009].

Retrieving the photon statistics p_n, n = 0, 1,…, ∞, corresponding to a measured click statistics c_m, m = 0, 1, …, M, represents the core problem of photon-statistics sensing. This generally ill-posed problem suffers from underdetermination and sampling error. Fortunately, we have additional constraints facilitating the retrieval; i.e., the photon-number probabilities are non-negative, normalized, and typically non-negligible only within a finite range. We have developed a novel approach, termed the expectation-maximization-entropy (EME) method, based on an expectation-maximization iterative algorithm weakly regularized by a maximum-entropy principle. The initial zeroth iteration is uniform for sufficiently large n_{max}, much greater than mean number of photons in the detected state. Each subsequent iteration is computed using the previous iteration, measured detection statistics $\mathbf{c}=\{c_m\}$, and the detection matrix **D** via the extremal equation derived by maximization of likelihood functional regularized by entropy of photon statistics $\mathbf{p}=\{p_n\}$, see the main text and supplemental materials of [Hloušek2019]. The supplement also includes the detailed derivation of the EME method and its comparison with a pure maximum-likelihood (the expectation-maximization algorithm), and a direct pseudoinverse approach.

The strength of the entropic regularization is controlled by a parameter λ, which scales roughly as $1/\sqrt{N}$ with the number of analyzed state copies (optical pulses); in our case $\lambda \approx 10^{-3}$. The exact optimum for given set of states can be determined by numerical simulations. Further research work is needed to remove the ambiguity of choosing the value of λ. In general, entropic regularization may appear as artificial smoothing of the retrieved photon statistics. However, pure maximum-likelihood reconstruction (without regularization) is known to bias solutions toward being more sparse (higher purity) than the true state, particularly when only a small number of copies is available [Schwemmer2015, Silva2017]. This bias arises in all methods that enforce positivity [Kalev2015]. Entropic regularization compensates for this effect, driving the reconstructed statistics closer to the true distribution.

To show the accuracy and the robustness of the EME method, we performed a numerical analysis for various photon statistics with different mean photon numbers. We compared the EME method with the EM and direct pseudoinversion. EME was found to be a unique estimator that guarantees non-negativity and the absence of numerical artifacts in the retrieved photon statistics. Total variation distance between the retrieved distribution and the true one is on the order of 10^{-3}, one order of magnitude smaller than in the case of the other approaches. Numerical simulations

yield average fidelity values F = 0.9996 using the EME algorithm and F = 0.997 using the maximum-likelihood approach, again an order of magnitude improvement in infidelity. We also found that the EME convergence is 10–1000× faster than the plain EM approach. For details, see supplemental materials of [Hloušek2019].

In experimental demonstration, we analyzed coherent states, thermal states, multimode thermal states, single-photon and multiple-photon-subtracted thermal states, and nonclassical multiphoton states. Furthermore, we have varied the mean number of photons, the number of modes, and the number of subtracted or superimposed photons. From this vast set of measured states, 25 of them are shown in Fig. 4.3, covering various values of mean photon number $\langle n \rangle$ and autocorrelation parameter $g^{(2)}$, evaluated from the measured photon statistics, $g^{(2)} = \langle n(n-1) \rangle / \langle n \rangle^2$. Photon-statistics sensing provides faithful g(2) values for states, where the commonly used Hanbury Brown–Twiss measurement would fail due to high multiphoton content [Stensson2018]. The unprecedented accuracy is achieved despite leaving all systematic errors (residual imbalance of the PNRD, dark counts of SPAD detectors, jitter of the electronics, etc.) uncorrected and operating with raw data.

The coherent states measured by the PNRD show almost perfect Poissonian statistics with $g^{(2)} = 1$ up to $\langle n \rangle = 20$ with average fidelity $\bar{F} = 0.996$. For the thermal states we measured almost ideal Bose-Einstein photon statistics depicted in Fig. 4.1 (b) with $g^{(2)} = 2$ up to $\langle n \rangle = 5$ with F = 0.997. We varied the number of the collected thermal modes, which yielded a signal governed by Mandel-Rice statistics, going from Bose-Einstein to Poisson distribution as the number of modes increased. Multiple-photon subtraction from the thermal state was implemented using a beam splitter with a 5% reflectance. When a (multi)coincidence was detected by a multichannel single-photon detector in the reflected port, the heralded optical signal in the transmitted port was analyzed by the reported PNRD. A typical result of two-photon subtraction is shown in Fig. 4.1 (c). Increasing the number of subtracted photons results in a transition from super-Poissonian chaotic light to a Poissonian signal [Allevi2010b, Zhai2013]. Furthermore, we generated multiphoton states by mixing incoherently several single-photon states from spontaneous parametric down-conversion using time multiplexing. N_p successive time windows, where a single photon was heralded, were merged into a single temporal detection mode. This source emulates the collective emission from identical independent single emitters [Straka2018a, Ta2015]. The resulting photon statistics measured for these highly nonclassical multiphoton states corresponds extremely well to the ideal attenuated N_p-photon states, see Figs. 4.1(d) and 4.1 (e) for $N_p = 1$ and 9 with F = 0.999. Also the $g^{(2)}$ parameter computed from the measured photon statistics perfectly agrees with the theoretical model $1 - 1/N_p$, see Fig. 4.3.

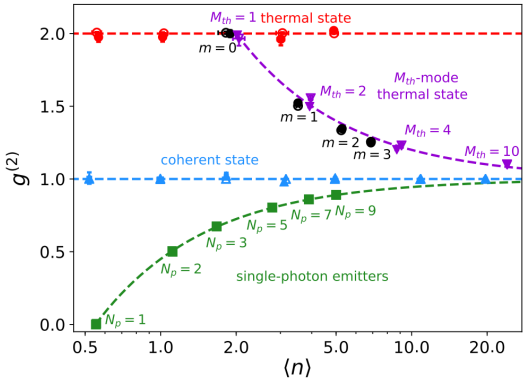


Figure 4.3. The autocorrelation parameter $g^{(2)}$ evaluated from the measured photon statistics (solid marker) and the corresponding ideal statistics (empty marker) of various optical signals with mean photon number ⟨n⟩ [Hloušek2019]. Shown are the coherent states with $g^{(2)} = 1$ (blue triangle up), thermal states (also termed chaotic light) with $g^{(2)} = 2$ (red circle), M_{th}-mode thermal states with $M_{th} = 1, 2, 4, 10$ (violet triangle down), and m-photon-subtracted thermal states for m = 0, 1, 2, 3 (black circle). The cases of $M_{th} = 1$ and m = 0 coincide with the thermal state. Furthermore, the emission from a cluster of N_p single-photon emitters is shown for $N_p = 1,\ldots,9$ with $g^{(2)} = 1 - 1/N_p$ (green squares).

Improving the whole experimental workflow, mainly the electronics circuitry, of the PNRD in [Hloušek2024a] allowed us to further increase the dynamic range and accuracy of photon-statistics sensing. We reached a fidelity larger than 0.999, i.e., errors as small as 10^{-3}, over the range of photon numbers up to 60, see Fig. 4.4. Currently, the largest PNRD developed in our lab consists of 20 SPADs [Večerka2025]. Their spatial multiplexing is still in free-space, as described above, but significantly miniaturized. The whole 20-channel PNRD occupies an area of 60×60 cm, less than the original 10-channel PNRD presented in [Hloušek2019, Hloušek2024a], see Fig. 4.5. Furthermore, the SPAD detectors employed are built in house including biasing, active quenching, and temperature stabilization. The detection signals are processed by a concidence unit with up to 32 channels, providing a complete information about coincidence patterns and the reduced histogram without channel identity, enabling higher processing rates.

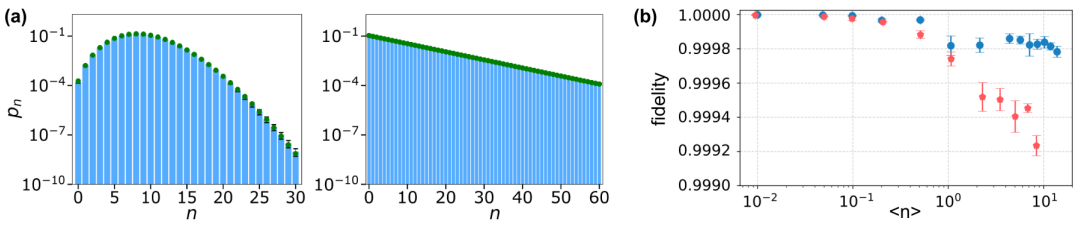


Figure 4.4. (a) Retrieved (blue bars) and theoretical (green circles) photon statistics of coherent state (left panel) and thermal state (right panel) with mean photon numbers of $\langle n\rangle_{coh} = 10.21(1)$ and $\langle n\rangle_{th} = 8.41(3)$ [Hloušek2024a]. (b) Accuracy of the retrieved photon statistics: fidelity as a function of mean photon number ⟨n⟩. Shown are the coherent states (blue circles) and thermal states (red pentagons). Note the vertical scale; all fidelities are higher than 99.9%.

Figure 4.5. (a) Photo of the 20-channel photon-number-resolving detector with free-space fully reconfigurable multiplex and two arrays of custom-built single-photon avalanche diodes. The whole PNRD system is fitted on a 60x60 cm optical breadboard. (b) A sneak peek of the in-house-built 32-channel coincidence unit assembly. Photos from our lab (QOLO).

To conclude, we demonstrated a fully reconfigurable, near-ideal photon-number-resolving detection scheme combining custom optoelectronics with a novel EME photon-statistics retrieval method. The design is free of systematic errors, which are either negligible or can be tuned away, and achieves accuracy beyond the conventional channel limit of PNRDs. We measured a wide variety of photonic states, from highly nonclassical light to large coherent and chaotic signals, retrieving photon statistics directly from raw data with no detector characterization and only the EME algorithm. This calibration-free operation is critical, as the full detector tomography at such scale (M=20, $n_{max}\approx100$) is extremely challenging. Despite uncorrected imperfections and variability of the input, the method consistently delivers fidelities larger than 99.8% for photon numbers up to 100, with $g^{(2)}$ values accurate to within a fraction of a percent. While demonstrated with common SPADs, the workflow is applicable to any on–off detector technology. The multichannel architecture also supports straightforward on-chip integration, with possible gains in speed, efficiency, and compactness expected from superconducting nanowire single-photon detectors integrated with waveguide platforms.

Besides photon-statistics sensing, we have used the developed PNRD for a direct assessment of strongly nonclassical (quantum non-Gaussian) properties of quantum states of light [Straka2018a, Lachman2019]. The reconfigurability of the PNRD allows changing the number of detection channels, which proved crucial for testing various orders of quantum non-Gaussianity criteria hiearchy, see Fig. 4.6 [Straka2018a]. Furthermore, a smaller variant of the PNRD with 8 channels (and detection efficiency reaching 70%) is currently used for sensing the emission statistics and statistical feature extraction of single organic molecules [Dostálová2025b]. We integrated the PNRD with a custom-built single-molecule high-numerical-aperture microscope and acquired the emission statistics and its long-term dynamics of single terrylene molecules and their clusters. We aim for improved emitter counting strategies and novel techniques of super-resolution imaging. Another current research focuses on significantly increasing the effective number of PNRD detection channels. The goal is to accurately retrieve photon statistics of large photonic states up to thousands of photons, entering the mesoscopic region. We will combine the technology of spatial multiplexing with an actively-switched time-multiplexed networks developed by our group [Švarc2019, Švarc2023, Švarc2024, Mazin2025] and modulation of PNRD properties.

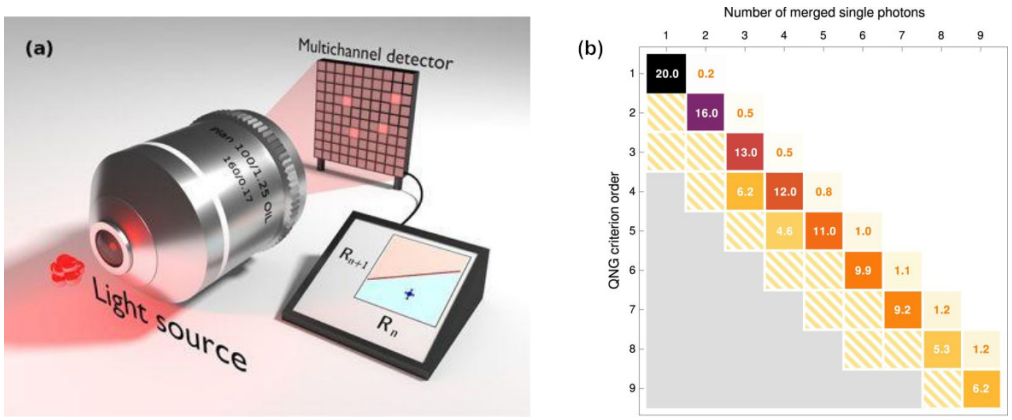
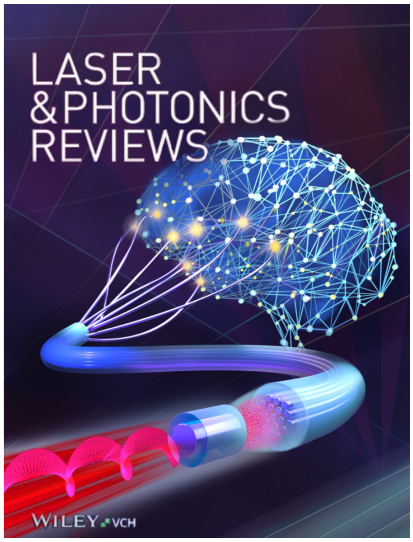


Figure 4.6. (a) A general proposal of the experimental quantum non-Gaussianity (QNG) witness [Straka2018a]. Multi-photon light is collected and brought to a balanced PNRD detector with m+1 channels, where coincidences c_m and c_{m+1} are compared to the QNG threshold. (b) Table of QNG depths (in dB), defined as the maximum attenuation, for which the quantum state is still QNG. The horizontal axis shows the number of single-photon states that comprise the measured state. The vertical axis represents the order m of the QNG criterion used to measure the state. Solid-colored tiles represent points with positively measured QNG despite statistical uncertainties. For points above the diagonal, the depth estimates are conservative and lower than the actual QNG depth, because c_{n+1} is no longer caused solely by noise. The upper white region represents combinations of measured states and criteria that did not show QNG. Orange stripes denote measurements where statistical uncertainty intersects with the QNG criterion border, making the result inconclusive. Data in the gray region contain no detections at all.

5. Optimal preparation, control, and sensing of polarization states of light



This chapter presents our recent progress in ultra-accurate and fast preparation, control, and measurement of light polarization at the single-photon level. We will show that twisted nematic liquid crystal devices, when carefully calibrated, can achieve polarization preparation and measurement with fidelities reaching 99.9%, enabling compact and precise polarimetry. Furthermore, machine-learning techniques can overcome the complex response of such devices, yielding bidirectional control with infidelities of 10^{-4} and enabling local and remote single-photon polarization state preparation. Finally, we will present an all-fiber sensor for highly accurate, fast, and in-situ polarization sensing.

The presented results are based on the published works:

- M. Bielak, R. Stárek, V. Krčmarský, M. Mičuda, and M. Ježek, Accurate polarization preparation and measurement using twisted nematic liquid crystals, Opt. Express 29, 33037 (2021);
- D. Vašinka, M. Bielak, M. Neset, and M. Ježek, Bidirectional deep learning of polarization transfer in liquid crystals with application to quantum state preparation, Phys. Rev. Applied 17, 054042 (2022);
- M. Bielak, D. Vašinka, and M. Ježek, All-fiber microsensor of polarization at single-photon level aided by deep-learning, Laser Photonics Rev. e01775 (2025).

Aiming at preparation of arbitrary polarization states of light, we developed a polarization modulator based on a sequence of twisted nematic liquid crystal (TNLC) cells harvested from commercial liquid crystal displays [Bielak2021]. For the experimental demonstration, we selected the Lumex LCD-S101D14TR seven-segment reflective TNLC display (33×50 mm) as the best trade-off between segment size and overall compactness, see Fig. 5.1 (a). The central segment has an approximately rectangular 15×4.5 mm clear aperture, which is perfectly sufficient for collimated optical beams with mode field diameter up to 2.5 mm. Because reflective TNLC displays cannot be used directly in transmission or as unitary elements, we removed the reflective backing, protective films, polarizers, and auxiliary layers, leaving the bare glass cells. Three such cells were then stacked to provide independent polarization transformations as shown in Fig. 5.1 (b). We mounted the cells on a printed-circuit assembly that integrates a microcontroller, 16-bit DAC, and voltage amplifier, see Fig 5.1 (c). Each cell is driven by a symmetric 1 kHz square wave (50% duty cycle) with 0–10 Vpp amplitude (we refer to the amplitude simply as

a control voltage), applied between the segment electrode (G) and the common back electrode (COM). The resulting TNLC device can address the polarization state of a single or even several parallel optical beams arranged in a line or a matrix in complex quantum information processing circuits [Lanyon2009, Mičuda2013, Xiao2017, Xiao2020, Stárek2018a, Stárek2018b].

We then modeled a single twisted-nematic cell as a cascade of infinitesimal retarders with an overall rotation and twist, and we replaced the common arctangent law for voltage-dependent retardance with a logistic form that described the measured retardance–voltage behavior more faithfully. Experimentally, we prepared a fixed input polarization, rotated the cell, scanned the control voltage, and tomographically read out the output using a reference polarimeter. The data showed that a single cell maps the input only onto a one-dimensional trajectory on the Bloch (Poincaré) sphere, so a single element cannot reach arbitrary pure states. Despite fitting, the forward model still exhibited systematic deviations from measurements and varied across nominally identical cells, likely due to edge fields, multiple reflections, and thickness inhomogeneity. We therefore concluded that no available forward model achieves sufficiently high fidelity with experimental data over all states, motivating the model-free calibration and multi-cell architecture used later in the work.

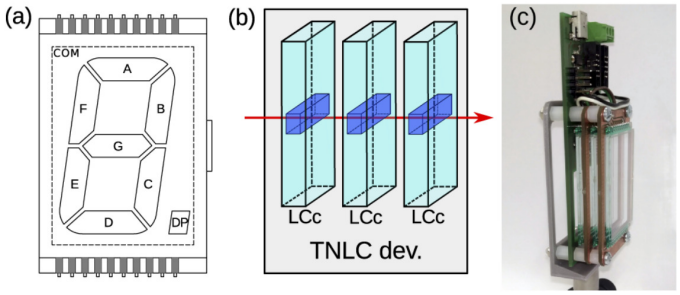
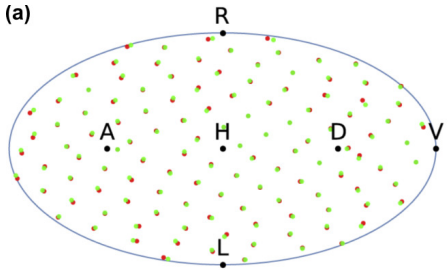


Figure 5.1. (a) Seven-segment liquid crystal display layout; elements of the digit are labeled from A to G, the decimal point as DP, and the common back electrode as COM. (b) Scheme of the presented TNLC device consisting of three TNLC cells (LCc); dark blue blocks show the active segments and the red arrow represents the optical beam. (c) Photo of the TNLC device mounted on printed-circuit board containing the electronic driver and communication interface [Bielak2021].

To overcome the intrinsic modeling difficulties of TNLCs, we implemented a model-free calibration procedure based on an evolutionary (genetic) algorithm that directly optimizes applied voltages. By cascading three TNLC cells, we realized arbitrary transformations of horizontally polarized input light into well-defined target states. In total, we prepared 120 states uniformly distributed across the Bloch sphere and verified their quality through independent measurements, achieving an average state fidelity of 0.999 with angular deviations < 0.5°, see Fig. 5.2 (a). We also demonstrated a preparation of mixed polarization states by combining deterministic settings with randomized sequences, thus enabling controlled depolarization with tunable degree of polarization. Calibrations proved stable under typical lab drifts: no measurable change for cell-temperature variations within ±3 °C or laser-wavelength changes within ±5 nm.

We further presented the use of the TNLC device as a reconfigurable polarimeter. To evaluate its performance, we tested both minimal tomography (based on four

projective measurements arranged in a tetrahedral configuration on the Poincaré sphere) [Řeháček2004] and overcomplete tomography (based on six measurements corresponding to the eigenstates of the Pauli operators: horizontal, vertical, diagonal, antidiagonal, right- and left-handed circular polarizations). Polarization reconstructions were carried out using maximum-likelihood estimation [Ježek2003, Hradil2004], ensuring physically valid density matrices. Our results showed that the TNLC polarimeter achieved reconstruction fidelities exceeding 0.999, comparable to those obtained with conventional motorized waveplate-based systems. Importantly, electronic reconfiguration allowed us to perform complete tomography sequences within about one second, roughly an order of magnitude faster than reference setups relying on mechanical rotation of birefringent elements. The reconfiguration times are detailed in Fig. 5.2 (b).



(b)

initial state \ final state	H	V	D	A	R	L
H		49	226	590	251	319
V	168		182	597	271	329
D	313	232		200	203	255
A	444	403	188		312	197
R	566	151	417	481		141
L	601	497	299	185	424	

Figure 5.2. (a) Demonstration of the Bloch (Poincaré) sphere coverage using the reported TNLC device [Bielak2021]. The target states are represented by red points, and green markers show the states prepared by the TNLC device. The sphere is visualized in the Hammer map projection. (b) Transition times in milliseconds between initial and final polarization projections of the TNLC device.

Finally, we demonstrated the utility of our device in quantum optics experiments at the single-photon level. Specifically, we applied the TNLC-based polarimeter to perform tomographic characterization of a three-qubit Greenberger–Horne–Zeilinger (GHZ) state generated by a photonic quantum Toffoli gate [Stárek2018a]. Using the overcomplete Pauli tomography and maximum-likelihood reconstruction, we obtained a state fidelity of 0.936 with respect to the ideal GHZ state, which is in excellent agreement with results from standard waveplate tomography. This experiment confirmed that TNLC-based devices are capable of ultra-accurate and fast polarization preparation and measurement even under demanding single-photon conditions, highlighting their promise as robust, vibration-free, and fast tools for both classical polarimetry and advanced quantum information processing.

In subsequent work published in 2022, we presented bidirectional deep-learning control of the three-cell TNLC device, learning both the direct mapping (voltages → polarization state) and the inverse mapping (target state → voltages) [Vašinka2022]. A fully connected direct network (predictor), trained on tomographically reconstructed polarization states, predicts output states with $\sim 4\times10^{-4}$ average infidelity, clearly outperforming linear and radial-basis interpolation baselines and delivering faster inference, while remaining robust as the dataset scales. For the inverse problem, we fixed the trained direct model and learned an inverse network (controller) that is validated in state space through the frozen forward model. This compound pipeline achieves $\sim 2\times10^{-4}$ average infidelity on unseen targets—again orders of magnitude better than interpolation—while maintaining low computation

cost. Analysis across model sizes shows these deep networks generalize well even in the over-parameterized regime, reducing the amount of calibration data needed for high accuracy.

The proposed bidirectional learning represents a universal approach to optimal control of quantum devices, not just TNLC polarization modulators. Typically, a quantum device is controlled by various classical signals—voltages, currents, radio-frequency fields, digital signals, etc. These signals affect the device operation and its quantum performance, whether in prepared quantum states or performed quantum operations. The transformation from control signals (e.g., voltages) to quantum states (described by density matrices) can be described by a direct predictor, such as a neural network trained on experimental data. However, the general goal is different: finding optimal control signals for a given quantum state. This inverse transformation maps the target quantum state to a set of control signals. When trained directly on experimental data, it may struggle with ambiguity. Namely, different control settings may lead to the same quantum state. In addition, the training error is evaluated in the space of control signals, which has an arbitrary structure. The number and ranges of control signals are technical parameters without a unified, fundamental physical interpretation. For example, when comparing two similar devices, one with three control voltages (overall error 20 mV) and another with four control voltages (error 30 mV), the latter can perform better despite the nominally higher error. Daisy-chaining the inverse and direct models removes these issues, because the mapping is performed from the space of target quantum states to the space of actually prepared states, see Fig. 5.3. Here the training error is evaluated in quantum-state space using physically meaningful metrics such as fidelity. The compound network architecture resembles an autoencoder trained in an unsupervised regime [Kramer1991, Hinton2006]. The supervision is implicit, provided by the already trained direct part, while the control signals play the role of the latent variables. After training, the first part of the autoencoder is extracted, yielding the optimal controller.

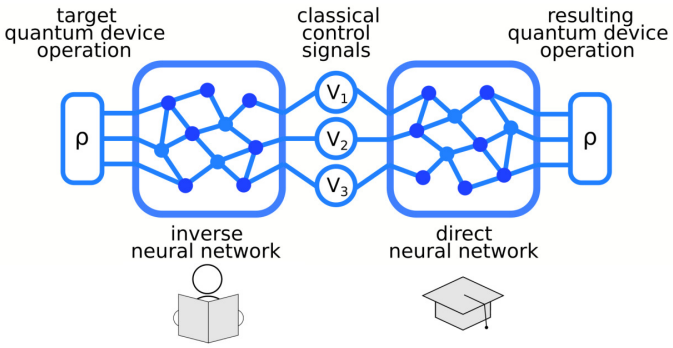


Figure 5.3. Representation of the compound model created by connecting the inverse model (controller) to the already-trained direct model (predictor) [Vašinka2022]. The inverse part transforms parameters of a polarization density matrix to three control voltages. The direct part then converts these voltages back into a polarization state. The direct part is fixed during the learning process, and only the inverse part is trainable.

The learned controller was then experimentally validated at the single-photon level, see Fig. 5.4. Locally, the TNLC device prepared >1000 target states (forming the UP logo on the Bloch sphere) with six-projection tomography confirming an average fidelity ≈ 0.998 to the modeled targets, see Fig. 5.5 (b). Remotely, using entangled photon pairs from a Sagnac source, two-qubit tomography showed high-quality

entanglement, and numerically propagated measurement operators yielded remote state preparation with average fidelity ≈ 0.988, see Fig. 5.5 (c). Together, these results demonstrated ultra-accurate, fast, and scalable polarization control—from continuous analog voltages to local and remote single-photon state engineering.

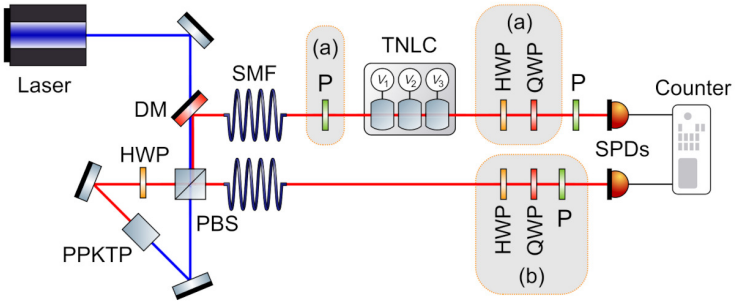


Figure 5.4. (a) Scheme of heralded single-photon polarization state preparation using the idler photon (lower red path) for heralding. The components denoted as (b) are not present in this configuration. The PPKTP crystal inside a Sagnac interferometer is pumped unidirectionally (blue line). The heralded signal photon (upper red line) in the horizontal polarization state (after polarizer P) is transformed by the LC device and analyzed by quantum tomography using wave plates (HWP and QWP), a polarizer (P), and a single-photon detector (SPD). The coincidence detection events of the signal and idler detectors are acquired by a counter. (b) Setup for remote quantum state preparation using an entangled pair of photons. The components denoted as (a) are not present in this configuration. The generation of entangled photon pairs requires bidirectional pumping of the PPKTP crystal. The signal photon is projected to a target polarization state by the TNLC device and the polarizer. The polarization state of the idler photon is analyzed by quantum tomography.

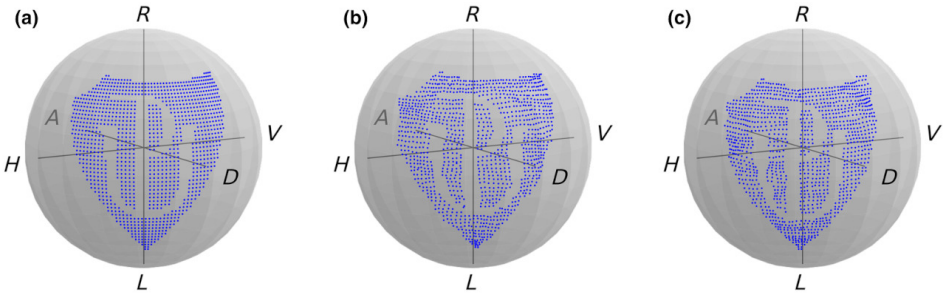


Figure 5.5. Visualization of the polarization states forming the Palacký University logo on the Bloch sphere [Vašinka2022]. (a) The target states for which we predict the optimal control voltages using the inverse part of the compound model. (b) The logo consisting of measured polarization states from heralded single-photon preparation, applying the predicted control voltages. (c) Polarization states obtained in the process of remote single-photon state preparation.

Recently, we introduced an all-fiber, single-shot polarization sensor that reconstructs complete polarization states, including partial polarization, at single-photon sensitivity [Bielak2025]. The sensor maps the polarization state of incident light onto a coarse speckle field generated by a short few-mode fiber and reads out only sparse samples of that field, see Fig. 5.6. Specifically, we collected light at the front face of a 5 cm step-index SMF28 fiber (NA 0.14, 8.2 µm core) chosen to keep the sensing area minimal while still supporting several interfering modes; other few-/multimode fibers, including shorter ones, would also work provided they generate polarization-

dependent speckle. After ~5 mm free-space propagation, seven GIF625 multimode fibers (62.5 µm cores) arranged in a 375 µm-diameter honeycomb array sampled the speckle and routed the signals to independent single-photon detectors (SPADs). We normalized the per-channel counts into a count distribution used as the sensor measurement outcome. The experimental scheme had two basic interchangeable front-end configurations: (i) a twisted-nematic liquid-crystal (TNLC) device driven by calibrated voltages (see previous part of this chapter) for generating training/validation states, and (ii) application specimens (e.g., birefringent samples). In all cases, the modified beam was coupled into the same 5 cm few-mode fiber, with the speckle sampled and detected as above.

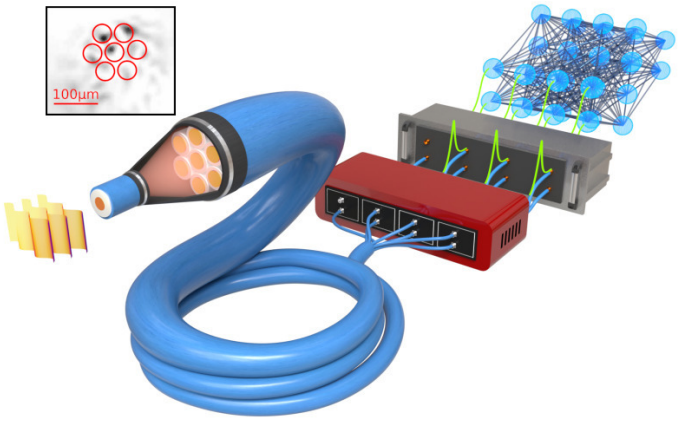


Figure 5.6. A visual representation of the all-fiber polarization sensor [Bielak2025]. Polarized light emitted from a specimen is collected by a short piece of a few-mode fiber. The intermodal interference generates a granular speckle pattern at the output (inset; inverted intensities). Several isolated samples of this pattern impinge on a fiber array and propagate to a corresponding number of single-photon detectors. The recorded detections are electronically processed using deep-learning methods to characterize the incident polarization state.

For data acquisition and calibration, we prepared ~30,000 polarization states uniformly over the Bloch sphere using a fiber-coupled 0.8 µm CW diode laser attenuated to the single-photon level, a linear polarizer, and the TNLC device to apply targeted unitary transformations. We calibrated the TNLC device via bidirectional learning [Vašinka2022]. We then measured each prepared state with our sensor to obtain the associated count distribution, typically collecting on the order of 1.5×10^5 detections per channel in 50 ms. To cover mixed states, we extended the dataset by weighted-summing count distributions of orthogonal preparations, operationally equivalent to measuring the corresponding partially polarized states, thereby populating the full volume of the Bloch sphere for downstream model training and evaluation.

We used the dataset to train a fully connected deep neural network for reconstructing the polarization state given the associated count distribution. Specifically, we trained a fully connected neural network (four hidden layers) to map normalized counts to a physically valid polarization state by outputting a triangular 2×2 matrix whose Cholesky product, after trace normalization, yielded the density matrix (polarization coherence matrix). We optimized the network with the Adam optimizer on experimentally acquired data and selected hyperparameters using a separate validation set that maximized state fidelity; performance was finally reported on a

held-out test set. This end-to-end scheme let us work and score in the state space (via fidelity/infidelity), rather than in the arbitrary space of detector counts, and returned either the coherence matrix or its equivalent Bloch parameters for downstream analysis. The trained sensor achieved an average infidelity $\sim 8\times10^{-4}$ (10th–90th percentile $[5\times10^{-5}, 2\times10^{-3}]$), close to the preparation-limited accuracy and corresponding to a ~0.01 error per Stokes parameter. As an external benchmark, we scanned a 1951 USAF birefringent target and compared the output of our sensor to a rotating-waveplate polarimeter, finding an average fidelity ≈0.987 between the two methods and estimating a spatial resolution of ~6(5) μm from edge fits.

We evaluated speed and data throughput by sweeping the acquisition window and photon budget. The sensor reached 0.999 average fidelity at 33 Hz, enabling accurate real-time operation. Alternatively, we measured >2,000 states per second with only 10^4 detected photons in a single shot while maintaining >0.99 fidelity. Crucially, accuracy is governed by the number of detected photons, not the nominal repetition rate; substantially higher rates are attainable by replacing SPAD detectors with faster SNSPDs.

In open-lab conditions and without special stabilization, the sensor with a 5 cm few-mode fiber remained stable for a week without recalibration. Using a shorter 12 mm long few-mode fiber encased in a ceramic ferrule further extended stability to >1 month, with a measured drift slope of 4.7×10^{-4} day^{-1} ($\sim 2\times10^{-5}$ h^{-1}). The robustness stems from the short propagation length, the rigid ferrule mount, and the use of a few-mode, coarse speckle that is less sensitive to bending and thermal gradients. Because sufficient intermodal mixing occurs over only a few hundred wavelengths, the sensing fiber can be shortened to <1 mm to improve stability further. When needed, fully automated recalibration restores peak performance within an hour.

We demonstrated the application potential of the all-fiber polarization sensor in three settings. First, we show imaging capability on a dense connective tissue (AmScope PS25W). We replaced the calibration stage with the specimen, positioned the few-mode fiber tip near the tissue, raster-scanned the area with motorized stages, and measured the polarization at each point as Bloch parameters (renormalized Stokes parameters). The polarization map, rendered as a false-color RGB image, revealed the characteristic anisotropic patterns that are invisible in a conventional intensity image and agreed with a stand-alone polarization microscope (with crossed polarizers), see Fig. 5.7. We then targeted dynamic, weakly scattering biological specimens by measuring a floating Actinoptychus heliopelta diatom. Using the same optical arrangement, we tracked the spatial modulation of all three Bloch parameters as the organism traversed the field of view, distinguishing the diatom signal from the background. This analysis can count the number of living specimens, estimate their size and morphology, and classify them based on their birefringent properties.

Beyond biological samples, the sensor can also find applications for fast in-situ material inspection, local strain analysis, and microcrystal growth monitoring. We showcased high-speed in-situ polarimetry by recording a voltage-driven transition in twisted-nematic liquid crystals, see Fig 5.8. With 5 ms acquisition windows, the sensor captured a smooth transition from horizontal to vertical polarization that completed in ~50 ms, resolving the full trajectory of all Bloch components in real time—capability that is inaccessible to rotating-waveplate systems. Together, these

studies establish the sensor as a compact tool for static imaging, dynamic biological monitoring, and fast in-situ material characterization at low photon fluxes.

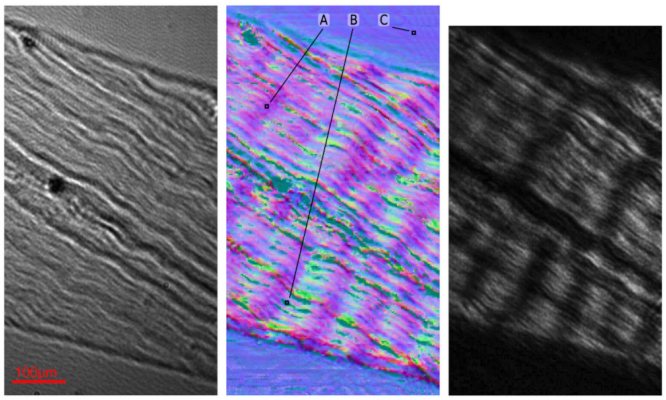


Figure 5.7. Visualization of dense connective tissue [Bielak2025]: (left) an intensity image, (middle) a scan using the all-fiber polarization sensor, (right) an image using a stand-alone polarization microscope. The resulting Bloch parameters of the all-fiber scan are represented as an RGB false-color image. The three highlighted pixels characterize the purple and green segments in the polarization structure alongside a reference background polarization. Their respective Bloch parameters are A = (0.34, −0.43, 0.82), B = (−0.90, 0.05, 0.06), and C = (−0.04, −0.18, 0.98).

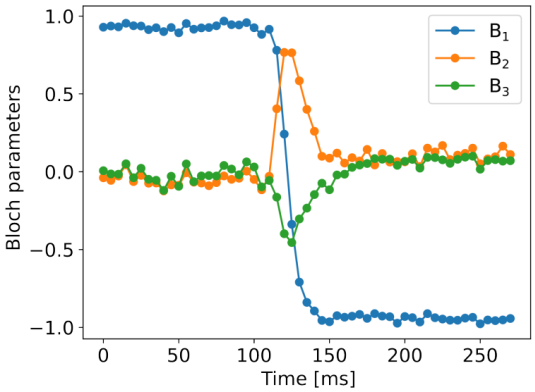


Figure 5.8. Time-resolved measurement of a polarization transition induced by voltage-controlled twisted nematic liquid crystals [Bielak2025]. The all-fiber sensor continuously captures the complete polarization information with a 5 ms acquisition window at the approximate power level of 10^5 detected photons. The transition from horizontal to vertical polarization is represented through the evolution of the three Bloch (normalized Stokes) parameters, each shown in a different color.

6. Quantum correlation sensing

Quantifying entanglement in a physical system is central to both foundational studies and quantum technologies, yet conventional approaches often rely on a priori knowledge of the system or resource-intensive procedures such as indirect quantum state tomography and collective measurements. Recently, we addressed this gap by using deep learning to quantify entanglement from local, incomplete measurement data. In this chapter, we will present our breakthrough results at the boundary of quantum physics and AI. We will introduce the application of deep artificial neural networks to directly estimate the degree of entanglement and quantum correlations in generic, partially mixed states from an informationally incomplete set of local projective measurements. The networks are trained entirely on numerically simulated data to prevent technical biases and speed up the workflow. They deliver quantitative estimates with errors up to an order of magnitude lower than state-of-the-art tomographic methods, even under substantial undersampling, thereby reducing measurement overhead while avoiding full quantum state reconstruction. Beyond fixed measurement settings, we will discuss a convolutional architecture that ingests heterogeneous measurement data and performs independently of the specific measurement device or chosen projectors, enabling estimation without retraining across different experimental scenarios. Our approach does not require a full description of the underlying quantum state, scales favorably with system complexity, and integrates naturally with existing experimental pipelines. We will further demonstrate immediate practical value for certification and benchmarking of entanglement sources, showcasing results with photonic states generated via spontaneous parametric down-conversion and with a semiconductor quantum dot.

The presented results are based on the work:

- D. Koutný, L. Ginés, M. Moczała-Dusanowska, S. Höfling, Ch. Schneider, A. Predojević, and M. Ježek, Deep learning of quantum entanglement from incomplete measurements, Sci. Adv. 9, eadd7131 (2023).

The publication was selected for the journal cover.
Media coverage: Physics World, PHYS.ORG, La Nueva Espana, etc.

SCIENCE ADVANCES

VOLUME 9 | ISSUE 29 | 21 JUL 2023

< PREVIOUS | ALL ISSUES | NEXT >

ONLINE COVER: Unlocking the quantum enigma—using deep neural networks to quantify entanglement. Koutny *et al*. experimentally demonstrate quantum entanglement quantification using machine learning, an approach that could outperform state-of-the-art methods with greater accuracy from fewer measurements. This work may provide deeper insights into quantum entanglement.

Credit: Monika Tomanová

Even in a well-understood system, such as a pair of qubits, a reliable quantification of entanglement requires full state tomography [Fiurášek2002, Carmeli2016, Lu2016, Yu2020]. In photonics this means recording all combinations of the three local Pauli bases, i.e., 36 local projectors. Omitting some projectors in this measurement scheme degrades tomographic reconstructions and, in turn, entanglement estimates. We therefore proposed to bypass state reconstruction altogether and to learn the degree of entanglement directly from incomplete local data using deep neural networks

(DNNs) [Koutný2023]. To demonstrate the advantage of the DNN approach, we used two target quantifiers tailored to the regimes we studied. For bipartite qubit systems, we quantify entanglement with the concurrence, widely used in experiments and bounded between zero (separable) and one (maximally entangled) [Horodecki2009]. For scalability beyond two qubits, we use the quantum mutual information matrix, which captures all two-party correlations and generalizes naturally to multipartite systems [Valdez2017]. These choices let us evaluate both canonical two-qubit sources and settings where a bipartite monotone is not readily applicable.

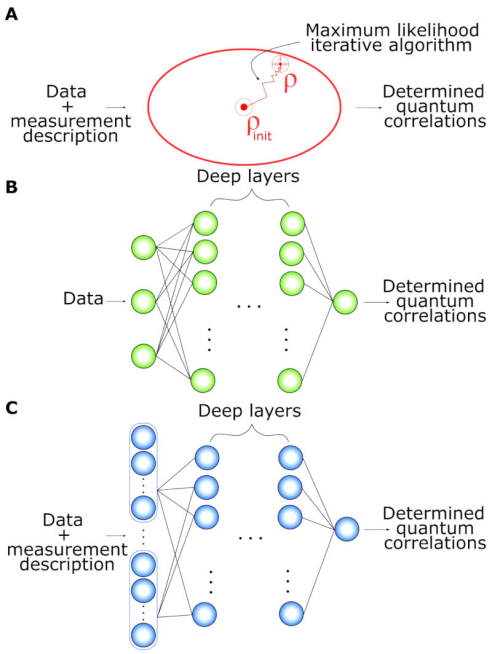


Figure 6.1. Schematics of the three methods used to infer the quantum correlations [Koutný2023]. (A) The maximum likelihood algorithm (MaxLik) finds the most likely quantum state ρ based on the measured data and an initial guess ρ_{init}. (B) The green DNN represents a fully connected neural network that infers directly the concurrence and the mutual information from specific measurements (specific measurement projectors), whereas (C) the blue DNN works with an arbitrary set of measurement projectors. The input for the former is the measured data. The measurement-independent DNN has a convolutional first layer and takes as input both the data and the measurement description.

We implemented three different approaches to determine the concurrence and the mutual information from an incomplete set of data. We show them schematically in Fig. 6.1. First, as a baseline we implemented maximum-likelihood (MaxLik) tomography [Ježek2003, Hradil2004]: starting from a maximally mixed seed, the algorithm iteratively updated the density operator to maximize the likelihood of the observed frequencies and then computed the desired quantifier from the reconstructed state. For incomplete measurements, we adopted a correct renormalization of the MaxLik approach [Hradil2004]. This provided a reconstruction-then-quantify benchmark that is known to work excellently with informationally complete data but to deteriorate when measurement data are incomplete. We also compared the MaxLik to other reconstruction methods, such as maximum-likelihood maximum-entropy and semidefinite programming with L1 and L2 norms. However, the MaxLik provides the most accurate predictions across the board, see Supplementary Materials of [Koutný2023]. Second, we developed measurement-specific DNNs, i.e., fully connected networks trained for a fixed,

predefined set of projectors. These networks directly output the concurrence or mutual information from the measured data. Third, we introduced a measurement-independent DNN whose first layer is convolutional and inputs both the measurement outcome and a machine-readable description of the corresponding projector; this architecture allowed us to predict from arbitrary subsets of settings without retraining.

For two-qubit and fixed-measurement models, we used fully connected networks with seven hidden layers containing (60, 40, 30, 30, 30, 20) neurons and on the order of tens of thousands of trainable parameters. We instantiated many such models, each matched to a particular subset of k measurement projections, from the full quorum of 36 down to 4 projections. For the measurement-independent network, we paired each measurement probability with a vectorized description of the corresponding projection operator and used a convolutional first layer (50 kernels) with a stride chosen to prevent crosstalk between probability-projector pairs, enabling the same network to accept arbitrary subsets by zero-padding missing entries. We trained and validated the DNNs entirely on simulated data. We generated large training datasets by sampling approximately one million random two-qubit states that broadly covered partially mixed states: four-fifths followed the Bures-induced ensemble [Bures1969] and one-fifth were random Haar-pure states [Mezzadri2007] mixed with white noise. For each state we computed the ideal outcome probabilities for the chosen Pauli projectors and the target quantifier (concurrence or mutual information). This large, purely simulated corpus ensured that the learned mappings were not biased by device-specific noise or drifts in any particular laboratory setup. We shuffled and split the data into 800,000 training and 200,000 validation examples, reserving an independent 5000-state test set. The training minimized the mean absolute error (MAE) between the predicted and true quantifiers, averaged over the test set, using Nesterov-accelerated adaptive moment estimation (NAdam). We also averaged over several combinations of measurement projections. Because the number of possible measurement subsets of k projections selected from all 36 Pauli projections grows combinatorially, we evaluated the networks on randomly sampled subsets rather than exhaustively enumerating all combinations.

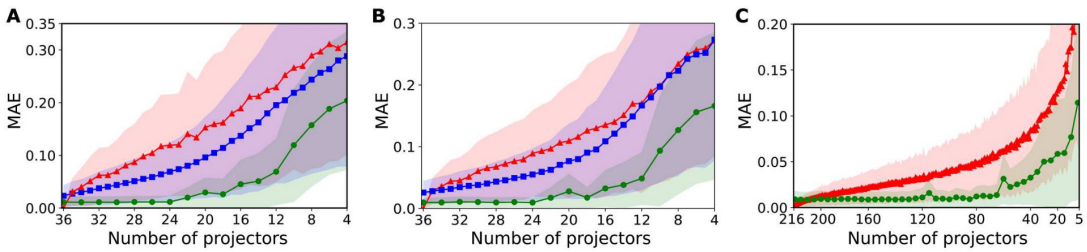


Figure 6.2. Entanglement quantification error for the two- and three-qubit systems [Koutný2023]. The mean absolute error (MAE) versus the number of measurement projections for (A) two-qubit concurrence, (B) two-qubit mutual information, and (C) three-qubit mutual information. Red triangles depict MAE for the MaxLik, blue squares stand for the values of MAE computed from measurement-independent DNN, and, lastly, green circles represent the values of MAE computed from measurement-specific DNNs. The uncertainty regions are depicted in the corresponding colors and may overlap. The DNNs outperform the MaxLik approach in terms of entanglement quantification accuracy and its consistency, given by smaller errors and uncertainty intervals, even for substantially incomplete measurements.

For performance assessment we compared MaxLik, measurement-specific DNNs, and the measurement-independent DNN on identically constructed test sets, always using the same incomplete local data. To capture variability across projector choices, we trained 12 independently initialized measurement-specific networks for each $k = 36,\ldots, 4$ (with increment of 2) and reported the average and standard deviation of their MAEs. For MaxLik and the measurement-independent DNN we averaged MAEs over hundreds of random subsets. The results are shown in Fig. 6.2 (A,B) for concurrence and mutual information, respectively. With full data, MaxLik essentially recovered the true state and yielded negligible quantifier error, as expected. However, MaxLik error increased rapidly once even a few measurement projections were omitted. In contrast, our measurement-specific DNNs remained accurate deep into the undersampled regime. With only two-thirds of the settings (24 projectors), they achieved MAEs around the percent level for both concurrence and mutual information, approximately an order of magnitude better than MaxLik under the same incompleteness, and with substantially tighter uncertainty bands across random subsets. The measurement-independent model, while typically less accurate than the measurement-specific networks, consistently outperformed MaxLik over a broad range of k and offered the practical benefit of working with arbitrary measurement choices. It is useful for quick, on-the-fly estimates before optionally investing computation time in a projector-specific model.

We generalized our approach beyond two qubits by switching from concurrence to the mutual information as the correlation quantifier and by demonstrating that the DNN pipeline scaled naturally to larger systems. We began with three qubits, where the mutual information forms a 3×3 adjacency matrix with three independent elements corresponding to the three possible bipartitions [Valdez2017]. A full tomographic determination of these three numbers would have required measuring all local Pauli settings on each qubit, i.e., 216 projections. Instead, we built measurement-specific DNNs that mapped incomplete sets of measured probabilities directly to the three components of the mutual-information vector, following the same design we had used for two-qubit mutual information, see Fig. 6.2 (C). Near the tomographically complete regime, our DNN predictions matched the MaxLik baseline on average. However, when we reduced the data to roughly one third of all three-qubit projections, the measurement-specific DNNs achieved about a fivefold lower MAE than MaxLik, averaged over random states and random subsets, indicating that the deep-learning-based approach retains far more information about correlations.

To assess computational demands and scaling, we analyzed the network sizes at a fixed measurement fraction (one quarter of all local Pauli projectors). The two- and three-qubit measurement-specific models used approximately 37,000 and 42,000 trainable parameters, respectively. We pushed this further by training four- and five-qubit measurement-specific networks (again with one quarter of all projectors), which required about 69,000 and 231,000 parameters. Because memory limited the dataset size at higher qubit numbers, we adopted incremental learning [Carpenter1991, vandeVen2022]. After convergence on one batch of 100,000 examples, we generated a new batch and continued training from the best checkpoint. Even our largest models at this stage remain modest by modern standards: the five-qubit measurement-specific network has slightly more than 230,000 parameters, and the entire training on a few million simulated samples fit on a single consumer-grade

GPU. Once trained, the inference (i.e., mapping measurement data to an entanglement estimate) is extremely fast, providing several orders of magnitude speedups over reconstruct-then-quantify pipelines such as the MaxLik and semidefinite-programming-based methods, which must be rerun from scratch on each new dataset.

The higher-qubit DNNs not only remained tractable but also improved the advantage over the MaxLik. Summarizing across system sizes at the same fractional sampling, the DNN MAE was lower than MaxLik by factors of about 2.2, 3.0, 3.8, and 4.3 for two, three, four, and five qubits, respectively (see Table 6.1). Based on this trend, we expect that if one keeps the ratio of MaxLik accuracy to DNN accuracy fixed, the required fraction of projectors relative to full tomography should decrease with system size.

Number of qubits	MAE		Ratio of MaxLik and DNN MAEs
	MaxLik	DNN	
2	0.20 ± 0.16	0.09 ± 0.09	2.2
3	0.068 ± 0.055	0.023 ± 0.020	3.0
4	0.019 ± 0.014	0.005 ± 0.001	3.8
5	0.039 ± 0.032	0.009 ± 0.001	4.3

Table 6.1. The summary of the mutual information quantification from incomplete measurements consisting of one-fourth of all possible Pauli projectors in each case [Koutný2023]. The MaxLik and the measurement-specific DNNs are compared up to five-qubit quantum systems. The ratio of the mean absolute errors (MAEs) of the methods shows an increasing improvement in the performance of the DNN approach for entanglement quantification in higher-dimensional systems.

All the results presented so far were based on measurement probabilities. To analyze the sensitivity of entanglement quantification to sampling errors, we repeated the evaluation using data sampled from multinomial distribution based on the ideal measurement probabilities. The total number of samples corresponds to the number of measurement runs, i.e., state copies used. Figure 6.3 shows the performance of measurement-specific DNNs for two-qubit systems and 1,000 and 10,000 measurement samples, panels (a) and (b), per all used measurement projections, and the ideal probabilities (c). The results remain on par with those obtained in the case of the ideal measurement probabilities (infinite sampled data). Surprisingly, for a small number of the total detections (e.g., one thousand samples), the measurement-specific DNNs outperform Maxlik and other methods even for the full 36-projector measurement.

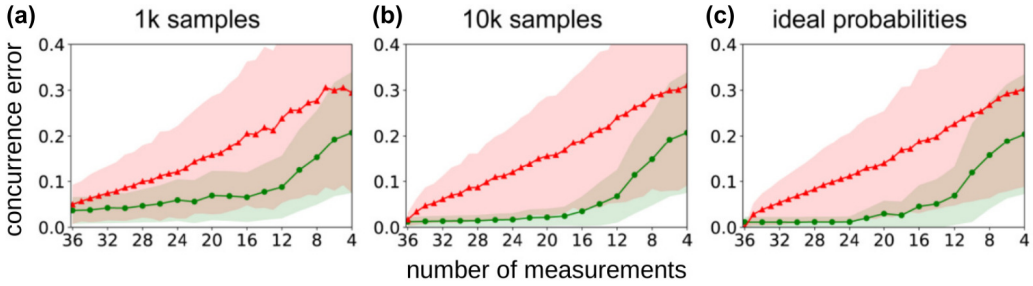


Figure 6.3. Two-qubit entanglement quantification error for various numbers of simulated measurement runs [Koutný2023]. The concurrence MAE versus the number of measurement projections for (a) 1,000 and (b) 10,000 measurement runs. For comparison, panel (c) shows MAE for the ideal measurement probabilities (same as in Fig. 6.2 A). Red triangles and green circles depict MAE for the MaxLik and measurement-specific DNNs, respectively. The uncertainty regions are depicted in the corresponding colors.

Furthermore, we validated the DNN-based quantification on real experimental data under non-ideal conditions and limited statistics. First, we measured polarization-entangled photon pairs produced by a continuous-wave spontaneous parametric down-conversion (SPDC) source, see Fig. 6.4 (a). The source used a type-II, collinear beta-barium borate (BBO) crystal at 810 nm; we interfered the two correlated photons on a balanced beam splitter to conditionally prepare the singlet Bell state, see Fig. 6.4 (b). For full reference data, we performed complete polarization tomography, i.e., all 36 projective settings formed from local projections onto H/V, D/A, and R/L states, using a standard analyzer (quarter- and half-wave plates, a polarizer, single-photon detectors) and coincidence detection. To create controlled degradations, we varied the coincidence window and injected a weak classical field into one arm as a tunable noise channel with an attenuated laser diode [Straka2015]. This allowed us to record datasets ranging from nearly maximally entangled to weakly entangled outputs. The specific experimental datasets we analyzed in [Koutný2023] had concurrences of 0.985 ± 0.001 and 0.201 ± 0.002.

On these SPDC datasets, we quantified entanglement with the DNNs and MaxLik while subsampling the number of measurements, see Fig. 6.4 (d,e). Across projector budgets from near-complete down to far-incomplete, both DNN approaches outperformed MaxLik for the nearly pure and the mixed states. Measurement-specific DNNs remained accurate deep into undersampling: their MAE stayed below about 0.04 even with only 14 projectors. The measurement-independent DNN also surpassed MaxLik for the generic partially mixed case at essentially any number of projectors. The only notable exception appeared for the almost maximally entangled state: here the MaxLik outperforms the measurement-independent DNN. We attribute this to data sparsity: when randomly choosing few projectors on a nearly pure state, many outcomes had near-zero counts, which limited the predictive strength of the measurement-independent model (as the unused measurement inputs are also filled by zeros). At the same time, the MaxLik approach is biased toward pure states in the case of heavily undersampled data [Schwemmer2015, Silva2017], and the positivity constraint tends to produce sparse (low-rank) states [Kalev2015], which can artificially inflate concurrence and reduce its reported error. This explains why the MaxLik curve could look optimistic in this corner case.

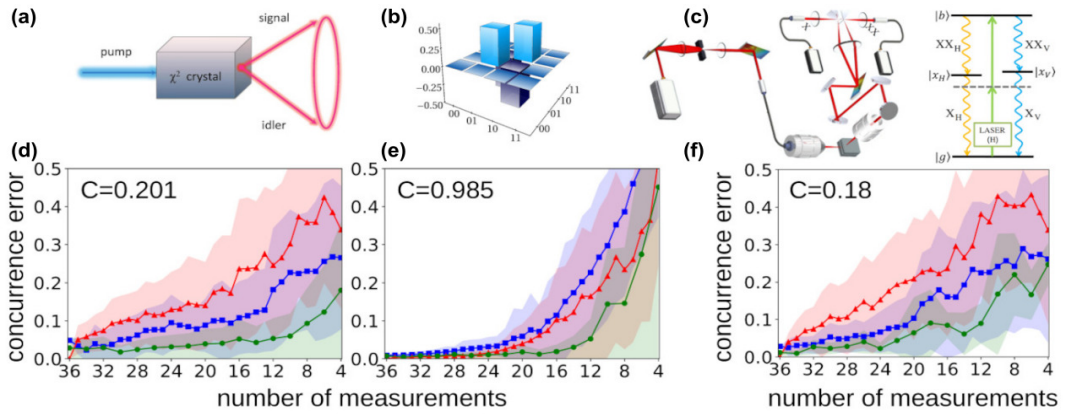


Figure 6.4. (a) Spontaneous parametric down-conversion (SPDC) source and (b) the density matrix plot of the produced entangled state reconstructed by the full quantum state tomography. (c) Semiconductor quantum dot source based on the biexciton–exciton cascade. (d-f) Performance of the MaxLik and DNN-based approaches for experimental datasets. We show the dependence of the concurrence MAE on the number of measurement projections for (d-e) the SPDC source and (f) the quantum dot source. The concurrence of experimentally prepared quantum states was determined from the full MaxLik tomography to (d) 0.201 ± 0.002, (e) 0.985 ± 0.001, and (f) 0.18 ± 0.01. The MAE for the measurement-specific DNNs is depicted in green circles, for measurement-independent DNN in blue squares, and for the MaxLik approach in red triangles.

The second experiment was based on a resonantly driven semiconductor quantum dot that emitted polarization-entangled photon pairs through the biexciton–exciton cascade, see Fig. 6.4 (c). The device featured a quantum dot embedded in a circular Bragg grating cavity for efficient collection [Ates2012]. The quantum dot was excited via two-photon resonant excitation of the biexciton [Jayakumar2013]. The excitation pulses were derived from a pulsed 80-MHz repetition rate Ti:Sapphire laser. We spectrally filtered out pump scatter, separated the exciton and biexciton lines, and coupled both paths to single-mode fibers. We analyzed polarization with the same tomography hardware and protocol as in the SPDC system. In this solid-state source, the observable entanglement was primarily limited by the residual fine-structure splitting (FSS); the data used in the paper had concurrence 0.18 ± 0.01. On the quantum-dot data, both DNN variants again outperformed MaxLik on average over incomplete measurement sets, see Fig. 6.4 (f). This was a stringent test because our networks had been trained only on ideal, simulated probabilities from noiseless states and measurements. Nevertheless, their inference transferred to real, imperfect experiments. Taken together, the SPDC and quantum-dot results show that our learning-based approach sustained lower errors than indirect reconstruction-based pipelines over a wide range of incomplete measurements and noise conditions.

In summary, we developed a learning-based route to quantify entanglement directly from incomplete local measurements. We used two estimators: a measurement-specific DNN trained for a fixed set of projectors and a measurement-independent DNN that accepted arbitrary projectors by encoding both probabilities and projector descriptions. Against MaxLik tomography, both models achieved lower errors once measurements became incomplete; the measurement-specific DNN often reduced MAE by ~10×, while the measurement-independent model traded a small accuracy loss for flexibility. We generalized beyond two qubits by predicting the three-entry

mutual-information vector for three qubits, obtaining ~5× lower MAE than MaxLik with about one-third of all settings. At a fixed fractional budget (¼ of Pauli projectors), the advantage of the DNNs over MaxLik grew with system size. We also validated the DNNs on experiments: polarization-entangled pairs from an SPDC process and a resonantly driven quantum dot. Across all undersampling levels, measurement-specific DNNs outperformed MaxLik. Simultaneously, the DNN inference was fast, enabling real-time estimation.

Looking ahead, AI will be central to characterizing emergent phenomena, discovering new quantum phases, refining theories of strongly correlated systems, surpassing conventional numerics, and ultimately probing fundamental questions in quantum science [Du2025]. Systematic progress along these fronts will make AI an essential tool for unlocking the computational potential of quantum technologies [Acampora2025].

Missing pieces include scaling to much larger qubit numbers, full device independence (our measurement-independent model still used characterized projectors and was tested only for two qubits), dimensional independence (i.e., a single model applicable to various numbers of qubits), and confidence (uncertainty) prediction. Currently we push the DNN framework to regimes of more than 12 near-pure qubits and more than 6 generic partially mixed qubits. We also explore various different entanglement metrics, going beyond pair-wise correlations and targeting genuine multipartite entanglement (convex-roof bounds). We have already fixed the issues of zero-padding measurement-device-independent models, and we currently discuss more AI-native ways toward device independence, such as symmetry-aware and graph-based architectures. Furthermore, we aim to predict uncertainties together with entanglement quantifiers within a single DNN model.

Complete device-agnostic modeling is perhaps the largest challenge. Are deep neural networks predicting a complex physical quantity able to generalize completely beyond the scope of the measurement devices used for producing the training data? We hear similar questions almost every time when presenting this project. This motivates another research project, fully classical but also very complex, discussed in the next chapter.

7. Device-agnostic super-resolution imaging

The spatial resolution of imaging and sensing is fundamentally constrained by diffraction of a signal used to transmit information from a sample to a detector. It does not matter whether we use photons or electrons, their wavelengths and the aperture of imaging systems limit the details we can observe. Super-resolution lets us distinguish structures smaller than the classical diffraction limit, revealing otherwise blurred details. Accessing this hidden spatial information enables new observations and more rigorous tests across biology, material science, and astronomy. Current super-resolution methods, however, strongly rely on prior calibration and detailed knowledge of the imaging setup. Even when calibrated, the system response varies across the field of view and over time, degrading the robustness and accuracy of these device-specific approaches. Recently, we developed a device-agnostic deep-learning model that reconstructs super-resolved images from a single diffraction-limited camera frame without calibration or knowledge of the optical system. We demonstrated superior reconstructions to classical deconvolution and current device-specific deep-learning methods. The model was tested on astronomy and single-molecule localization microscopy data, and on dense samples of organic molecules at room temperature and quantum dots at cryogenic temperatures. It recovered sub-diffraction structures without any system-specific inputs, indicating a path toward universal, calibration-free super-resolution across instruments.

The presented results are based on the preprint:

- D. Vašinka, F. Juráň, J. Běhal, and M. Ježek, From stars to molecules: AI guided device-agnostic super-resolution imaging, arXiv:2502.18637, bioRxiv:2025.02.25.640182 (2025).

The following recent preprints are also briefly summarized:

- A. Dostálová, D. Vašinka, R. Stárek, and M. Ježek, Calibration-free single-frame super-resolution fluorescence microscopy, arXiv:2505.13293, bioRxiv:2025.05.20.655080 (2025).
- D. Vašinka, J. Lee, C. Stalker, V. Mitryakhin, I. Solovev, S. Stephan, S. Höfling, F. Eilenberger, S. A. Tongay, C. Schneider, M. Ježek, and A. Predojević, Universal super-resolution framework for imaging of quantum dots, arXiv:2510.06076 (2025).

Numerous research and real-world applications rely on imaging point-like or single-emitter sources. In biology, single-molecule localization microscopy reveals cellular structures at the nanoscale [Rust2006, Betzig2006, Balzarotti2017, Lelek2021, Chen2022, Bodén2024]. In quantum physics, super-resolution characterizes quantum dots [Lidke2005] and images cold atoms in optical lattices [McDonald2019, Impertro2023]. Astronomy benefits by resolving individual stars and galaxies [Cava2018, Zhang2024].

Super-resolution can be obtained with linear inverse or Bayesian reconstructions [Bertero2021], tomographic synthesis [Chen2022, Luo2015, Bianco2023], and blinking-based localization [Rust2006, Betzig2006, Balzarotti2017]. Deep learning

methods have been extensively used [Rivenson2017, Nehme2018, Moen2019, Barbastathis2019, Speiser2021, vonChamier2021], and AI is even discovering new microscope designs [Rodríguez2024]. Broadly, methods fall into two categories: reconstruction and parameter estimation. Reconstruction aims to directly restore the whole super-resolved image [Bertero2021, Nehme2018], while estimation focuses on extracting key parameters and features, such as emitter localization [Speiser2021, Smith2010, Hekrdla2025]. The localization performance rapidly decreases for images with higher number of overlapping emitters, which typically limits its application to stochastically blinking or photoactivated samples. In the following discussion, we will focus exclusively on the former, more universal approach: image reconstruction.

Accurate single-image reconstruction hinges on a correct imaging model, especially the point spread function (PSF). PSF mismodeling degrades reconstructions; reliable PSFs typically require dedicated calibration with isolated emitters and can vary across the field of view [Booth2015, Xu2020]. Blind reconstruction eliminates explicit calibration by assuming a parametric PSF and estimating it jointly with the object, but the resulting problem is non-convex, initialization-sensitive, and often underperforms well-calibrated non-blind approaches in practice [Campisi2007]. Deep-learning reconstructions inherit similar device dependence: the learned mapping is tied to the training PSF, photon budget, background statistics, and emitter density. Deploying on a different imaging system usually requires new calibration data and retraining, making adaptation costly [Rivenson2017, Nehme2018, Belthangady2019, Speiser2021].

In [Vašinka2025a], we developed a device-agnostic modeling network (DAMN) that reconstructs super-resolved images of point-like emitters from a single intensity frame, see Fig. 7.1. The model needs no optical-parameter inputs, calibration, or retraining, and it accepts images of arbitrary size. DAMN is fully convolutional: 35 hidden layers with 71 channels and 7×7 kernels, each followed by LeakyReLU [Maas2013, Xu2015] (slope 0.05) and 0.01 dropout [Srivastava2014]. An optional variant adds three upsampling layers that double each image dimension, yielding an eightfold increase in pixel count. Inputs are normalized to unit sum, and the output uses a softmax to keep non-negativity and the normalization constraint. The presented model architecture results from extensive optimization utilizing a mesh adaptive direct search using the Nomad library [Nomad2022], as well as numerous additional manual adjustments.

We trained DAMN on numerically generated pairs of ground-truth objects and their resolution-limited images, see Fig. 7.1. The simulation spanned wide ranges of emitter power (1–10^5 photons), background (1–100), emitter concentration (5–500 per 50×50 field), and PSF width ($10^{-0.25}$–$10^{1.25}$ px ~ 0.5–17.75 px), using randomly either Gaussian or Airy shapes. The generation process of a single simulated data sample followed these steps. First, we generated the emitter concentration, followed by assigning each emitter its power and pixel position in the image. Next, we performed a convolution with the point spread function to simulate the effects of finite resolution. Subsequently, shot noise was added to each pixel of the image. For the first time in the emitter visualization community, data with such a broad scale of imaging parameters were generated and used to train an artificial neural network. Also, the level of network complexity is unprecedented in the field of emitter

reconstruction and localization. Nevertheless, the model and dataset remain relatively modest compared to modern large language models, suggesting further potential to integrate additional features and parameters, such as larger upsampling, PSF asymmetry and aberrations, etc.

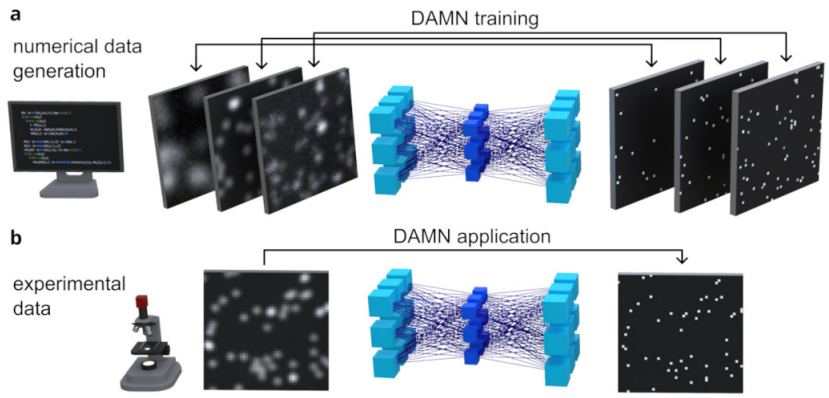


Figure 7.1. Schematic representation of the DAMN approach and its application to intensity images of point-like emitting sources [Vašinka2025a]. (a) The model is trained using numerically simulated data pairs comprising resolution-limited noisy images alongside their super-resolved counterparts. Each training sample represents a unique combination of underlying optical parameters, such as the width of the point spread function and the signal-to-noise ratio. (b) Following training, this model is applied to enhance the resolution of experimental images acquired using a real-life imaging system.

The training used backpropagation with mean-square-error (MSE) loss (batch size 128) and the Adam optimizer [Kingma2015], while mean-absolute-error (MAE) monitors validation. The model has over 8 million trainable parameters and was trained incrementally [vandeVen2022] on nearly three million simulated images (with a 25% validation set split) so that the data distribution continually refreshed. As a result, a single network (with one set of weights) generalizes across imaging devices and remains robust to parameter changes, spatial PSF non-uniformity, and temporal instabilities such as drift.

First, we evaluated the DAMN performance on simulated data. We compared DAMN to Richardson–Lucy deconvolution (Bayesian, uniform prior) supplied with the true PSF. Performance was measured on a held-out simulated test set (not used for training) as the MAE between each reconstruction and its target, averaged over entire 50×50 images, see Fig. 7.2. Curves show averages with 90% confidence intervals. Panels (a–c) plot MAE versus (a) emitter power, (b) PSF width, and (c) emitter concentration. Panel (a) contains dual horizontal axes representing the varying emitter power using both the signal-to-noise ratio (SNR) and the peak-to-noise ratio (PNR). Across all three, DAMN (red) outperforms Richardson–Lucy (green) by up to two orders of magnitude despite using no device information. Panel (d) varies the PSF shape continuously from Airy to Gaussian at equal width. The performance of both methods remains approximately constant during this transition. Richardson–Lucy is expected to be stable because it always receives the correct PSF profile. DAMN was trained only on the endpoints (pure Airy or pure Gaussian), yet generalizes smoothly to intermediate shapes, indicating robustness to unseen PSF profiles.

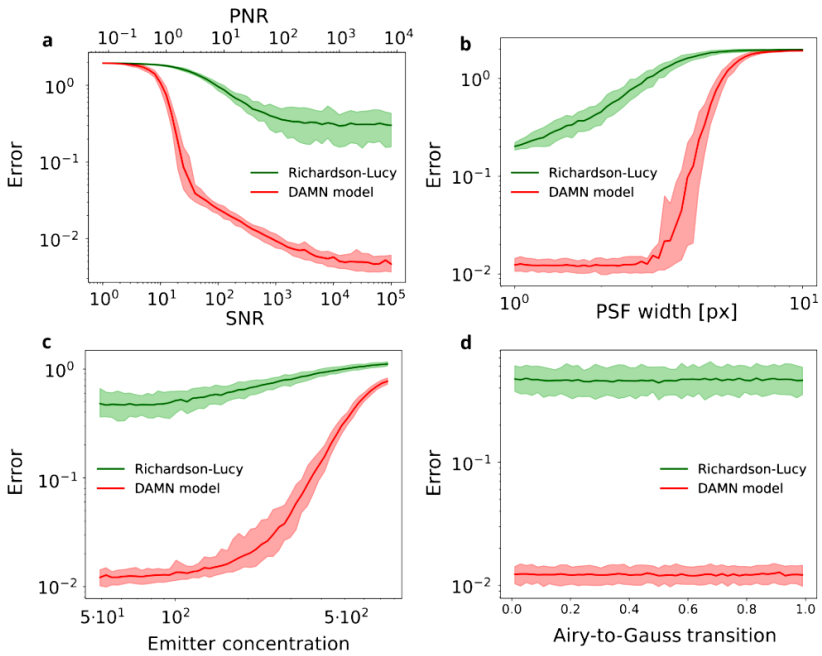


Figure 7.2. The dependence of the mean absolute error on (a) the signal-to-noise ratio (SNR), (b) the width of a Gaussian point spread function (PSF), (c) the number of emitters in the image (concentration), and (d) the continuous transition between an Airy and a Gaussian PSF, respectively. The resulting averages of the Richardson-Lucy algorithm (green) and the DAMN model (red) are accompanied by their 90% confidence intervals over the test set. Panel (a) includes a secondary horizontal axis recalculating the SNR values to the peak-to-noise ratio (PNR). Across all panels, the DAMN model consistently outperforms the Richardson-Lucy deconvolution by up to two orders of magnitude. The optical parameters not investigated in a given panel have the following values: SNR = 500, the average noise intensity = 10, the concentration = 50, and the Gaussian PSF with $\sigma = 2$ px.

We also validated DAMN on experimentally acquired data using a custom microscope that generates paired low-resolution inputs and the corresponding ground truth, see Fig. 7.3. A digital micromirror device (DMD) encoded a binary mask defining emitter positions. Preparation optics re-imaged this mask into the sample plane to create point-like emitters. A low-resolution microscope then formed a resolution-limited camera frame, further down-scaled to 50×50 to match the masks. Each camera image was paired with its DMD mask, enabling quantitative evaluation. Typical operating parameters were SNR ≈ 2300, average background ≈ 10, an Airy-shaped PSF with $\sigma \approx 2.05$ px (FWHM ≈ 3.4 px), and Rayleigh limit $R \approx 3.9$ px. We recorded series across emitter concentrations and acquired a separate high-SNR low-concentration calibration images only for the Richardson–Lucy algorithm to estimate the PSF. For more details on the experiment, see Methods section in [Vašinka2025a].

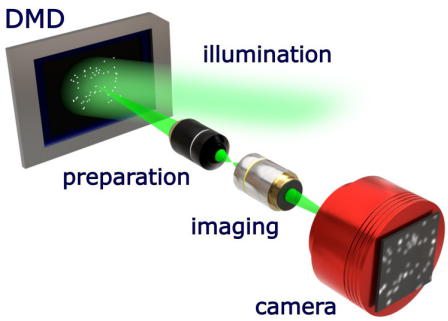


Figure 7.3. Schematic illustration of the optical setup used to collect experimental data pairs [Vašinka2025a]. A ground-truth mask is imposed on the digital micromirror device (DMD) by configuring its mirrors. An incoherent illumination light reflected by these mirrors impinges a high-resolution preparation system. The DMD-imposed mask is re-imaged into the front sample plane of the preparation system, creating point-like emitters with the intended spatial distribution. The imaging part of the setup, comprised of a low-resolution microscope objective, images the sample-plane emitters onto a camera. The resulting camera-captured intensity image and the DMD-imposed mask represent the experimental data pair.

Figure 7.4 (a) plots mean absolute error versus emitter concentration: dots are experimental measurements; solid lines are matched-parameter simulations. Despite Richardson–Lucy being supplied the measured PSF, DAMN achieves far lower errors across all concentrations. Figure 7.4 (b) shows a representative frame with ~200 emitters, its DMD mask, and reconstructions: DAMN closely reproduces the mask, resolving emitters spaced well below the Rayleigh limit, whereas Richardson–Lucy exhibits artifacts and misses sources.

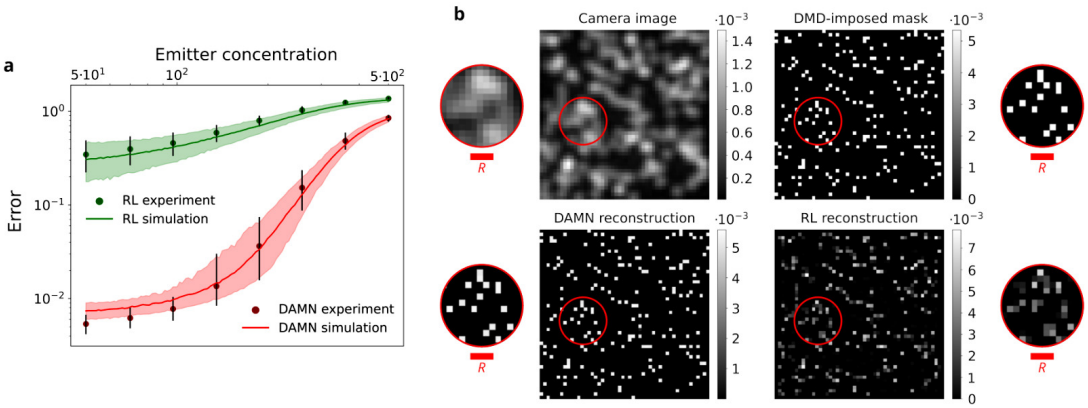


Figure 7.4. (a) The red and green dots represent the mean absolute error between the DMD-imposed masks and the super-resolved images reconstructed by each method. These errors were evaluated across various emitter concentrations. The accompanying continuous lines depict error values derived from simulated data using optical parameters estimated for our imaging system. (b) A typical camera image containing nearly 200 emitters, alongside its corresponding DMD-imposed mask and each method reconstructions. The circled areas contain a magnified region for easier visual comparison. It is evident that the DAMN model significantly outperforms the Richardson-Lucy (RL) algorithm even in regions where the mutual emitter distance is well below the Rayleigh resolution limit R = 3.9 px (inset scalebars).

Furthermore, we explored the benefits of incorporating upsampling layers into the DAMN convolutional architecture. The modified network reconstructs images with an eightfold increase in dimensions, leading to an even more prominent improvement in the resolution. We applied this model to an astronomy image of the galaxy NGC 300 by the European Southern Observatory [ESO2010], see Fig. 7.5 (a). Panel (b) depicts a magnified segment of the galaxy with marked regions of interest, while panel (c) presents their high-resolution reconstruction produced by the DAMN model. Despite the stars in all three regions being below the resolution limit of the imaging system, the DAMN reconstruction clearly reveals them. In region III, DAMN resolves two peaks separated by about 0.11 arcsec, suggesting a close pair that appears as one source in the original image. Confirming this hypothesis would require further analysis using additional data.

Additionally, we applied the same DAMN model, without retraining, to a publicly available dataset [SMLM-Challenge, Sage2015] acquired for the single-molecule localization microscopy challenge. Figure 7.6 depicts the performance evaluation using 500 tubulin images of 128 × 128 pixels with high molecule concentration, see panel (a) for an integrated image. Panel (b) shows a reference map from SOSplugin (least-squares localization with a Gaussian PSF) [Reuter2014], and panel (c) shows Deep-STORM [Nehme2018], a device-specific deep-learning method requiring setup details (camera, PSF model, SNR, density). DAMN needs none of this and directly produces a super-resolved tubulin image, see panel (d). Panel (e) visualizes the projection of a microtubule profile over the yellow segment, demonstrating the resolution improvement. Surprisingly, the device-agnostic DAMN without any calibration or information on microscope parameters outperforms the state-of-the-art device-specific Deep-STORM, which has been trained on calibration data and requires a detailed microscope characterization.

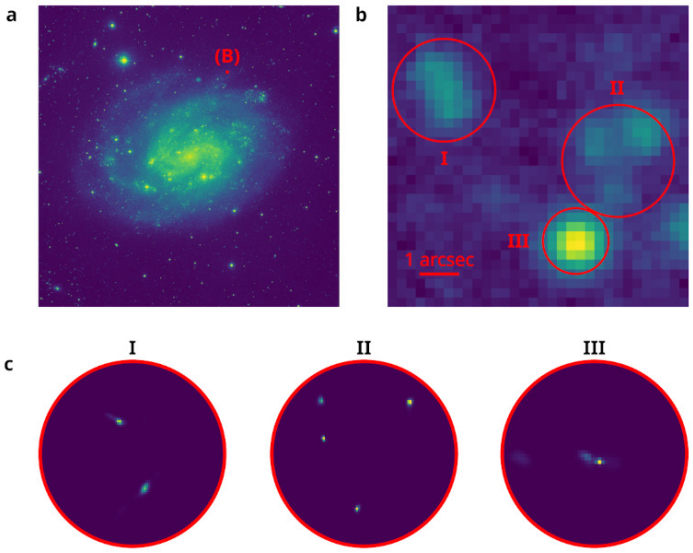


Figure 7.5. Demonstration of the DAMN model super-resolving capabilities on an astronomy image of a dense star cluster [Vašinka2025a]. (a) Intensity image of the spiral galaxy NGC 300 acquired by European Southern Observatory. (b) Magnified region of the galaxy, highlighting key areas of interest. (c) High-resolution reconstructions of these regions generated by the DAMN model, revealing details of individual stars. These reconstructions do not use any calibration data or prior information on the employed imaging system.

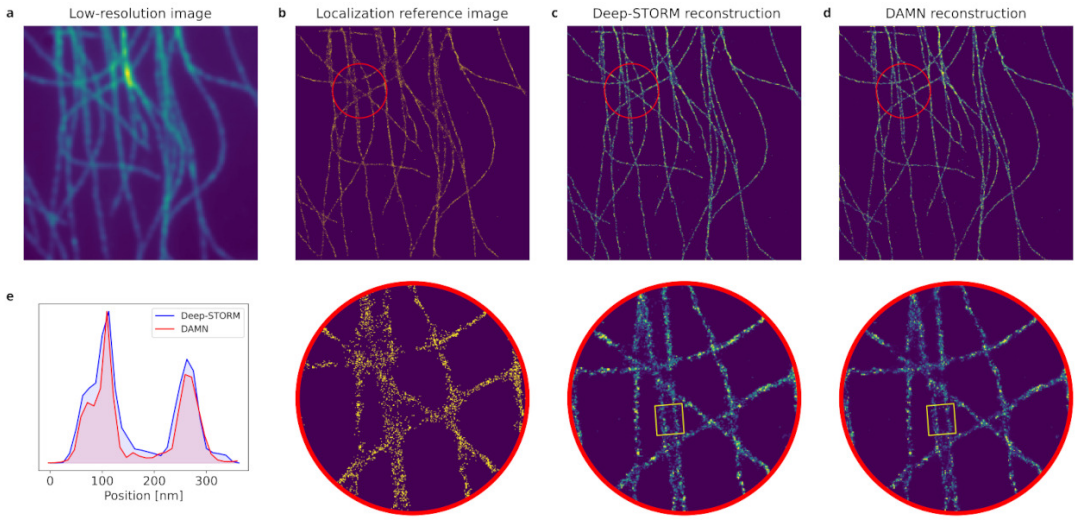


Figure 7.6. Demonstration of the DAMN model on a high-concentration tubulin dataset from the single-molecule localization microscopy challenge [Vašinka2025a]. (a) A low-resolution image integrated from the set of 500 measured images, containing numerous emitting molecules. (b) A reference binary map of localized molecule positions derived from the dataset. (c) A reconstruction generated by Deep-STORM, a device-dependent deep-learning approach calibrated using detailed information on the imaging setup. (d) A super-resolved image produced from the calibration-free, device-agnostic model without prior information. Panels (b)-(d) include inset images with a magnified region highlighting the resolution details. (e) Projection of the microtubule profile over the yellow rectangle segment.

Recently, we have applied the DAMN super-resolution framework for imaging of organic molecules and semiconductor quantum dots. Specifically, in [Dostálová2025a], we used terrylene molecules at various concentrations embedded in a polyvinyl alcohol layer. These samples were imaged using a high-resolution fluorescence microscope (100×/NA 1.4) and scientific CMOS camera. DAMN consistently outperformed Richardson–Lucy deconvolution, ThunderSTORM multi-emitter fitting, and DECODE deep-learning localization method [Speiser2021]. The state-of-the-art methods were supplied with the measured PSF and other necessary information, while DAMN operated in a completely calibration-free regime. Our model resolved molecules separated by 35 nm at 580 nm wavelength (from a single frame acquired within 150 ms), corresponding to seven-fold resolution improvement beyond the Rayleigh limit. These results demonstrate fast, calibration-free, single-frame super-resolution that is practical for dynamic or light-sensitive samples and broadly deployable on standard wide-field setups. Finally, in [Vašinka2025b], we validated the approach on low- and high-density In(Ga)As quantum dots and strain-induced dots in 2D monolayer WSe_2 at cryogenic temperatures using a low-resolution imaging system (NA 0.7) and electron-multiplying CCD camera, resolving overlapping emitters even under low signal-to-noise and inhomogeneous backgrounds. Chiefly, strain-induced dots were reconstructed along an etched hole edge with deviations of only 9 nm, significantly below the Rayleigh limit. By eliminating calibration and iterative acquisitions, this single-shot strategy enables rapid, robust super-resolution for nanomaterial characterization and quantum photonic device fabrication.

To conclude, we demonstrated a device-agnostic deep learning approach to super-resolution of emitters that eliminates the need for calibration data or system-

parameter estimation. Trained solely on numerically simulated data, the DAMN model reconstructs super-resolved images from a single camera frame without explicit knowledge of the optics, handles arbitrary image sizes, and remains robust to spatial nonuniformities and temporal drifts. On real-world data, from molecules to galaxies, calibration-free DAMN consistently outperforms calibration-dependent super-resolution methods.

The model complexity (and thus its universality) is currently limited by available computational resources, yet it already outperforms state-of-the-art super-resolution methods. With greater resources, further gains are expected from richer simulations and larger architectures. By taking this first step, our work lays a foundation for universal, optics-independent image reconstruction across domains. Most importantly, the results confirm the feasibility of the device-independent concept in physics. Classical imaging served as a proof of concept, and we are now ready to explore device-agnostic learning in the quantum domain.

8. Summary, outlook, and open questions

The presented thesis explored how to turn light into information with maximal efficiency and minimal resources, and provided demonstrations of several bleeding-edge photonic sensing and imaging solutions. We operate within a unified sensor scheme: a physical input (field, sample, emitter, or correlated source) is routed through an optical network (fixed or programmable, ordered or disordered, spatial or temporal) that produces an information-bearing pattern. Single-photon-sensitive detection acquires the raw statistics. A decoding stage (statistical estimators with known constraints or machine learning models trained on synthetic or experimental data) outputs the target quantity (photon statistics, polarization, entanglement, or emitter maps) without reconstructing unnecessary degrees of freedom. This encoder–decoder view clarifies design trade-offs, guides what to calibrate (and what to avoid calibrating), and explains why deep learning can substitute much of the optical complexity with digital processing. In sum, by treating optical networks as programmable encoders and computational networks as adaptable decoders, we move a step closer to the ultimate goal of optical sensing: **every photon counts, and every count becomes knowledge with as little calibration pain as possible.**

8.1. What we learned

In Chapter 2, we surveyed the state-of-the-art in network-based photonic sensing, single- and multi-photon detection, vector-field metrology, nanoscale emitters, and AI-assisted readout. We also summarized enabling hardware and methods developed in our lab.

In Chapter 3, we established a direct, single-source standard for measuring detector nonlinearity, avoiding detector-dependent models, revealing previously unreported supralinearity in widely used single-photon avalanche diodes and mapping sub- and supra-linear regimes in superconducting nanowire single-photon detectors, including faint-illumination supralinearity. These findings correct the common assumption of linear or sub-linear response and set requirements for precision photon counting, metrology, and quantum experiments: characterize the detector response empirically and do not rely on textbook models.

In Chapter 4, we built a large, reconfigurable, no-crosstalk spatially multiplexed photon-number-resolving detector and paired it with an expectation-maximization-entropy retrieval algorithm. The algorithm enforces positivity and mitigates maximum-likelihood sparsity bias, yielding order-of-magnitude improvements in fidelity and convergence over standard approaches. We validated the algorithm on a broad set of classical and nonclassical states and showed accurate estimates of the $g^{(2)}$ parameter even for states with high multiphoton content. High-resolution coincidence electronics rounds out a practical, metrology-grade photon-number-resolving detection pipeline.

In Chapter 5, we demonstrated compact, fast, and accurate polarization control and sensing. Twisted-nematic liquid-crystal stacks enabled preparation and measurement of arbitrary polarization states with ~99.9% fidelity; bidirectional neural calibration further improved the fidelity to ~99.99% while simplifying control. We then realized an all-fiber polarimetric micro-sensor that maps polarization to speckle, samples it with a fiber array, and decodes with deep learning at single-photon levels. We demonstrated in-situ, fast, and robust polarization readout.

In Chapter 6, we introduced direct, learning-based quantification of entanglement and all pairwise correlations from incomplete local measurements, bypassing full tomography. Two deep neural network variants (measurement-specific and measurement-independent) consistently outperformed maximum-likelihood reconstructions even under massive undersampling, both in measurement projectors and state copies. We demonstrated the real-world performance in experiments based on spontaneous parametric down-conversion and a resonantly driven quantum dot. The approach offers real-time estimates and a path toward scalable correlation sensing, while highlighting open issues in higher-qubit scaling and device independence.

In Chapter 7, we developed device-agnostic deep-learning super-resolution that reconstructs sub-diffraction structures from a single image without knowledge of any system parameters. On both astronomy and single-molecule data, the method surpassed the state-of-the-art baselines. The result is a concrete, working instance of measurement-device-independent inference in classical imaging and a template we aim to lift into quantum sensing.

Collectively, these results show that photonic sensing benefits from deliberately shifting complexity from optical parts of a sensor and its hardware calibration to computational decoders, preferably universally trained on synthetic data. Done correctly, the shift increases robustness, reduces calibration debt, and pushes accuracy toward physical limits while retaining speed.

8.2. Outlook: toward programmable, multiparameter, and self-calibrating sensors

A programmable, multiparameter encoder. The next step is to design optical networks that jointly encode several observables, such as photon statistics, spatial distribution, polarization, and correlations, into a single detection signature. Practically, this means combining stable (but possibly random) interferometric networks and decoding with a multi-task model that outputs several target parameters from the same raw clicks. The hardware and methods presented in this work are natural building blocks for this integration.

Scaling channels and dynamic range. Accurate statistics for mesoscopic states (hundreds to thousands of photons) require many effective channels and low-loss routing. We plan to combine spatial multiplexing with active temporal networks and detector-property modulation to expand the dynamic range while keeping crosstalk

negligible. A similar goal applies to other degrees of freedom of light, such as polarization and its correlations, where the number of measurement projections can be further increased. This will potentially improve the accuracy of the measurement to the ultimate quantum limit, support emitter counting at ultra-high densities, and allow large-scale, multi-parameter sensing.

Hybrid statistical-neural decoders. Pure machine learning can overfit or ignore constraints; pure statistical estimators can be slow or brittle. A practical path is hybrid: embed the physics (positivity, normalization, noise statistics) as differentiable layers or penalties, and let a compact network learn residual structure and denoise. Our expectation-maximization-entropy method already formalizes a hard-coded, principled prior; the network can supply learned regularization, uncertainty estimates, and fast amortized inference for repeated measurements.

From calibration-lite to calibration-free. Our device-agnostic modeling network demonstrates that training on vast physics-respecting simulations can remove the need for prior calibration of the measurement device while surviving real-world nonidealities. The same idea can regularize photon-counting inversions and correlation estimators: train decoders on synthetic distributions pushed through measured or simulated detector transfer matrices, and deploy without per-session detector tomography.

Reliability, uncertainty, and interpretability. For broader deployment in metrology and biomedicine, decoders must quantify confidence and expose salient features. Likelihood-based training and Bayesian deep learning can provide variances (concentrations) as latent outputs. Post-hoc explainability can identify which detection features drive outputs and flag distribution shifts. Such diagnostics will be essential for regulatory and scientific acceptance.

Real-time and adaptive sensing. The learning-based sensing suggests a broader paradigm: inferring the target quantity (such as entanglement measures and correlation graphs) directly from incomplete, noisy local data, fast enough for closed-loop experiments. Beyond producing a point estimate, the model also outputs an uncertainty and proposes the next measurement projection that maximizes expected information gain for the target. This active policy iteratively selects measurement projections under experimental constraints, updates the estimate after each acquisition, and stops once a preset confidence or error threshold is reached, enabling efficient, data-frugal sensing.

8.3. Open questions and challenges

Information flow in encoder–decoder sensors. We lack a general theory that tracks how task-relevant information (e.g., Fisher information) propagates through an optical network and a neural decoder under realistic noise, loss, and model mismatch [Hüpfl2024, Weimar2025]. Developing tight, experimentally testable bounds, covering randomized encoders and learned decoders, would guide both optical design (which mixing patterns and bandwidths to prefer) and training (which priors and augmentations to use to preserve information).

Self-calibrating hardware via in-situ learning. Can we close the loop and train the photonic encoder itself by backpropagating a task loss into programmable optics (e.g., phase shifters, fast polarization elements), while the decoder adapts concurrently? Proof-of-principle optical neural networks exist, but integrating them with well-characterized detectors and stable, low-latency control remains to be done. A successful implementation would reduce explicit calibration to "bring-up only" and make sensors drift-tolerant by design.

Physics-aware and explainable neural decoding. How best to embed physical constraints (positivity, conservation laws, Poisson counting, monotonicities) into trainable architectures without sacrificing speed or generalization? Furthermore, deployable decoders must return credible intervals alongside point estimates and support out-of-distribution detection. They should also provide human-legible attributions so experimentalists can debug failures and refine hardware. Tools from interpretable and explainable learning offer a starting point, but domain-specific measures that respect optical symmetries, quantum-state constraints, and counting statistics are still missing.

Scaling and device independence in the quantum regime. Our entanglement estimators outperform tomography under undersampling for calibrated measurement projections and a few qubits, but full device independence and genuine multipartite quantification at scale remain open. Symmetry-aware, graph-based models and convex-roof surrogates are plausible directions, yet require careful validation on high-quality experimental datasets. Translating super-resolution DAMN-style device-agnosticism into quantum sensing is another challenge. The need for massive computational resources for multi-qubit scenarios and device-agnostic models represents a serious roadblock.

References

[Aaronson2013] S. Aaronson and A. Arkhipov, The computational complexity of linear optics, Theory of Computing. 9, 143 (2013).

[Aaronson2018] S. Aaronson, Shadow tomography of quantum states, in STOC 2018: Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, 325 (Association for Computing Machinery, 2018).

[Abbaszadeh2021] H. Abbaszadeh, et al., Liquid-crystal-based topological photonics, Proc. Natl. Acad. Sci. 118, e2020525118 (2021).

[Acampora2025] G. Acampora, et al., Quantum computing and artificial intelligence: status and perspectives, arXiv:2505.23860 (2025).

[Acerbi2019] F. Acerbi and S. Gundacker, Understanding and simulating SiPMs, Nucl. Instr. Meth. Phys. Res. A. 926, 16 (2019).

[Acernese2019] F. Acernese, et al. (Virgo Collaboration), Increasing the astrophysical reach of the advanced Virgo detector via the application of squeezed vacuum states of light, Phys. Rev. Lett. 123, 231108 (2019).

[Achilles2003] D. Achilles, et al., Fiber-assisted detection with photon number resolution, Opt. Lett. 28, 2387 (2003).

[Ackley1985] D. H. Ackley, et al., A learning algorithm for Boltzmann machines, Cogn. Sci. 9, 147 (1985).

[Adamson2007] R. B. A. Adamson, et al., Preparation of pure and mixed polarization qubits, Phys. Rev. A 75, 012104 (2007).

[Adamson2010] R. B. A. Adamson and A. M. Steinberg, Improving quantum state estimation with mutually unbiased bases, Phys. Rev. Lett. 105, 030406 (2010).

[Akhlaghi2009] M. K. Akhlaghi and A. H. Majedi, Semiempirical modeling of dark count rate and quantum efficiency of superconducting nanowire single-photon detectors, IEEE Trans. Appl. Supercond. 19, 361–366 (2009).

[Akhlaghi2011] M. K. Akhlaghi, et al., Nonlinearity in single photon detection: modeling and quantum tomography, Opt. Express 19, 21305 (2011).

[Allevi2010a] A. Allevi, et al., Photon-number correlations by photon-number resolving detectors, Opt. Lett. 35, 1707 (2010).

[Allevi2010b] A. Allevi, et al., Reliable source of conditional states from single-mode pulsed thermal fields by multiple-photon subtraction, Phys. Rev. A 82, 013816 (2010).

[Altepeter2011] J. B. Altepeter, et al., Entangled photon polarimetry, Opt. Express 19, 26011 (2011).

[Amawi2024] M. T. Amawi, et al., Three-dimensional magnetic resonance tomography with sub-10 nanometer resolution, npj Quantum Inf. 10, 16 (2024).

[Amico2008] L. Amico, et al., Entanglement in many-body systems, Rev. Mod. Phys. 80, 517 (2008).

[Amini2020] A. Amini, et al., Deep evidential regression, Adv. Neural Inf. Process. Syst. 33, 14927 (2020).

[Arnold2005] M. A. Arnold and G. W. Small, Noninvasive glucose sensing, Anal. Chem. 77, 5429 (2005).

[Arute2019] F. Arute, et al., Quantum supremacy using a programmable superconducting processor, Nature 574, 505 (2019).

[Asavanant2019] W. Asavanant, et al., Generation of time-domain-multiplexed two-dimensional cluster state Science 366, 373 (2019).

[Aslam2023] N. Aslam, et al., Quantum sensors for biomedical applications, Nat. Rev. Phys. 5, 157 (2023).

[Aspuru-Guzik2012] A. Aspuru-Guzik and P. Walther, Photonic quantum simulators, Nat. Phys. 8, 285 (2012).

[Ates2012] S. Ates, et al., Bright single-photon emission from a quantum dot in a circular bragg grating microcavity, IEEE J. Sel. Top. Quantum Electron. 18, 1711 (2012).

[Audenaert2006] K. M. R. Audenaert, M. B. Plenio, When are correlations quantum? —Verification and quantification of entanglement by simple measurements, New J. Phys. 8, 266 (2006).

[Avenhaus2009] M. Avenhaus, et al., Fiber-assisted single-photon spectrograph, Opt. Lett. 34, 2873 (2009).

[Ballard2021] Z. Ballard et al., Machine learning and computation-enabled intelligent sensor design, Nat. Mach. Intell. 3, 556 (2021).

[Balzarotti2017] F. Balzarotti, et al., Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes, Science 355, 606 (2017).

[Banaszek1998] K. Banaszek, Reconstruction of photon distribution with positivity constraints, Acta Phys. Slov. 48, 185 (1998); Maximum-likelihood estimation of photon-number distribution from homodyne statistics, Phys. Rev. A 57, 5013 (1998).

[Banaszek2003] K. Banaszek and I. A. Walmsley, Photon counting with a loop detector, Opt. Lett. 28, 52 (2003).

[Banner2024] P. R. Banner, et al., Number-state reconstruction with a single single-photon avalanche detector, Opt. Quantum 2, 110 (2024).

[Barbastathis2019] G. Barbastathis, et al., On the use of deep learning for computational imaging, Optica 6, 921 (2019).

[Barbieri2003] M. Barbieri, et al., Detection of entanglement with polarized photons: Experimental realization of an entanglement witness, Phys. Rev. Lett. 91, 227901 (2003).

[Barreiro2005] J. T. Barreiro, et al., Generation of hyperentangled photon pairs, Phys. Rev. Lett. 95, 260501 (2005).

[Barron1993] A. R. Barron, Universal approximation bounds for superpositions of a sigmoidal function, IEEE Trans. Inf. Theory 39, 930–945 (1993).

[Barry2020] J. F. Barry, et al., Sensitivity optimization for NV-diamond magnetometry, Rev. Mod. Phys. 92, 015004 (2020).

[BassoBasset2021] F. Basso Basset, et al., Quantum key distribution with entangled photons generated on demand by a quantum dot, Sci. Adv. 7, eabe6379 (2021).

[Baumann2010] K. Baumann, et al., Dicke quantum phase transition with a superfluid gas in an optical cavity, Nature 464, 1301 (2010).

[Běhal2022] J. Běhal, M. Valentino, L. Miccio, V. Bianco, S. Itri, R. Mossotti, G. Dalla Fontana, E. Stella, and P. Ferraro, Polarimetric detection in environmental optics, ACS Photonics 9, 694 (2022).

[Běhal2025] J. Běhal and M. Ježek, Incoherent lateral shearing digital holographic microscopy, Optica 12, 620 (2025).

[Belkin2019] M. Belkin, et al., Reconciling modern machine-learning practice and the classical bias–variance trade-off, Proc. Natl. Acad. Sci. U.S.A. 116, 15849 (2019).

[Bell1964] J. S. Bell, On the Einstein Podolsky Rosen paradox, Phys. Phys. Fiz. 1, 195 (1964).

[Belthangady2019] C. Belthangady and L. A. Royer, Applications, promises, and pitfalls of deep learning for fluorescence image reconstruction, Nat. Methods 16, 1215 (2019).

[Benedetti2013] C. Benedetti, et al., Experimental estimation of quantum discord for a polarization qubit and the use of fidelity to assess quantum correlations, Phys. Rev. A 87, 052136 (2013).

[Bertero2021] M. Bertero, et al., Introduction to Inverse Problems in Imaging (2nd ed.), CRC Press (2021).

[Bertolotti2012] J. Bertolotti, et al., Non-invasive imaging through opaque scattering layers, Nature 491, 232 (2012).

[Bertolotti2022] J. Bertolotti and O. Katz, Imaging in complex media, Nat. Phys. 18, 1008 (2022).

[Betzig2006] E. Betzig, et al., Imaging intracellular fluorescent proteins at nanometer resolution, Science 313, 1642 (2006).

[Bianco2023] V. Bianco, et al., Label-free intracellular multi-specificity in yeast cells by phase-contrast tomographic flow cytometry, Small Methods 7, 2300447 (2023).

[Bielak2021] M. Bielak, et al., Accurate polarization preparation and measurement using twisted nematic liquid crystals, Opt. Express. 29, 33037 (2021).

[Bielak2025] M. Bielak, et al., All-fiber microsensor of polarization at single-photon level aided by deep learning, Laser Photonics Rev. e01775 (2025).

[Bodén2024] A. Bodén, et al., Super-sectioning with multi-sheet RESOLFT microscopy, Nat. Methods 21, 882 (2024).

[Bogaerts2020] W. Bogaerts, et al., Programmable photonic circuits, Nature 586, 207 (2020).

[Bohnet2016] J. G. Bohnet, et al., Quantum spin dynamics and entanglement generation with hundreds of trapped ions, Science 352, 1297 (2016).

[Bommasani2021] R. Bommasani, et al., On the opportunities and risks of foundation models, arXiv:2108.07258 (2021).

[Booth2015] M. Booth, et al., Aberrations and adaptive optics in super-resolution microscopy, Microscopy 64, 251 (2015).

[Borhani2018] N. Borhani, et al., Learning to see through multimode fibers, Optica 5, 960 (2018).

[Bottou2010] L. Bottou, Large-scale machine learning with stochastic gradient descent, In Y. Lechevallier and G. Saporta, eds., Proceedings of COMPSTAT'2010, pp. 177–186 (Physica-Verlag HD, 2010).

[Bouwmeester1997] D. Bouwmeester, et al., Experimental quantum teleportation, Nature 390, 575 (1997).

[Brasselet2019] S. Brasselet, Polarization-resolved microscopy in the life sciences, Opt. Photonics News 30, 34 (2019).

[Breitenbach1997] G. Breitenbach, et al., Measurement of the quantum states of squeezed light, Nature 387, 471 (1997).

[Brida2010a] G. Brida, et al., Experimental estimation of entanglement at the quantum limit, Phys. Rev. Lett. 104, 100501 (2010).

[Brida2010b] G. Brida, et al., Experimental realization of sub-shot-noise quantum imaging, Nat. Photonics 4, 227 (2010).

[Brida2011] G. Brida, et al., Optimal estimation of entanglement in optical qubit systems, Phys. Rev. A 83, 052301 (2011).

[Bromberg2009] Y. Bromberg, et al, Quantum and classical correlations in waveguide lattices, Phys. Rev. Lett. 102, 253904 (2009).

[Bruschini2019] C. Bruschini, et al., Single-photon avalanche diode imagers in biophotonics: review and outlook, Light Sci. Appl. 8, 87 (2019).

[Brydges2019] T. Brydges, et al., Probing Rényi entanglement entropy via randomized measurements, Science 364, 260 (2019).

[Budker2007] D. Budker and M. Romalis, Optical magnetometry, Nat. Phys. 3, 227 (2007).

[Bueno2000] J. M. Bueno, Polarimetry using liquid-crystal variable retarders, J. Opt. A 2, 216 (2000).

[Buller2010] G. S. Buller and R. J. Collins, Single-photon generation and detection, Meas. Sci. Technol. 21, 012002 (2010).

[Bures1969] D. Bures, An extension of Kakutani's theorem on infinite product measures to the tensor product of semifinite w*-algebras, Trans. Am. Math. Soc. 135, 199 (1969).

[Cahall2017] C. Cahall, et al., Multi-photon detection using a conventional superconducting nanowire single-photon detector, Optica 4, 1534 (2017).

[Calkins2013] B. Calkins, et al., High quantum-efficiency photon-number-resolving detector for photonic on-chip information processing, Opt. Express 21, 22657 (2013).

[Campisi2007] P. Campisi and K. Egiazarian, Blind Image Deconvolution: Theory and Applications, CRC Press (2007).

[Cao2022] H. Cao, et al., Wavefront shaping in complex media, Nat. Phys. 18, 994 (2022).

[Cao2023] H. Cao, et al., Controlling light propagation in multimode fibers for imaging, spectroscopy, and beyond, Adv. Opt. Photon. 15, 524 (2023).

[Caramazza2019] P. Caramazza, et al., Transmission of natural scene images through a multimode fibre, Nat. Commun. 10, 2029 (2019).

[Caravaca-Aguirre2013] A. M. Caravaca-Aguirre et al., Real-time resilient focusing through a bending multimode fiber, Opt. Express 21, 12881 (2013).

[Cardano2016] F. Cardano, et al., Statistical moments of quantum-walk dynamics reveal topological quantum transitions, Nat. Commun. 7, 11439 (2016).

[Carleo2017] G. Carleo and M. Troyer, Solving the quantum many-body problem with artificial neural networks, Science 355, 602 (2017).

[Carleo2018] G. Carleo, Y. Nomura, M. Imada, Constructing exact representations of quantum many-body systems with deep neural networks, Nat. Commun. 9, 5322 (2018).

[Carleo2019] G. Carleo, et al., Machine learning and the physical sciences, Rev. Mod. Phys. 91, 045002 (2019).

[Carmeli2016] C. Carmeli, et al., Verifying the quantumness of bipartite correlations, Phys. Rev. Lett. 116, 230403 (2016).

[Carolan2015] J. Carolan, et al., Universal linear optics, Science 349, 711 (2015).

[Carpenter1991] G. A. Carpenter, et al., Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system, Neural Netw. 4, 759 (1991).

[Carpenter2016] J. Carpenter, et al., Complete spatiotemporal characterization and optical transfer matrix inversion of a 420 mode fiber, Opt. Lett. 41, 5580 (2016).

[Cava2018] A. Cava, et al., The nature of giant clumps in distant galaxies probed by the anatomy of the Cosmic Snake, Nat. Astron. 2, 76 (2018).

[Chen2016] Z.-Y. Chen, et al., High-capacity transmission using polarization freedom, Light: Sci. Appl. 6, e16207 (2016).

[Chen2018] H. Chen et al., High-performance photogating photodetectors based on perovskite–graphene hybrid materials, Nat. Commun. 9, 1920 (2018).

[Chen2020] T. Chen, et al., A simple framework for contrastive learning of visual representations, Proc. ICML 119, 1597 (2020).

[Chen2021] X. Chen and K. He, Exploring simple siamese representation learning, Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 15750 (2021).

[Chen2022] B. Chen, et al., Resolution doubling in light-sheet microscopy via oblique-plane structured illumination, Nat. Methods 19, 1419 (2022).

[Cheng2023] R. Cheng, et al., A 100-pixel photon-number-resolving detector unveiling photon statistics, Nat. Photonics 17, 112 (2023).

[Childs2004] A. M. Childs and J. Goldstone, Spatial search by quantum walk, Phys. Rev. A 70, 022314 (2004).

[Childs2013] A. M. Childs, et al., Universal computation by multiparticle quantum walk, Science 339, 791 (2013).

[Chiles2022] J. Chiles, et al., New constraints on dark photon dark matter with superconducting nanowire detectors in an optical haloscope. Phys. Rev. Lett. 128, 231802 (2022).

[Čižmár2012] T. Čižmár and K. Dholakia, Exploiting multimode waveguides for pure fibre-based imaging, Nat. Commun. 3, 1027 (2012).

[Cohen2016] T. S. Cohen and M. Welling, Group equivariant convolutional networks, Proceedings of ICML (2016).

[Coppola2024] S. Coppola, et al., Compact systems for polarization-based inspection, ACS Appl. Mater. Interfaces 16, 19453 (2024).

[Cortez2009] P. Cortez, et al., Modeling wine preferences by data mining from physicochemical properties, Decis. Support Syst. 47, 547 (2009).

[Cotter2017] J. P. Cotter, et al., In search of multipath interference using large molecules," Sci. Adv. 3, e1602478 (2017).

[Cova1996] S. Cova, et al., Avalanche photodiodes and quenching circuits for single-photon detection, Appl. Opt. 35, 1956 (1996).

[Cova2004] S. Cova, et al., Evolution and prospects for single-photon avalanche diodes and quenching circuits, J. Mod. Opt. 51, 1267 (2004).

[Coy1996] J. A. Coy, et al., Characterization of a liquid crystal television as a programmable spatial light modulator, Opt. Eng. 35, 15 (1996).

[Cramer2010] M. Cramer, et al., Efficient quantum state tomography, Nat. Commun. 1, 149 (2010).

[Cranmer2020] M. Cranmer, et al., Discovering symbolic models from deep learning with inductive biases, 34th Conference on Neural Information Processing Systems (NeurIPS), Vancouver (2020).

[Crespi2011] A. Crespi, et al., Integrated photonic quantum gates for polarization qubits, Nat. Commun. 2, 566 (2011).

[Cybenko1989] G. Cybenko, Approximation by superpositions of a sigmoidal function, Math. Control Signals Syst. 2, 303–314 (1989).

[Dahmardeh2023] M. Dahmardeh, et al., Self-supervised machine learning pushes the sensitivity limit in label-free detection of single proteins below 10 kDa, Nat. Methods 20, 442 (2023).

[Danaci2021] O. Danaci, et al., Machine learning pipeline for quantum state estimation with incomplete measurements, Mach. Learn. Sci. Technol. 2, 035014 (2021).

[Das2011] R. S. Das and Y. K. Agrawal, Raman spectroscopy: Recent advancements, techniques and applications, Vib. Spectrosc. 57, 163 (2011).

[Datta2020] R. Datta, et al., Fluorescence lifetime imaging microscopy: fundamentals and advances in instrumentation, analysis, and applications, J. Biomed. Opt. 25, 071203 (2020).

[Dawid2025] A. Dawid, et al., Machine Learning in Quantum Sciences (Cambridge University Press, 2025).

[Defienne2014] H. Defienne, et al., Nonclassical light manipulation in a multiple-scattering medium, Opt. Lett. 39, 6090 (2014).

[Defienne2016] H. Defienne, et al., Two-photon quantum walk in a multimode fiber, Sci. Adv. 2, e1501054 (2016).

[Defienne2021] H. Defienne, et al., Polarization entanglement-enabled quantum holography, Nat. Phys. 17, 591 (2021).

[Degen2017] C. L. Degen, et al., Quantum sensing, Rev. Mod. Phys. 89, 035002 (2017).

[DeMartino2003] A. De Martino, et al., Optimized Mueller polarimeter with liquid crystals, Opt. Lett. 28, 616 (2003).

[Dempster1977] A. P. Dempster, et al., Maximum likelihood from incomplete data via the EM algorithm, J. Royal Stat. Soc. B 39, 1 (1977).

[Deng2017] D.-L. Deng, et al., Quantum entanglement in neural network states, Phys. Rev. X 7, 021021 (2017).

[Deng2021] Y. Deng, et al., In-situ material inspection using polarization imaging, Light: Sci. Appl. 10, 76 (2021).

[Deng2023] Y.-H. Deng, et al., Gaussian boson sampling with pseudo-photon-number-resolving detectors and quantum computational advantage, Phys. Rev. Lett. 131, 150601 (2023).

[Digman2008] M. Digman, et al., Mapping the number of molecules and brightness in the laser scanning microscope, Biophys. J. 94, 2320-2332 (2008).

[Divochiy2008] A. Divochiy, et al., Superconducting nanowire photon-number-resolving detector at telecommunication wavelengths, Nat. Photonics 2, 302 (2008).

[Doradla2013] P. Doradla, et al., Polarization-resolved detection of gastrointestinal cancer, J. Biomed. Opt. 18, 090504 (2013).

[Dostálová2025a] A. Dostálová, et al., Calibration-free single-frame super-resolution fluorescence microscopy, preprint arXiv:2505.13293, bioRxiv:2025.05.20.655080 (2025).

[Dostálová2025b] A. Dostálová, et al., Sensing and imaging with single organic molecules with photon-number-resolved detection and deep learning, Bristol Quantum Information Technologies (BQIT:25), conference poster (2025).

[Dozat2016] T. Dozat, Incorporating Nesterov Momentum into Adam, International Conference on Learning Representations (ICLR) Workshop (2016).

[Du2025] Y. Du, et al., Artificial intelligence for representing and characterizing quantum systems, arXiv:2509.04923 (2025).

[Eaton2023] M. Eaton, et al., Resolution of 100 photons and quantum generation of unbiased random numbers, Nat. Photonics 17, 106 (2023).

[Edrei2016] E. Edrei and G. Scarcelli, Optical imaging through dynamic turbid media using the Fourier-domain shower-curtain effect, Optica 3, 71 (2016).

[Einstein1935] A. Einstein, et al., Can quantum-mechanical description of physical reality be considered complete?, Phys. Rev. 47, 777 (1935).

[Eisaman2011] M. D. Eisaman, et al., Single-photon sources and detectors, Rev. Sci. Instrum. 82, 071101 (2011).

[Eisert2020] J. Eisert, et al., Quantum certification and benchmarking, Nat. Rev. Phys. 2, 382 (2020).

[Ekert2002] A. K. Ekert, et al., Direct estimations of linear and nonlinear functionals of a quantum state, Phys. Rev. Lett. 88, 217901 (2002).

[Elben2018] A. Elben, et al., Rényi entropies from random quenches in atomic Hubbard and spin models, Phys. Rev. Lett. 120, 050406 (2018).

[Elben2023] A. Elben, et al., The randomized measurement toolbox, Nat. Rev. Phys. 5, 9 (2023).

[Elmalem2025] S. Elmalem, et al., Massively multiplexed wide-field photon correlation sensing, Optica 12, 451 (2025).

[Elson2007] J. Elson, et al., Asirra: A CAPTCHA that exploits interest-aligned manual image categorization, Proc. 14th ACM Conf. Comput. Commun. Secur. (CCS), 366 (2007).

[Endo2025] M. Endo, et al., Optically sampled superconducting-nanostrip photon-number resolving detector for non-classical quantum state generation, Opt. Express 33, 32545 (2025).

[Erhard2020] M. Erhard, et al., Advances in high-dimensional quantum entanglement, Nat. Rev. Phys. 2, 365 (2020).

[ESO2010] European Southern Observatory, Wide Field Imager view of the southern spiral NGC 300. https://www.eso.org/public/images/eso1037a/ (2010).

[Evered2023] S. J. Evered, et al., High-fidelity parallel entangling gates on a neutral-atom quantum computer, Nature 622, 268 (2023).

[Facchin2020] M. Facchin, et al., Speckle-based determination of the polarisation state of single and multiple laser beams, OSA Contin. 3, 1302 (2020).

[Fadrný2024] J. Fadrný, et al., Experimental preparation of multiphoton-added coherent states of light, npj Quantum Inf. 10, 89 (2024).

[Feng2021] Y. Feng and Y. Tu, The inverse variance–flatness relation in stochastic gradient descent is critical for finding flat minima, Proc. Natl. Acad. Sci. U.S.A. 118(9) (2021).

[Filip2002] R. Filip, Overlap and entanglement-witness measurements, Phys. Rev. A 65, 062320 (2002).

[Fishman2023] R. E. K. Fishman, et al., Photon-emission-correlation spectroscopy as an analytical tool for solid-state quantum defects, PRX Quantum 4, 010202 (2023).

[Fitch2003] M. J. Fitch, et al., Photon-number resolution using time-multiplexed single-photon detectors, Phys. Rev. A 68, 043814 (2003).

[Fiurášek2002] J. Fiurášek, Structural physical approximations of unphysical maps and generalized quantum measurements, Phys. Rev. A 66, 052315 (2002).

[Fiurášek2004] J. Fiurášek and N. J. Cerf, How to measure squeezing and entanglement of Gaussian states without homodyning, Phys. Rev. Lett. 93, 063601 (2004).

[Fiurášek2025] J. Fiurášek, Fundamental limits on determination of photon number statistics from measurements with multiplexed on/off detectors, Phys. Rev. Research 7, 033147 (2025).

[Flamini2019] F. Flamini, et al., Photonic quantum information processing: a review, Rep. Prog. Phys. 82, 016001 (2019).

[Freire2023] P. Freire, et al., Artificial neural networks for photonic applications – from algorithms to implementation: tutorial, Adv. Opt. Photon. 15, 739-834 (2023).

[Freund1990] I. Freund, Looking through walls and around corners, Physica A Stat. Mech. Appl. 168, 49 (1990).

[Furusawa1998] A. Furusawa, et al., Unconditional quantum teleportation, Science 282, 706 (1998).

[Fyrillas2024] A. Fyrillas, et al., Scalable machine learning-assisted clear-box characterization for optimally controlled photonic circuits, Optica 11, 427 (2024).

[Gao2018] J. Gao, et al., Experimental machine learning of quantum states, Phys. Rev. Lett. 120, 240501 (2018).

[Gao2025] D. Gao, et al., Establishing a new benchmark in quantum computational advantage with 105-qubit Zuchongzhi-3.0 processor, Phys. Rev. Lett. 134, 090601 (2025).

[Gebhart2023] V. Gebhart, et al., Learning quantum systems, Nat. Rev. Phys. 5, 141 (2023).

[Gilmer2017] J. Gilmer, et al., Neural message passing for quantum chemistry, ICML'17: Proceedings of the 34th International Conference on Machine Learning – volume 70, 1263 (2017).

[Giovannetti2004] V. Giovannetti, et al., Quantum-Enhanced Measurements: Beating the Standard Quantum Limit, Science 306, 1330 (2004).

[Giovannetti2006] V. Giovannetti, et al., Quantum Metrology, Phys. Rev. Lett. 96, 010401 (2006).

[Giovannetti2011] V. Giovannetti, et al., Advances in quantum metrology, Nat. Photonics 5, 222 (2011).

[Giustina2015] M. Giustina, et al., Significant-loophole-free test of Bell's theorem with entangled photons, Phys. Rev. Lett. 115, 250401 (2015).

[Goel2024] S. Goel, et al., Inverse design of high-dimensional quantum optical circuits in a complex medium, Nat. Physics 20, 232 (2024).

[Gol'tsman2001] G. N. Gol'tsman, et al., Picosecond superconducting single-photon optical detector, Appl. Phys. Lett. 79, 705 (2001).

[Google2023] Google Quantum AI, Suppressing quantum errors by scaling a surface code logical qubit, Nature 614, 676 (2023).

[Gorman2014] K. B. Gorman, et al., Ecological sexual dimorphism and environmental variability within a community of Antarctic penguins (Genus Pygoscelis), PLoS One 9, e90081 (2014).

[Grill2020] J.-B. Grill, et al., Bootstrap your own latent: A new approach to self-supervised learning, Adv. Neural Inf. Process. Syst. 33, 21271 (2020).

[Gross2010] D. Gross, et al., Quantum state tomography via compressed sensing. Phys. Rev. Lett. 105, 150401 (2010).

[Grygar2022] J. Grygar, et al., Quantum non-Gaussianity certification of photon-number-resolving detectors, Opt. Express 30, 33097 (2022).

[Gu2019] Y. Gu et al., A graphene-based broadband optical detector with high gain and low noise, Nat. Commun. 10, 265 (2019).

[Guccione2020] G. Guccione, et al., Connecting heterogeneous quantum networks by hybrid entanglement swapping, Sci. Adv. 6, eaba4508 (2020).

[Guhne2002] O. Gühne, et al., Detection of entanglement with few local measurements, Phys. Rev. A 66, 062305 (2002).

[Gundacker2020] S. Gundacker and A. Heering, The silicon photomultiplier: fundamentals and applications of a modern solid-state photon detector, Phys. Med. Biol. 65, 17TR01 (2020).

[Guo2020] Y. Guo, et al., High-responsivity photodetectors based on few-layer MoS2, ACS Nano 14, 264–271 (2020).

[Guo2024] S.-A. Guo, et al., A site-resolved two-dimensional quantum simulator with hundreds of trapped ions, Nature 630, 613 (2024).

[Gupta2020] R. K. Gupta, et al., Deep learning enabled laser speckle wavemeter with a high dynamic range, Laser & Photonics Rev. 14, 2000120 (2020).

[Haderka2009] O. Haderka, et al., Fast time-domain balanced homodyne detection of light, Appl. Opt. 48, 2884 (2009).

[Hadfield2009] R. H. Hadfield, Single-photon detectors for optical quantum information applications, Nat. Photonics 3, 696 (2009).

[Hadfield2016] R. H. Hadfield and G. Johansson, Eds., Superconducting Devices in Quantum Optics, edited by (Springer, 2016).

[Hadfield2023] R. H. Hadfield, et al., Single-photon detection for long-range imaging and sensing, Optica 10, 1124 (2023).

[Häffner2005] H. Häffner, et al., Scalable multiparticle entanglement of trapped ions, Nature 438, 643 (2005).

[Hansen2001] H. Hansen, et al., Ultrasensitive pulsed, balanced homodyne detector: application to time-domain quantum measurements, Opt. Letters 26, 1714 (2001).

[Harder2016] G. Harder, et al., Single-mode parametric-down-conversion states with 50 photons as a source for mesoscopic quantum optics, Phys. Rev. Lett. 116, 143601 (2016).

[Harney2020] C. Harney, et al., Entanglement classification via neural network quantum states, New J. Phys. 22, 045001 (2020).

[Hartmann2019] M. J. Hartmann, G. Carleo, Neural-network approach to dissipative quantum many-body dynamics, Phys. Rev. Lett. 122, 250502 (2019).

[He2017] Y. He, et al., Time-bin encoded boson sampling with a single-photon device, Phys. Rev. Lett. 118, 190501 (2017).

[Hekrdla2025] M. Hekrdla, et al., Optimized molecule detection in localization microscopy with selected false positive probability, Nat. Commun. 16, 601 (2025).

[Hensen2015] B. Hensen, et al., Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres, Nature 526, 682 (2015).

[Hinton2002] G. E. Hinton, Training products of experts by minimizing contrastive divergence, Neural Comput. 14, 1771 (2002).

[Hinton2006] G. E. Hinton and R. R. Salakhutdinov, Reducing the Dimensionality of Data with Neural Networks, Science 313, 504 (2006).

[Hloušek2019] J. Hloušek, et al., Accurate detection of arbitrary photon statistics, Phys. Rev. Lett. 123, 153604 (2019).

[Hloušek2021] J. Hloušek, et al., Direct experimental certification of quantum non-Gaussian character and Wigner function negativity of single-photon detectors, Phys. Rev. Lett. 126, 043601 (2021).

[Hloušek2022] J. Hloušek, Photon-number-resolving detectors and their applications in quantum technologies, Ph.D. thesis, Palacký University Olomouc (2022).

[Hloušek2023] J. Hloušek, et al., Experimental observation of anomalous supralinear response of single-photon detectors, Appl. Phys. Rev. 10, 011412 (2023).

[Hloušek2024a] J. Hloušek, et al., High-resolution coincidence counting system for large-scale photonics applications, Phys. Rev. Applied 21, 024023 (2024).

[Hloušek2024b] J. Hloušek, et al., A many-particle bosonic quantum Maxwell demon, arXiv:2412.03241 (2024).

[Hoffman2021] D. P. Hoffman, et al., The promise and peril of deep learning in microscopy, Nat. Methods 18, 131 (2021).

[Hopfield1982] J. J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, Proc. Natl. Acad. Sci. U.S.A. 79, 2554 (1982).

[Hornik1991] K. Hornik, Approximation capabilities of multilayer feedforward networks, Neural Netw. 4, 251–257 (1991).

[Horodecki2002] P. Horodecki, A. Ekert, Method for direct detection of quantum entanglement, Phys. Rev. Lett. 89, 127902 (2002).

[Horodecki2003] P. Horodecki, Measuring quantum entanglement without prior state reconstruction, Phys. Rev. Lett. 90, 167901 (2003).

[Horodecki2009] R. Horodecki, et al., Quantum entanglement, Rev. Mod. Phys. 81, 865 (2009).

[Hošák2018] R. Hošák and M. Ježek, Arbitrary digital pulse sequence generator with delay-loop timing, Rev. Sci. Instrum. 89, 045103 (2018).

[Hradil2004] Z. Hradil, et al., Maximum-Likelihood Methods in Quantum Mechanics. In: Paris, M., Řeháček, J. (eds) Lecture Notes in Physics, page 649 (Springer, 2004).

[Huang2020] H.-Y. Huang, et al., Predicting many properties of a quantum system from very few measurements, Nat. Phys. 16, 1050 (2020).

[Huang2024] Z. Huang and L. Cao, Quantitative phase imaging based on holography: trends and new perspectives, Light Sci. Appl. 13, 145 (2024).

[Huisman2014] S. R. Huisman, et al., Programming balanced optical beam splitters in white paint, Opt. Express 22, 8320 (2014).

[Huisman2015] S. R. Huisman, et al., Programmable multiport optical circuits in opaque scattering materials, Opt. Express 23, 3102 (2015).

[Hüpfl2024] J. Hüpfl, et al., Continuity equation for the flow of Fisher information in wave scattering, Nat. Phys. 20, 1294 (2024).

[Impertro2023] A. Impertro, et al., An unsupervised deep-learning algorithm for single-site reconstruction in quantum gas microscopes, Commun. Phys. 6, 166 (2023).

[Irwin2005] K. D. Irwin and G. C. Hilton, Transition-Edge Sensors, In: Cryogenic Particle Detection, Topics in Applied Physics 99, 63 (Springer, 2005).

[Islam2015] R. Islam, et al., Measuring entanglement entropy in a quantum many-body system, Nature 528, 77 (2015).

[Israel2017] Y. Israel, et al, Quantum correlation enhanced super-resolution localization microscopy enabled by a fibre bundle camera, Nat Commun. 8, 14786 (2017).

[Jakeman1986] E. Jakeman and J. Rarity, The use of pair production processes to reduce quantum noise in transmission measurements," Opt. Commun. 59, 219 (1986).

[Jameson2010] D. M. Jameson and J. A. Ross, Fluorescence polarization in biomolecular research, Chem. Rev. 110, 2685 (2010).

[Jayakumar2013] H. Jayakumar, et al., Deterministic photon pairs and coherent optical control of a single quantum dot, Phys. Rev. Lett. 110, 135505 (2013).

[Ježek2003] M. Ježek, et al., Quantum inference of states and processes, Phys. Rev. A 68, 012305 (2003).

[Jin2014] H. Jin, et al., On-chip generation and manipulation of entangled photons based on reconfigurable lithium-niobate waveguide circuits, Phys. Rev. Lett. 113, 103601 (2014).

[Joshi2023] M. K. Joshi, et al., Exploring large-scale entanglement in quantum simulation, Nature 624, 539 (2023).

[Juhl2019] M. Juhl and K. Leosson, Polarimetry with disordered photonic structures, ACS Photonics 7, 203 (2019).

[Juhl2020] M. Juhl and K. Leosson, Polarization information in speckle patterns, ACS Photonics 7, 203 (2020).

[Jumper2021] J. Jumper, et al., Highly accurate protein structure prediction with AlphaFold, Nature 596, 583 (2021).

[Jungnitsch2011] B. Jungnitsch, et al., Taming multiparticle entanglement, Phys. Rev. Lett. 106, 190502 (2011).

[Kalashnikov2011] D. A. Kalashnikov, et al., Accessing photon bunching with a photon number resolving multi-pixel detector, Opt. Express 19, 9352 (2011).

[Kalashnikov2016] D. A. Kalashnikov, et al., Infrared spectroscopy with visible light," Nat. Photonics 10, 98 (2016).

[Kalev2015] A. Kalev, et al., Quantum tomography protocols with positivity are compressed sensing protocols, npj Quantum Inf. 1, 15018f (2015).

[Kaneda2019] F. Kaneda and P. G. Kwiat, High-efficiency single-photon generation via large-scale active time multiplexing, Sci. Adv. 5, aaw8586 (2019).

[Karamlou2024] A. H. Karamlou, et al., Probing entanglement in a 2D hard-core Bose–Hubbard lattice, Nature 629, 561 (2024).

[Kardynał2008] B. E. Kardynał, et al., An avalanche-photodiode-based photon-number-resolving detector, Nat. Photonics 2, 425 (2008).

[Karniadakis2021] G. E. Karniadakis, et al., Physics-informed machine learning, Nat. Rev. Phys. 3, 422 (2021).

[Kaszlikowski2008] D. Kaszlikowski, et al., Quantum correlation without classical correlations, Phys. Rev. Lett. 101, 070502 (2008).

[Katz2014] O. Katz, et al., Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations, Nat. Photonics 8, 784 (2014).

[Kaufman2016] A. M. Kaufman, et al., Quantum thermalization through entanglement in an isolated many-body system, Science 353, 794 (2016).

[Kauten2014] T. Kauten, et al., Measurement and modeling of the nonlinearity of photovoltaic and Geiger-mode photodiodes," Rev. Sci. Instrum. 85, 063102 (2014).

[Kauten2017] T. Kauten, et al., Obtaining tight bounds on higher-order interferences with a 5-path interferometer, New J. Phys. 19, 033017 (2017).

[Kawaguchi2022] K. Kawaguchi, et al., Generalization in Deep Learning, in Mathematical Aspects of Deep Learning, pp. 112–148 (Cambridge University Press, 2022).

[Kawasaki2025] A. Kawasaki, et al., Real-time observation of picosecond-timescale optical quantum entanglement towards ultrafast quantum information processing, Nat. Photonics 19, 271 (2025).

[Kendall2017] A. Kendall, Y. Gal, What uncertainties do we need in Bayesian deep learning for computer vision?, Adv. Neural Inf. Process. Syst. 30, 5574 (2017).

[Kerman2013] A. J. Kerman, et al., Readout of super-conducting nanowire single-photon detectors at high count rates, J. Appl. Phys. 113, 144511 (2013).

[Ketterer2019] A. Ketterer, et al., Characterizing multipartite entanglement with moments of random correlations, Phys. Rev. Lett. 122, 120505 (2019).

[Khemani2017] V. Khemani, et al., Critical properties of the many-body localization transition, Phys. Rev. X 7, 021013 (2017).

[Kim2012] M. Kim, et al., Maximal energy transport through disordered media with the implementation of transmission eigenchannels, Nat. Photonics 6, 581 (2012).

[Kingma2015] D. P. Kingma, et al., Adam: A method for stochastic optimization, International Conference on Learning Representations (ICLR), arXiv:1412.6980 (2015).

[Kitagawa2012] T. Kitagawa, et al., Observation of topologically protected bound states in photonic quantum walks, Nat. Commun. 3, 882 (2012).

[Knips2020] L. Knips, et al., Multipartite entanglement analysis from random correlations, npj Quantum Inf. 6, 51 (2020).

[Kokail2021] C. Kokail, et al., Entanglement Hamiltonian tomography in quantum simulation, Nat. Phys. 17, 936 (2021).

[Kornilov2014] V. Kornilov, Effects of dead time and afterpulses in photon detector on measured statistics of stochastic radiation, J. Opt. Soc. Am. A 31, 7–15 (2014).

[Korzh2020] B. Korzh, et al., Demonstration of sub-3 ps temporal resolution with a superconducting nanowire single-photon detector, Nat. Photonics 14, 250 (2020).

[Koutný2023] D. Koutný, et al., Deep learning of quantum entanglement from incomplete measurements, Sci. Adv. 9, eadd7131 (2023).

[Kramer1991] M. A. Kramer, Nonlinear principal component analysis using autoassociative neural networks, AIChE J. 37, 233 (1991).

[Krenn2023] M. Krenn, et al., Artificial intelligence and machine learning for quantum technologies, Phys. Rev. A 107, 010101 (2023).

[Krenn2025] M. Krenn, et al., Digital discovery of interferometric gravitational wave detectors, Phys. Rev. X 15, 021012 (2025).

[Kröger2017] J. Kröger, et al., High intensity click statistics from a 10 × 10 avalanche photodiode array, J. Phys. B: At. Mol. Opt. Phys. 50, 214003 (2017).

[Kudyshev2021] Z. A. Kudyshev, et al., Machine learning for integrated quantum photonics, ACS Photonics 8, 34 (2021).

[Kumar2012] R. Kumar, et al., Versatile wideband balanced detector for quantum optical homodyne tomography, Optics Communications 285, 5259 (2012).

[Lachman2019] L. Lachman, et al., Faithful hierarchy of genuine n-photon quantum non-gaussian light, Phys. Rev. Lett. 123, 043601 (2019).

[Lambert2013] N. Lambert, et al., Quantum biology, Nat. Phys. 9, 10 (2013).

[Lanyon2009] B. P. Lanyon, et al., Simplifying quantum logic using higher-dimensional Hilbert spaces, Nat. Phys. 5, 134 (2009).

[Lanyon2017] B. P. Lanyon, et al., Efficient tomography of a quantum many-body system. Nat. Phys. 13, 1158–1162 (2017).

[Larsen2019] M. V. Larsen, et al., Deterministic generation of a two-dimensional cluster state, Science 366, 369 (2019).

[Larsen2021] M. V. Larsen, et al., Deterministic multi-mode gates on a scalable photonic quantum computing platform, Nat. Phys. 17, 1018 (2021).

[LeCun1998] Y. LeCun, et al., Gradient-based learning applied to document recognition, Proc. IEEE, 86, 2278 (1998).

[LeCun2015][2] Y. LeCun, et al., Deep learning, Nature 521, 436 (2015).

[Lee2016] J. D. Lee, et al., Gradient descent only converges to minimizes, In V. Feldman, A. Rakhlin and O. Shamir, eds., 29th Annual Conference on Learning Theory, vol. 49 of Proceedings of Machine Learning Research, pp. 1246–1257 (PMLR, Columbia University, 2016).

[Leedumrongwatthanakun2020] S. Leedumrongwatthanakun, et al., Programmable linear quantum networks with a multimode fibre, Nat. Photonics 14, 139 (2020).

[Leedumrongwatthanakun2025] S. Leedumrongwatthanakun, et al., Classification of quantum states of light using random measurements through a multimode fiber, Phys. Rev. Res. 7, 023222 (2025).

[Lelek2021] M. Lelek, et al., Single-molecule localization microscopy, Nat. Rev. Methods Primers 1, 39 (2021).

[Lerner2018] E. Lerner, et al., Toward dynamic structural biology: Two decades of single-molecule Förster resonance energy transfer, Science 359, 288 (2018).

[Leshno1993] M. Leshno, et al., Multilayer feedforward networks with a nonpolynomial activation function can approximate any function, Neural Netw. 6, 861–867 (1993).

[Li2018] B.-B. Li, et al., Quantum enhanced optomechanical magnetometry, Optica 5, 850 (2018).

[Li2023] W. Li, et al., Estimation of the number of single-photon emitters for multiple fluorophores with the same spectral signature, AVS Quantum Sci. 5, 041401 (2023).

[Li2025] Z. Li, et al., Boosting photon-number-resolved detection rates of transition-edge sensors by machine learning, Optica Quantum 3, 246 (2025).

[Liao2017] S.-K. Liao, et al., Satellite-to-ground quantum key distribution, Nature 549, 43 (2017).

[Lidke2005] K. A. Lidke, et al., Superresolution by localization of quantum dots using blinking statistics, Opt. Express 13, 7052 (2005).

[LIGO2016] B. P. Abbott, et al. (LIGO Scientific Collaboration and Virgo Collaboration), Observation of gravitational waves from a binary black hole merger, Phys. Rev. Lett. 116, 061102 (2016).

[Ling2006] A. Ling, et al., Experimental polarization state tomography using optimal polarimeters, Phys. Rev. A 74, 022309 (2006).

[LiQ2023] Q. Li, et al., Single-photon absorption and emission from a natural photosynthetic complex, Nature 619, 300 (2023).

[LiR2024] R. Li, et al., Photonics for neuromorphic computing: Fundamentals, devices, and opportunities, Advanced Materials 37, e2312825 (2024).

[LiS2024] S. Li, et al., Localising two sub-diffraction emitters in 3D using quantum correlation microscopy, New J. Phys. 26, 033036 (2024).

[Lita2022] A. Lita, et al., Development of superconducting single-photon and photon-number resolving detectors for quantum applications, J. Light. Technol. 40, 7578 (2022).

[Liu2021] J. Liu, et al., Broadband spectroscopic ellipsometry, Light: Sci. Appl. 10, 55 (2021).

[Liu2025] T. Liu, et al., Bayesian deep-learning structured illumination microscopy enables reliable super-resolution imaging with uncertainty quantification, Nat. Commun. 16, 5027 (2025).

[LiuH2025] H.-L. Liu, et al., Robust quantum computational advantage with programmable 3050-photon Gaussian boson sampling, preprint Research Square DOI: 10.21203/rs.3.rs-7435859/v1 (2025).

[LiuZ2025] Z.-H. Liu, et al., Quantum learning advantage on a scalable photonic platform, Science 389, 1332 (2025).

[Lohrmann2019] A. Lohrmann, et al., Manipulation and measurement of quantum states with liquid crystal devices, Opt. Express 27, 13765 (2019).

[Los2024] J. W. Niels Los, et al., High-performance photon number resolving detectors for 850–950 nm wavelength range, APL Photonics 9, 066101 (2024).

[Lu2009] C.-Y. Lu, et al., Demonstrating anyonic fractional statistics with a six-qubit quantum simulator, Phys. Rev. Lett. 102, 030502 (2009).

[Lu2016] D. Lu, et al., Tomography is necessary for universal entanglement detection with single-copy observables, Phys. Rev. Lett. 116, 230501 (2016).

[Lu2017] Z. Lu, et al., The expressive power of neural networks: A view from the width, Adv. Neural Inf. Process. Syst. 30 (2017).

[Lu2019] X. Lu, et al., Chip-integrated visible–telecom entangled photon pair source for quantum communication, Nat. Phys. 15, 373 (2019).

[Lundberg2017] S. M. Lundberg and S.-I. Lee, A unified approach to interpreting model predictions, Advances in Neural Information Processing Systems (NeurIPS), vol. 30, arXiv:1705.07874 (2017).

[Lundberg2020] S. M. Lundberg, et al., From local explanations to global understanding with explainable AI for trees, Nat. Mac. Intell. 2, 56-67 (2020).

[Lundeen2009] J. S. Lundeen, et al., Tomography of quantum detectors, Nat. Phys. 5, 27 (2009).

[Luo2015] W. Luo, et al., Synthetic aperture-based on-chip microscopy, Light: Sci. Appl. 4, e261 (2015).

[Luo2019] K.-H. Luo, et al., Nonlinear integrated quantum electro-optic circuits, Sci. Adv. 5, eaat1451 (2019).

[Lvovsky2001] A. I. Lvovsky, et al., Quantum state reconstruction of the single-photon Fock state, Phys. Rev. Lett. 87, 050402 (2001).

[Lvovsky2009] A. I. Lvovsky and M. G. Raymer, Continuous-variable optical quantum-state tomography, Rev. Mod. Phys. 81, 299 (2009).

[Lvovsky2020] A. I. Lvovsky, et al., Production and applications of non-Gaussian quantum states of light, arXiv:2006.16985 (2020).

[Ma2021] W. Ma, et al., Deep learning for the design of photonic structures, Nat. Photonics 15, 77 (2021).

[Maas2013] A. Maas, et al., Rectifier nonlinearities improve neural network acoustic models, Proceedings of the 30 th International Conference on Machine Learning, Atlanta, Georgia, USA; JMLR W&CP 28 (2013).

[Madsen2022] L. S. Madsen, et al., Quantum computational advantage with a programmable photonic processor, Nature 606, 75 (2022).

[Magana-Loaiza2016] O. S. Magana-Loaiza, et al., Exotic looped trajectories of photons in three-slit interference, Nat. Commun. 7, 13987 (2016).

[Maier2019] Ch. Maier, et al., Environment-assisted quantum transport in a 10-qubit network, Phys. Rev. Lett. 122, 050501 (2019).

[Main2025] D. Main, et al., Distributed quantum computing across an optical network link, Nature 638, 383 (2025).

[Malinin2018] A. Malinin and M. Gales, Predictive uncertainty estimation via prior networks, Adv. Neural Inf. Process. Syst. 31, 7047 (2018).

[Maring2024] N. Maring, et al., A versatile single-photon-based quantum computing platform, Nat. Photonics 18, 603 (2024).

[Mark2024] P. Mark, et al., Nonlinear response of telecom-wavelength superconducting single-photon detectors, APL Quantum 1, 046109 (2024).

[Marquez2000] A. Marquez, et al., Characterization of edge effects in twisted nematic liquid crystal displays, Opt. Eng. 39, 3301 (2000).

[Marquez2001] A. Marquez, et al., Prediction of modulation behavior in TNLCs, Opt. Eng. 40, 2558 (2001).

[Marshall2025] M. C. Marshall, et al., High-stability single-ion clock with 5.5×10^{-19} systematic uncertainty, Phys. Rev. Lett. 135, 033201 (2025).

[Marsili2009] F. Marsili, et al., Physics and application of photon number resolving detectors based on superconducting parallel nanowires, New J. Phys. 11, 045022 (2009).

[Marsili2016] F. Marsili, et al., Hotspot relaxation dynamics in a current-carrying superconductor, Phys. Rev. B 93, 094518 (2016).

[Matthews2009] J. C. F. Matthews, et al., Manipulation of multiphoton entanglement in waveguide quantum circuits, Nat. Photonics 3, 346 (2009).

[Matthews2013] J. C. F. Matthews, et al., Observing fermionic statistics with photons in arbitrary processes, Sci. Rep. 3, 1539 (2013).

[Mattioli2016] F. Mattioli, et al., Photon-counting and analog operation of a 24-pixel photon number resolving detector based on superconducting nanowires, Opt. Express 24, 9067 (2016).

[Mazin2021] G. Mazin, et al., Non-blocking programmable delay line with minimal dead time and tens of picoseconds jitter, Rev. Sci. Instrum. 92, 114712 (2021).

[Mazin2025] G. Mazin, et al., Low loss single-photon multi-level programmable electro-optical light control enabled by FPGA technology, in preparation (2025).

[McCulloch1943] W. S. McCulloch and W. Pitts, A logical calculus of the ideas immanent in nervous activity, Bull. Math. Biophys. 5, 115 (1943).

[McDonald2019] M. McDonald, et al., Super-resolution microscopy of cold atoms in an optical lattice, Phys. Rev. X 9, 021001 (2019).

[Meany2015] T. Meany, et al., Laser written circuits for quantum photonics, Laser Photonics Rev. 9, 363 (2015).

[Metzger2017] N. K. Metzger, et al., Harnessing speckle for a sub-femtometre resolved broadband wavemeter and laser stabilization, Nat. Commun. 8, 15610 (2017).

[Mezzadri2007] F. Mezzadri, How to generate random matrices from the classical compact groups, Not. Am. Math. Soc. 54, 592–604 (2007).

[Mičuda2008] M. Mičuda, et al., High-efficiency photon-number-resolving multichannel detector, Phys. Rev. A 78, 025804 (2008).

[Mičuda2013] M. Mičuda, et al., Efficient Experimental Estimation of Fidelity of Linear Optical Quantum Toffoli Gate, Phys. Rev. Lett. 111, 160407 (2013).

[Mičuda2019] M. Mičuda, et al., Verifying genuine multipartite entanglement of the whole from its separable parts, Optica 6, 896 (2019).

[Migdal2013] A. Migdall, et al., Single-photon generation and detection (Academic Press, 2013).

[Ming2019] Y. Ming, et al., Quantum topology identification with deep neural networks and quantum walks, npj Comput. Mater. 5, 88 (2019).

[Moen2019] E. Moen, et al., Deep learning for cellular image analysis, Nat. Methods 16, 1233 (2019).

[Mogilevtsev1998] D. Mogilevtsev, Diagonal element inference by direct detection, Opt. Commun. 156, 307 (1998).

[Monz2011] T. Monz, et al., 14-Qubit entanglement: Creation and coherence, Phys. Rev. Lett. 106, 130506 (2011).

[Moreau2019] P.-A. Moreau, et al., Imaging with quantum states of light, Nat. Rev. Phys. 1, 367 (2019).

[Morvan2024] A. Morvan, et al., Phase transitions in random circuit sampling, Nature 634, 328 (2024).

[Mose2023] S. A. Moses, et al., A race-track trapped-ion quantum processor, Phys. Rev. X 13, 041052 (2023).

[Mosk2012] A. P. Mosk et al., Controlling waves in space and time for imaging and focusing in complex media, Nat. Photonics 6, 283 (2012).

[Mukundarajan2017] H. Mukundarajan, et al., Using mobile phones as acoustic sensors for high-throughput mosquito surveillance, eLife 6, e27854 (2017).

[Naito2006] K. Naito, et al., Single-molecule detection of airborne singlet oxygen, J. Am. Chem. Soc. 128, 16430–16431 (2006).

[Natarajan2012] C. M. Natarajan, et al., Superconducting nanowire single-photon detectors: physics and applications, Supercond. Sci. Technol. 25, 063001 (2012).

[Nehme2018] E. Nehme, et al., Deep-STORM: super-resolution single-molecule microscopy by deep learning, Optica 5, 458 (2018).

[Nix1994] D. A. Nix and A. S. Weigend, Estimating the mean and variance of the target probability distribution, Proc. 1994 IEEE Int. Conf. Neural Networks (ICNN'94) 1, 55 (1994).

[NobelPrize2022] A. Zeilinger, A. Aspect, and J. F. Clauser, Nobel Prize lectures, Aula Magna, Stockholm University, 8 Dec 2022.

[Nomad2022] C. Audet, et al., Algorithm 1027: NOMAD version 4: Nonlinear optimization with the MADS algorithm, ACM Trans. Math. Softw. 48, 1 (2022).

[O'Brien2003] J. L. O'Brien, et al., Demonstration of an all-optical quantum controlled-NOT gate, Nature 426, 264 (2003).

[Ono2013] R. Ono, et al., An entanglement-enhanced microscope, Nat. Commun. 4, 2426 (2013).

[Oren2017] D. Oren, et al., Quantum state tomography with a single measurement setup, Optica 4, 993 (2017).

[Orieux2013] A. Orieux, et al., Experimental detection of quantum channels, Phys. Rev. Lett. 111, 220501 (2013).

[Oripov2023] B. G. Oripov, et al., A superconducting nanowire single-photon camera with 400,000 pixels, Nature 622, 730 (2023).

[Osborne2002] T. J. Osborne and M. A. Nielsen, Entanglement in a simple quantum phase transition, Phys. Rev. A 66, 032110 (2002).

[Osterloh2002] A. Osterloh, et al., Scaling of entanglement close to a quantum phase transition, Nature 416, 608 (2002).

[Ourjoumtsev2006] A. Ourjoumtsev, et al., Generating optical Schrödinger kittens for quantum information processing, Science 312, 83 (2006).

[Palmieri2020] A. M. Palmieri, et al., Experimental neural network enhanced quantum tomography, npj Quantum Inf. 6, 20 (2020).

[Pan1998] J.-W. Pan, et al., Experimental entanglement swapping: Entangling photons that never interacted, Phys. Rev. Lett. 80, 3891 (1998).

[Panitchayangkoon2010] G. Panitchayangkoon, et al., Long-lived quantum coherence in photosynthetic complexes at physiological temperature, Proc. Natl. Acad. Sci. U.S.A. 107, 12766 (2010).

[Paris2004] M. G. A. Paris and J. Řeháček, Quantum State Estimation (Springer, 2004).

[Paul1996] H. Paul, et al., Photon chopping: New way to measure the quantum state of light, Phys. Rev. Lett. 76, 2464 (1996).

[Peinado2010] A. Peinado, et al., Optimization of Stokes polarimeter based on retarders, Opt. Express 18, 9815 (2010).

[Pelliccione2016] M. Pelliccione, et al., Scanned probe imaging of nanoscale magnetism at cryogenic temperatures with a single-spin quantum sensor, Nat. Nanotechnol. 11, 700 (2016).

[Perets2008] H. B. Perets, et al., Realization of quantum walks with negligible decoherence in waveguide lattices, Phys. Rev. Lett. 100, 170506 (2008).

[Peruzzo2010] A. Peruzzo, et al., Quantum walks of correlated photons, Science 329, 1500 (2010).

[Peters2005] N. A. Peters, et al., Remote state preparation of photon polarization, Phys. Rev. Lett. 94, 150502 (2005).

[Pierangeli2023] D. Pierangeli and C. Conti, Single-shot polarimetry of vector beams by supervised learning, Nat. Commun. 14, 1831 (2023).

[Pinkus1999] A. Pinkus, Approximation theory of the MLP model in neural networks, Acta Numerica 8, 143–195 (1999).

[Pirandola2020] S. Pirandola, et al., Advances in quantum cryptography, Adv. Opt. Photonics. 12, 1012 (2020).

[Pitsios2017] I. Pitsios, et al., Photonic simulation of entanglement growth and engineering after a spin chain quench, Nat. Commun. 8, 1569 (2017).

[Polino2020] E. Polino, et al., Photonic quantum metrology, AVS Quantum Sci. 2, 024703 (2020).

[Politi2008] A. Politi, et al., Silica-on-Silicon Waveguide Quantum Circuits, Science 320, 646 (2008).

[PonceBobadilla2024] A. V. Ponce-Bobadilla, et al., Practical guide to SHAP analysis: Explaining supervised machine learning model predictions in drug development, Clin. Transl. Sci. 17, e11513550 (2024).

[Poonia2024] M. Poonia, et al., Raman and Surface-Enhanced Raman Scattering Detection in Flowing Solutions for Complex Mixture Analysis, Annu. Rev. Anal. Chem. 17, 411 (2024).

[Popoff2010] S. M. Popoff, et al., Measuring the transmission matrix in optics: An approach to the study and control of light propagation in disordered media, Phys. Rev. Lett. 104, 100601 (2010).

[PsiQuantum2025] PsiQuantum team, A manufacturable platform for photonic quantum computing, Nature 641, 876 (2025).

[Qi2017] J. Qi and D. S. Elson, Polarization imaging in biomedicine: emerging trends, J. Biophotonics 10, 950–982 (2017).

[Qi2018] L. Qi et al., Multiphoton nonclassical light from clusters of single-photon emitters, New J. Phys. 20, 073013 (2018).

[Qi2023] J. Qi, et al., Polarimetric endoscopy for cancer staging, Nat. Biomed. Eng. 7, 971–985 (2023).

[Qiang2018] X. Qiang, et al., Large-scale silicon quantum photonics implementing arbitrary two-qubit processing, Nat. Photonics 12, 534 (2018).

[Qiao2021] C. Qiao, et al., Evaluation and development of deep neural networks for image super-resolution in optical microscopy, Nat. Methods 18, 194 (2021).

[Rattenbacher2023] D. Rattenbacher, et al., On-chip interference of scattering from two individual molecules, Optica 10, 1595 (2023).

[Redding2013] B. Redding, et al., Compact spectrometer based on a disordered photonic chip, Nature Photonics 7, 746 (2013).

[Redding2014] B. Redding, et al., High-resolution and broadband all-fiber spectrometers, Optica 1, 175 (2014).

[Reddy2020] D. V. Reddy, et al., Superconducting nanowire single-photon detectors with 98% system detection efficiency at 1550 nm, Optica 7, 1649 (2020).

[Reh2021] M. Reh, et al., Time-dependent variational principle for open quantum systems with artificial neural networks, Phys. Rev. Lett. 127, 230501 (2021).

[Řeháček2003] J. Řeháček, et al., Multiple-photon resolving fiber-loop detector, Phys. Rev. A 67, 061801(R) (2003).

[Řeháček2004] J. Řeháček, et al., Minimal qubit tomography, Phys. Rev. A 70(5), 052321 (2004).

[Reiser2022] P. Reiser, et al., Graph neural networks for materials science and chemistry, Commun. Mater. 3, 93 (2022).

[Ren2022] P. Ren, et al., Photonic-circuited resonance fluorescence of single molecules with an ultrastable lifetime-limited transition, Nat. Commun. 13, 3982 (2022).

[Reuter2014] M. Reuter, et al., BRCA2 diffuses as oligomeric clusters with RAD51 and changes mobility after DNA damage in live cells, Journal of Cell Biology 207, 599 (2014).

[Rimoli2022] C. V. Rimoli, et al., Polarization-resolved nonlinear microscopy in cell biology, Nat. Commun. 13, 301 (2022).

[Rivenson2017] Y. Rivenson, et al., Deep learning microscopy, Optica 4, 1437 (2017).

[Rivenson2018] Y. Rivenson, et al., Phase recovery and holographic image reconstruction using deep learning in neural networks, Light: Sci. Appl. 7, 17141 (2018).

[Robbins1951] H. Robbins and S. Monro, A stochastic approximation method, Ann. Math. Stat. 22, 400 (1951).

[Rocchetto2019] A. Rocchetto, et al., Experimental learning of quantum states, Sci. Adv. 5, eaau1946 (2019).

[Rodríguez2024] C. Rodríguez, et al., Automated discovery of experimental designs in super-resolution microscopy with XLuminA, Nat. Commun. 15, 10658 (2024).

[Roesel2025] T. Roesel, et al., Label-free single-molecule optical detection, npj Biosens. 2, 32 (2025).

[Roik2021] J. Roik, et al., Accuracy of entanglement detection via artificial neural networks and human-designed entanglement witnesses, Phys. Rev. Appl. 15, 054006 (2021).

[Ronceray2025] N. Ronceray, et al., Wide-field fluorescence lifetime imaging of single molecules with a gated single-photon camera, Light Sci. Appl. 14, 258 (2025).

[Ronneberger2015] O. Ronneberger, et al., U-Net: Convolutional networks for biomedical image segmentation, MICCAI 2015, LNCS 9351, pp. 234–241 (2015).

[Rosenblatt1958] F. Rosenblatt, The perceptron: A probabilistic model for information storage and organization in the brain, Psychol. Rev. 65, 386 (1958).

[Rossia2005] A. R. Rossia and M. G. A. Paris, A two-step MaxLik-MaxEnt strategy to infer photon distribution from on/off measurement at low quantum efficiency, Eur. Phys. J. D 32, 223 (2005).

[Rotter2017] S. Rotter and S. Gigan, Light fields in complex media: Mesoscopic scattering meets wave control, Rev. Mod. Phys. 89, 015005 (2017).

[Rubin2022] N. A. Rubin, et al., Dielectric metasurfaces for polarization analysis, Opt. Express 30, 9389 (2022).

[Rust2006] M. J. Rust, et al., Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM), Nat. Methods 3, 793 (2006).

[Sabines-Chesterking2019] J. Sabines-Chesterking, et al., Twin-beam sub-shot-noise raster-scanning microscope, Opt. Express 27, 30810 (2019).

[Sage2015] D. Sage, et al., Quantitative evaluation of software packages for single-molecule localization microscopy, Nat. Methods 12, 717 (2015).

[Sahl2017] S. J. Sahl, et al., Fluorescence nanoscopy in cell biology, Nat. Rev. Mol. Cell Biol. 18, 685 (2017).

[Sansoni2010] L. Sansoni, et al., Polarization entangled state measurement on a chip, Phys. Rev. Lett. 105, 200503 (2010).

[Sansoni2012] L. Sansoni, et al., Two-particle Bosonic-Fermionic quantum walk via integrated photonics, Phys. Rev. Lett. 108, 010502 (2012).

[Scarselli2009] F. Scarselli, et al., The graph neural network model, IEEE Transactions on Neural Networks 20, 61 (2009).

[Schaefer1983] A. R. Schaefer, et al., Silicon detector nonlinearity and related effects," Appl. Opt. 22, 1232 (1983).

[Schapeler2024] T. Schapeler, et al., Electrical trace analysis of superconducting nanowire photon-number-resolving detectors, Phys. Rev. Applied 22, 014024 (2024).

[Schapeler2025] T. Schapeler, et al., Scalable quantum detector tomography by high-performance computing, Quantum Sci. Technol. 10 015018 (2025).

[Schreiber2010] A. Schreiber, et al., Photons walking the line: A quantum walk with adjustable coin operations, Phys. Rev. Lett. 104, 050502 (2010).

[Schreiber2011] A. Schreiber, et al., Decoherence and disorder in quantum walks: From ballistic spread to localization, Phys. Rev. Lett. 106, 180403 (2011).

[Schreiber2012] A. Schreiber, et al., A 2D quantum walk simulation of two-particle dynamics, Science 336, 55 (2012).

[Schuck2016] C. Schuck, et al., Quantum interference in heterogeneous superconducting-photonic circuits on a silicon chip, Nat. Commun. 7, 10352 (2016).

[Schwemmer2015] C. Schwemmer, et al., Systematic errors in current quantum state tomography tools, Phys. Rev. Lett. 114, 080403 (2015).

[Segev2013] M. Segev, et al., Anderson localization of light, Nat. Photonics 7, 197 (2013).

[Sensoy2018] M. Sensoy, et al., Evidential deep learning to quantify classification uncertainty, Adv. Neural Inf. Process. Syst. 31, 3183 (2018).

[Shalm2015] L. K. Shalm, et al., Strong loophole-free test of local realism, Phys. Rev. Lett. 115, 250402 (2015).

[Shastri2021] B. J. Shastri, et al., Photonics for artificial intelligence and neuromorphic computing, Nat. Photonics 15, 102 (2021).

[Shibata2014] H. Shibata, et al., Quantum key distribution over a 72 dB channel loss using ultralow dark count superconducting single-photon detectors, Opt. Lett. 39, 5078 (2014).

[Shibata2017] H. Shibata, et al., SNSPD with ultimate low system dark count rate using various cold filters, IEEE Trans. Appl. Supercond. 27, 1 (2017).

[Srivastava2014] N. Srivastava, et al, Dropout: A simple way to prevent neural networks from overfitting, J. Mach. Learn. Res. 15, 1929 (2014).

[Silva2017] G. B. Silva, et al., Investigating bias in maximum-likelihood quantum-state tomography, Phys. Rev. A 95, 022107 (2017).

[Silverstone2014] J. W. Silverstone, et al., On-chip quantum interference between silicon photon-pair sources, Nat. Photonics 8, 104 (2014).

[Singh2025] A. Singh, et al., Photonic quantum information with time-bins: Principles and applications, arXiv:2507.08102 (2025).

[Sinha2010] U. Sinha, et al., Ruling out multi-order interference in quantum mechanics, Science 329, 418 (2010).

[Sluijterman2024] L. A. E. Sluijterman, et al., Optimal training of mean–variance estimation neural networks, Neurocomputing 597, 127929 (2024).

[Slussarenko2017] Slussarenko, et al., Unconditional violation of the shot-noise limit in photonic quantum metrology, Nat. Photonics 11, 700–703 (2017).

[Slussarenko2019] S. Slussarenko and G. J. Pryde, Photonic quantum information processing: A concise review, Appl. Phys. Rev. 6, 041303 (2019).

[Smith2009] B. J. Smith, et al., Phase-controlled integrated photonic quantum circuits, Opt. Express 17, 13516 (2009).

[Smith2010] C. S. Smith, et al., Fast, single-molecule localization that achieves theoretically minimum uncertainty, Nat. Methods 7, 373 (2010).

[Smithey1993] D. T. Smithey, et al., Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: Application to squeezed states and the vacuum, Phys. Rev. Lett. 70, 1244 (1993).

[SMLM-Challenge] N. Olivier, et al., Super-resolution microscopy hub, https://srm.epfl.ch/srm/dataset/challenge-2D-real/Real_High_Density/index.html (2013).

[Smolensky1986] P. Smolensky, Information processing in dynamical systems: Foundations of harmony theory, in: Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1: Foundations, eds. D. E. Rumelhart & J. L. McClelland, pp. 194–281 (MIT Press, 1986).

[Solli2008] D. R. Solli, et al., Amplified wavelength-time transformation for real-time spectroscopy, Nat. Photonics 2, 48 (2008).

[Solli2012] D. R. Solli, et al., Fluctuations and correlations in modulation instability, Nat. Photonics 6, 463 (2012).

[Sollner2012] I. Sollner, et al., Testing Born's rule in quantum mechanics for three mutually exclusive events, Found. Phys. 42, 742 (2012).

[Song2017] C. Song, et al., 10-Qubit Entanglement and parallel logic operations with a superconducting circuit, Phys. Rev. Lett. 119, 180511 (2017).

[Špačková2022] B. Špačková, et al., Label-free nanofluidic scattering microscopy of size and mass of single diffusing molecules and nanoparticles, Nat. Methods 19, 751 (2022).

[Spagnolo2008] N. Spagnolo, et al., Polarization-preserving ultrafast optical shutter, Opt. Express 16, 17609 (2008).

[Speiser2021] A. Speiser, et al., Deep learning enables fast and dense single-molecule localization with high accuracy, Nat. Methods 18, 1082 (2021).

[Sperling2012] J. Sperling, et al., True photocounting statistics of multiple on–off detectors, Phys. Rev. A 85, 023820 (2012).

[Stárek2018a] R. Stárek, et al., Nondestructive detector for exchange symmetry of photonic qubits, npj Quantum Inf. 4, 35 (2018).

[Stárek2018b] R. Stárek, et al., Experimental realization of SWAP operation on hyper-encoded qubits, Opt. Express 26, 8443 (2018).

[Stárek2022] R. Stárek, Experimental implementation of complex multiqubit quantum logic gates on platform of linear optics and their applications, Ph.D. thesis, Palacký University Olomouc (2022).

[Stárek2025] R. Stárek, et al., Measurement-device agnostic quantum tomography, Quantum Sci. Technol. 10, 045073 (2025).

[Starkov2009] V. N. Starkov, et al., Numerical reconstruction of photon-number statistics from photocounting statistics: Regularization of an ill-posed problem, Phys. Rev. A 80, 013813 (2009).

[Stensson2018] K. Stensson and G. Björk, Measurement of the two-time intensity-correlation function of arbitrary states, Phys. Rev. A 98, 033812 (2018).

[Stibůrek2023] M. Stibůrek, et al., 110 μm thin endo-microscope for deep-brain in vivo observations of neuronal connectivity, activity and blood flow dynamics, Nat. Commun. 14, 1897 (2023).

[Stipčević2013] M. Stipčević, et al., Characterization of a commercially available large area, high detection efficiency single-photon avalanche diode, J. Lightwave Technol. 31, 3591 (2013).

[Straka2015] I. Straka, et al., Conditional cooling limit for a quantum channel going through an incoherent environment, Sci. Rep. 5, 16721 (2015).

[Straka2018a] I. Straka, et al., Quantum non-Gaussian multiphoton light, npj Quant. Information 4, 4 (2018).

[Straka2018b] I. Straka, et al., Generator of arbitrary classical photon statistics, Opt. Express 26, 8998 (2018).

[Straka2020] I. Straka, et al., Counting statistics of actively quenched SPADs under continuous illumination, J. Lightwave Technol. 38, 4765-4771 (2020).

[Struchalin2021] G. I. Struchalin, et al., Experimental estimation of quantum state properties from classical shadows, PRX Quantum 2, 010307 (2021).

[Sun2024] B. Sun, et al., Single-photon time-stretch infrared spectroscopy, Laser Photonics Rev. 18, 2301272 (2024).

[Švarc2019] V. Švarc, et al., Fully tunable and switchable coupler for photonic routing in quantum detection and modulation, Opt. Lett. 44, 5844 (2019).

[Švarc2020] V. Švarc, et al., Feedforward-enhanced Fock state conversion with linear optics, Opt. Express 28, 11634 (2020).

[Švarc2023] V. Švarc, et al., Sub-0.1 degree phase locking of a single-photon interferometer, Opt. Express 31, 12562 (2023).

[Švarc2024] V. Švarc, Photonic switching for quantum information processing, Ph.D. thesis, Palacký University Olomouc (2024).

[Szabó2020] A. Szabó and C. Castelnovo, Neural network wave functions and the sign problem, Phys. Rev. Res. 2, 033075 (2020).

[Ta2015] H. Ta, et al., Mapping molecules in scanning far-field fluorescence nanoscopy, Nat. Commun. 6, 7977 (2015).

[Takeda2019] S. Takeda, et al., On-demand photonic entanglement synthesizer, Sci. Adv. 5, aaw4530 (2019).

[Tanabe2016] M. Tanabe, et al., Spectral supralinearity of silicon photodiodes in visible light due to surface recombination," Appl. Opt. 55, 3084 (2016).

[Taylor2013] M. A. Taylor, et al., Biological measurement beyond the quantum limit, Nat. Photonics 7, 229 (2013).

[Thomas2010] O. Thomas, et al., Efficient photon number detection with silicon avalanche photodiodes, Appl. Phys. Lett. 97, 031102 (2010).

[Thomas2012] O. Thomas, et al., Practical photon number detection with electric field-modulated silicon avalanche photodiodes, Nat. Commun. 3, 644 (2012).

[Tiedau2019] J. Tiedau, et al., A high dynamic range optical detector for measuring single photons and bright light, Opt. Express 27, 1 (2019).

[Tikhonov1995] A. N. Tikhonov, et al., Numerical Methods for the Solution of Ill-Posed Problems (Springer-Verlag, 1995).

[Titchener2018] J. G. Titchener, et al., Scalable on-chip quantum state tomography, npj Quant. Information 4, 19 (2018).

[Tiunov2020] E. S. Tiunov, et al., Experimental quantum homodyne tomography via machine learning, Optica 7, 448 (2020).

[Toninelli2021] C. Toninelli, et al., Single organic molecules for photonic quantum technologies, Nat. Mater. 20, 1615–1628 (2021).

[Torlai2018] G. Torlai, et al., Neural-network quantum state tomography, Nat. Phys. 14, 447 (2018).

[Torlai2019] G. Torlai, et al., Integrating neural networks with a quantum simulator for state reconstruction, Phys. Rev. Lett. 123, 230504 (2019).

[Tóth2010] G. Tóth, et al., Permutationally invariant quantum tomography. Phys. Rev. Lett. 105, 250403 (2010).

[Tse2019] M. Tse, et al., Quantum-enhanced advanced LIGO detectors in the era of gravitational-wave astronomy, Phys. Rev. Lett. 123, 231107 (2019).

[Turtaev2018] S. Turtaev, et al., In vivo endoscopy with multimode fibers, Light: Sci. Appl. 7, 1 (2018).

[Ursin2007] R. Ursin, et al., Entanglement-based quantum communication over 144 km, Nat. Physics 3, 481 (2007).

[Valdez2017] M. A. Valdez, et al., Quantifying complexity in quantum phase transitions via mutual information complex networks, Phys. Rev. Lett. 119, 225301 (2017).

[Valentino2022] M. Valentino, et al., Polarization-based environmental monitoring, Sci. Total Environ. 815, 152708 (2022).

[vandeVen2022] G. M. van de Ven, et al., Three types of incremental learning, Nat. Mach. Intell. 4, 1185 (2022).

[vanEnk2012] S. J. van Enk and C. W. J. Beenakker, Measuring Tr ρ^n on single copies of ρ using random measurements, Phys. Rev. Lett. 108, 110503 (2012).

[vanLeent2022] T. van Leent, et al., Entangling single atoms over 33 km telecom fibre, Nature 607, 69 (2022).

[Vardi1993] Y. Vardi and D. Lee, From image deblurring to optimal investments: Maximum likelihood solutions for positive linear inverse problems, J. Royal Stat. Soc. B 55, 569 (1993).

[Vašinka2022] D. Vašinka, et al., Bidirectional deep learning of polarization transfer in liquid crystals with application to quantum state preparation, Phys. Rev. Applied 17, 054042 (2022).

[Vašinka2025a] D. Vašinka, et al., From stars to molecules: AI guided device-agnostic super-resolution imaging, preprint arXiv:2502.18637, bioRxiv:2025.02.25.640182 (2025).

[Vašinka2025b] D. Vašinka, et al., Universal super-resolution framework for imaging of quantum dots, preprint arXiv:2510.06076 (2025).

[Vasquez-Lozano2024] J. J. Vasquez-Lozano, et al., Scanning quantum correlation microscopy with few emitters, New J. Phys. 26, 043001 (2024).

[Večerka2025] R. Večerka, Photon counting and photon statistics reconstruction, Bc. thesis, Palacký University Olomouc (2025).

[Vellekoop2015] I. M. Vellekoop, et al., Feedback-based wavefront shaping, Opt. Express 23, 12189 (2015).

[Venza2025] F. P. Venza and M. Colangelo, Research trends in single-photon detectors based on superconducting wires, APL Photonics 10, 040901 (2025).

[Villar2020] A. Villar, et al., Entanglement demonstration on board a nano-satellite, Optica 7, 734 (2020).

[Vogel1989] K. Vogel and H. Risken, Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase. Phys. Rev. A 40, 2847 (1989).

[vonChamier2021] L. von Chamier, et al., Democratising deep learning for microscopy with ZeroCostDL4Mic, Nat. Commun. 12, 2276 (2021).

[Vynck2023] K. Vynck, et al., Scattering in correlated disordered media, Rev. Mod. Phys. 95, 045003 (2023).

[Walborn2006] S. P. Walborn, et al., Experimental determination of entanglement with a single measurement, Nature 440, 1022 (2006).

[Walter2013] M. Walter, et al., Entanglement polytopes: Multiparticle entanglement from single-particle information, Science 340, 1205 (2013).

[Walther2005] P. Walther, et al., Experimental one-way quantum computing, Nature 434, 169 (2005).

[Wang2014] J. Wang, et al., Gallium arsenide (GaAs) quantum photonic waveguide circuits, Opt. Commun. 327, 49 (2014).

[Wang2016a] F. Wang, et al., Non-Markovian property of afterpulsing effect in single-photon avalanche detector, J. Lightave Technol. 34, 3610–3615 (2016).

[Wang2016b] J. Wang, Advances in space-division multiplexed optical communications, Photonics Res. 4, B14 (2016).

[Wang2020] J. Wang, et al., Integrated photonic quantum technologies, Nat. Photonics 14, 273 (2020)

[Wang2022] Y. Wang, et al., Molecular contrastive learning of representations via graph neural networks, Nat. Mach. Intell. 4, 279–287 (2022).

[Wang2025] H. Wang, et al., Scalable photonic quantum technologies, Nat. Mater. (2025).

[WangH2019] H. Wang, et al., Boson sampling with 20 input photons and a 60-mode interferometer in a 10^{14}-dimensional Hilbert space, Phys. Rev. Lett. 123, 250503 (2019).

[WangK2018] K. Wang, et al., Quantum metasurface for multiphoton interference and state reconstruction, Science 361, 1104 (2018).

[WangK2019] K. Wang, et al., Simulating dynamic quantum phase transitions in photonic quantum walks, Phys. Rev. Lett. 122, 020501 (2019).

[WangX2018] X.-L. Wang, et al., 18-qubit entanglement with six photons' three degrees of freedom, Phys. Rev. Lett. 120, 260502 (2018).

[Wehner2018] S. Wehner, et al., Quantum internet: A vision for the road ahead, Science 362, eaam9288 (2018).

[Weimar2025] M. Weimar, et al., Fisher information flow in artificial neural networks, Phys. Rev. X 15, 031072 (2025).

[Wen2023] Z. Wen, et al., Single multimode fibre for in vivo light-field-encoded endoscopic imaging, Nat. Photonics 17, 679 (2023).

[Wen2024] S. Wen, et al., Metasurface array for single-shot spectroscopic ellipsometry, Light: Sci. Appl. 13, 88 (2024).

[Wengerowsky2018] S. Wengerowsky, et al., An entanglement-based wavelength-multiplexed quantum communication network, Nature 564, 225 (2018).

[Westerhout2020] T. Westerhout, et al., Generalization properties of neural network approximations to frustrated magnet ground states, Nat. Commun. 11, 1593 (2020).

[Wetzel2017] S. J. Wetzel, Unsupervised learning of phase transitions: from principal component analysis to autoencoders, Phys. Rev. E 96, 022140 (2017).

[Whittaker2017] R. Whittaker, et al., Absorption spectroscopy at the ultimate quantum limit from single-photon states, New J. Phys. 19, 023013 (2017).

[Wolterink2016] T. A. W. Wolterink, et al., Programmable two-photon quantum interference in 103 channels in opaque scattering media, Phys. Rev. A 93, 053817 (2016).

[Wrigge2008] G. Wrigge, et al., Efficient coupling of photons to a single molecule and the observation of its resonance fluorescence, Nat. Phys. 4, 60 (2008).

[Xiao2017] L. Xiao, et al., Observation of topological edge states in parity–time-symmetric quantum walks, Nat. Phys. 13, 1117 (2017).

[Xiao2020] L. Xiao, et al., Non-Hermitian bulk–boundary correspondence in quantum dynamics, Nat. Phys. 16, 761 (2020).

[Xie2022] M. Xie, et al., Ultrasensitive detection of local acoustic vibrations at room temperature by plasmon-enhanced single-molecule fluorescence. Nat. Commun. 13, 3330 (2022).

[Xu2015] B. Xu, et al., Empirical evaluation of rectified activations in convolutional network, arXiv:1505.00853 (2015).

[Xu2019] J. Xu, et al., Automated stoichiometry analysis of a single-molecule fluorescence imaging traces via deep learning, J. Am. Chem. Soc. 141, 6976–6985 (2019).

[Xu2020] F. Xu, et al., Three-dimensional nanoscopy of whole cells and tissues with in situ point spread function retrieval, Nat. Methods 17, 531 (2020).

[Xu2021] J. Xu, et al., Magnetic sensitivity of cryptochrome 4 from a migratory songbird, Nature 594, 535 (2021).

[Yamauchi2005] M. Yamauchi, Jones-matrix models for twisted nematic liquid crystal devices, Appl. Opt. 44, 4484 (2005).

[Yao2023] C. Yao, et al., Integrated reconstructive spectrometer with programmable photonic circuits, Nat. Commun. 14, 6376 (2023).

[Yariv1984] A. Yariv and A. P. Yeh, Optical Waves in Crystals, Wiley (1984).

[Yin2017] J. Yin, et al., Satellite-based entanglement distribution over 1200 kilometers, Science 356, 1140 (2017).

[Yin2020] J. Yin, et al., Entanglement-based secure quantum cryptography over 1,120 kilometres, Nature 582, 501 (2020).

[You2020] C. You, et al., Identification of light sources using machine learning, Appl. Phys. Rev. 7, 021404 (2020).

[Yu2020] N. Yu, Multipartite entanglement certification, with or without tomography, IEEE Trans. Inf. Theory 66, 6369 (2020).

[Yu2024] H. Yu, et al., Time-encoded photonic quantum states: Generation, processing, and applications, Appl. Phys. Rev. 11, 041318 (2024).

[Yuan2023] S. Yuan, et al., Geometric deep optical sensing, Science 379, eade1220 (2023).

[Zadeh2021] I. E. Zadeh, et al., Superconducting nanowire single-photon detectors: A perspective on evolution, state-of-the-art, future developments, and applications, Appl. Phys. Lett. 118, 190502 (2021).

[Zambra2005] G. Zambra, et al., Experimental reconstruction of photon statistics without photon counting, Phys. Rev. Lett. 95, 063602 (2005).

[Zappa2007] F. Zappa, et al., Principles and features of single-photon avalanche diode arrays, Sens. Actuators A: Phys. 140, 103 (2007).

[Zavatta2004] A. Zavatta, et al., Quantum-to-classical transition with single-photon-added coherent states of light, Science 306, 660 (2004).

[Zbontar2021] J. Zbontar, et al., Barlow twins: Self-supervised learning via redundancy reduction, Proc. ICML 139, 12310 (2021).

[Zhai2013] Y. Zhai, et al., Photon-number-resolved detection of photon-subtracted thermal light, Opt. Lett. 38, 2171 (2013).

[Zhan2021] Z. Zhan, et al., Optical polarization–based seismic and water wave sensing on transoceanic cables, Science 371, 931 (2021).

[Zhang2012] L. Zhang, et al., Recursive quantum detector tomography, New J. Phy, 115005 (2012).

[Zhang2021] Y. Zhang, et al., Single-photon detection in quantum dot photodiodes, Adv. Opt. Mater. 9, 2100592 (2021).

[Zhang2024] K. Zhang, et al., An Earth-mass planet and a brown dwarf in orbit around a white dwarf, Nat. Astron. 8, 1575 (2024).

[ZhangY2025] Y. Zhang, et al., Subcutaneous depth-selective spectral imaging with mμSORS enables noninvasive glucose monitoring, Nat. Metab. 7, 42 (2025).

[ZhangZ2025] Z. Zhang, et al., Structured light meets machine intelligence, eLight 5, 26 (2025).

[Zhong2020] H.-S. Zhong, et al., Quantum computational advantage using photons, Science 370, 1460 (2020).

[Zhu2018] D. Zhu, et al., A scalable multi-photon coincidence detector based on superconducting nanowires, Nat. Nanotechnol. 13, 596 (2018).

[Zhu2022] L. Zhu, et al., Crystal Twins: self-supervised learning for crystalline material property prediction, npj Comput. Mater. 8, 131 (2022).

[Zhuang1999] Z. Zhuang, et al., Polarization controller using nematic liquid crystals, Opt. Lett. 24, 694 (1999).

[Zimmermann2023] J. Zimmermann, et al., Finding the semantic similarity in single-particle diffraction images using self-supervised contrastive projection learning, npj Comput. Mater. 9, 24 (2023).

Appendix to the habilitation thesis / Příloha habilitační práce

From light to knowledge: Sensing with photonic and neural networks

RNDr. Miroslav Ježek, Ph.D.

The habilitation thesis is based on the following scientific works of the applicant. Copies of the underlying publications and preprints are attached.

Habilitační práce je založena na následujících vědeckých pracích uchazeče. Příslušné publikace a preprinty jsou přiloženy.

J. Hloušek, I. Straka, M. Ježek, Experimental observation of anomalous supralinear response of single-photon detectors, Appl. Phys. Rev. 10, 011412 (2023).

J. Hloušek, M. Dudka, I. Straka, M. Ježek, Accurate detection of arbitrary photon statistics, Phys. Rev. Lett. 123, 153604 (2019).

J. Hloušek, J. Grygar, M. Dudka, M. Ježek, High-resolution coincidence counting system for large-scale photonics applications, Phys. Rev. Applied 21, 024023 (2024).

M. Bielak, R. Stárek, V. Krčmarský, M. Mičuda, M. Ježek, Accurate polarization preparation and measurement using twisted nematic liquid crystals, Opt. Express 29, 33037 (2021).

D. Vašinka, M. Bielak, M. Neset, M. Ježek, Bidirectional deep learning of polarization transfer in liquid crystals with application to quantum state preparation, Phys. Rev. Applied 17, 054042 (2022).

M. Bielak, D. Vašinka, M. Ježek, All-fiber microsensor of polarization at single-photon level aided by deep-learning, Laser Photonics Rev. e01775 (2025).

D. Koutný, L. Ginés, M. Moczała-Dusanowska, S. Höfling, Ch. Schneider, A. Predojević, M. Ježek, Deep learning of quantum entanglement from incomplete measurements, Sci. Adv. 9, eadd7131 (2023).

D. Vašinka, F. Juráň, J. Běhal, M. Ježek, From stars to molecules: AI guided device-agnostic super-resolution imaging, arXiv:2502.18637, bioRxiv:2025.02.25.640182 (2025).

Applied Physics Reviews

Experimental observation of anomalous supralinear response of single-photon detectors

Cite as: Appl. Phys. Rev. **10**, 011412 (2023); https://doi.org/10.1063/5.0106987
Submitted: 01 July 2022 • Accepted: 22 February 2023 • Published Online: 21 March 2023

Josef Hloušek, Ivo Straka and Miroslav Ježek

Applied Physics Reviews ARTICLE scitation.org/journal/are

Experimental observation of anomalous supralinear response of single-photon detectors

Cite as: Appl. Phys. Rev. **10**, 011412 (2023); doi: 10.1063/5.0106987
Submitted: 1 July 2022 · Accepted: 22 February 2023 ·
Published Online: 21 March 2023

Josef Hloušek,[a] Ivo Straka, and Miroslav Ježek[b]

AFFILIATIONS

Department of Optics, Faculty of Science, Palacký University, 17. listopadu 12, 77146 Olomouc, Czechia

[a]Author to whom correspondence should be addressed: hlousek@optics.upol.cz
[b]Electronic address: jezek@optics.upol.cz

ABSTRACT

The linearity of single-photon detectors allows accurate optical measurements at low light levels and using non-classical light in spectroscopy, biomedical imaging, optical communication, and sensing. However, in practice, the response of single-photon detectors can exhibit intriguing nonlinear effects that may influence the performed measurements. Here, we demonstrate a direct single-source measurement of the absolute nonlinearity of single-photon detectors with unprecedented accuracy. We discover a surprising supralinear behavior of single-photon avalanche diodes and show that it cannot be explained using known theoretical models. We also fully characterize sub- and supra-linear operation regimes of superconducting nanowire single-photon detectors and uncover the supralinearity under faint continuous illumination. The results identify new detector anomalies that supersede existing knowledge of nonlinear effects at the single-photon level.

I. INTRODUCTION

A majority of radiometric, spectroscopic, imaging, and optical communication methods rely on comparing two or more levels of light intensity measured by a photodetector and assuming that its response is proportional to the incident radiation. Transmittance measurement represents the simplest example where the optical power is detected with and without the sample under test, see Fig. 1(a). For an ideal detector, the power ratio would be the same as the actual transmittance of the sample. The measurement accuracy is, however, impaired by any deviation from a perfectly linear response of the detector. One can correct the detection imperfections on the condition that the model of the nonlinear effect is known and accurate enough.

With the advent of ultra-sensitive detectors and quantum-enhanced metrology, we tend to perform measurements at the ultimate sensitivity levels dictated by the laws of physics.[1,2] The goal is to reach the quantum advantage regime—that is, to improve the sensitivity of measurement beyond the shot-noise limit, or to relax the requirements of the measurement, such as the minimum required detection efficiency. Shaping the statistics of light and using nonclassical optical signals as measurement probes allow for increasing the precision of length measurements,[3,4] imaging and particle tracking,[5–9] and spectrophotometry.[10,11] The optical transmittance measurement assisted by correlated photons and single-photon detectors (SPDs) can serve as a prominent example of a quantum-enhanced measurement scheme.[12–15]

The measurement precision at the single-photon level is severely affected by the nonlinearity of the employed photonic detectors. The reason is that the other systematic errors need to be eliminated to reach the quantum regime, while the SPDs themselves maintain strong inherent nonlinearity. One can observe a complex interplay of detector-specific phenomena, such as dark counts, dead time, recovery transition, multi-photon response, and latching. These effects cause highly nontrivial nonlinear behavior that is much stronger compared to classical photodiodes—see Fig. 1(b). Not only does the nonlinear response distort the measurement of the average photon flux, but it also distorts the measured photon statistics and prevents us from reaching the ultimate precision of quantum-enhanced measurements. Notable experiments susceptible to detection nonlinearity are the tests of Born's rule in quantum mechanics.[16–21]

Here, we explore the nonlinearity of actively and passively quenched single-photon avalanche diodes (SPADs) and a superconducting nanowire single-photon detector (SNSPD) for various bias currents. The employed nonlinearity measurement does not require a calibrated reference or time-resolved detection, which considerably simplifies the task. The nonlinearity characterization is performed with unprecedented accuracy; we reliably detect nonlinearities smaller than 1:1000 and cover seven orders of magnitude of incident illumination. We discover a supralinear region of SPAD operation, which is not consistent with any known theoretical models and has not been

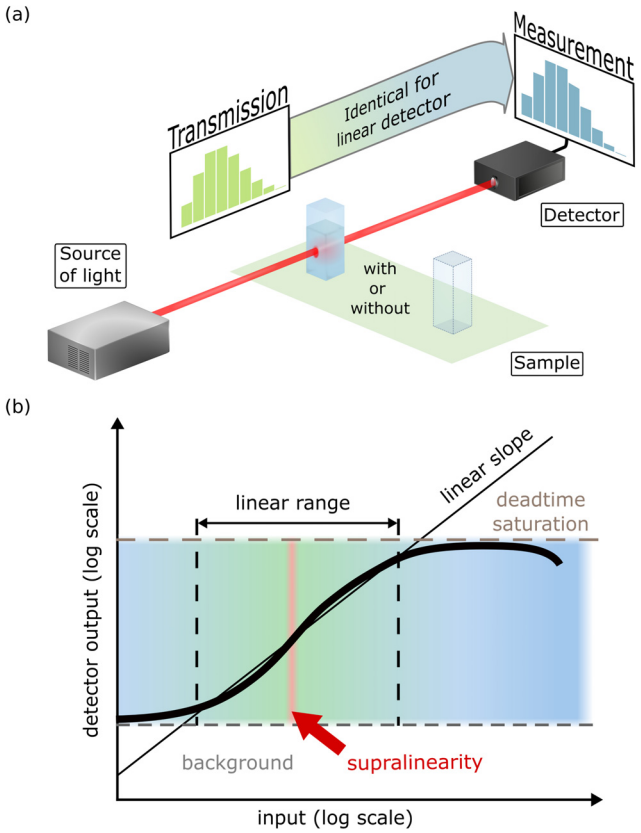


FIG. 1. (a) Optical transmittance measurement of a sample relies on comparing two levels of intensity captured by a detector. Any deviation from the ideal linear detection then critically affects the measurement accuracy. Nonlinearity is a particularly vexing problem in quantum metrology when nonclassical statistics of light is often required together with SPDs. Their inherently nonlinear response distorts the photon statistics and makes it difficult to reach the quantum advantage. (b) Nonlinear behavior of a SPD. Apart from the background detections and saturation effects, SPDs may exhibit supralinearity.

reported yet. We also characterize the nonlinear behavior of an SNSPD with a complex structure of sub- and supra-linear operation regions. We observed, for the first time, the supralinearity of an SNSPD under continuous illumination at very low detection rates.

A. Single-photon detectors

Before discussing the main results, let us briefly review the fundamentals of SPDs and the basic principles of nonlinearity characterization.

SPDs convert single-photon absorptions into macroscopic voltages or currents. The current state-of-the-art technology and techniques that are available to achieve single-photon detection are built on various device structures, materials, and non-trivial physical phenomena. The ones enjoying the most widespread practical use are based on avalanche diodes and superconducting circuits.[22–26] Among these, we investigate detectors that operate in a counting regime (Geiger mode); that is SPADs and SNSPDs. There are also emerging SPDs based on low-dimensional materials.[27,28] Such detectors do not yet have established detection models that could consistently model their nonlinearity, which underlines the need to characterize their response by the direct measurement.

SPD in a counting regime outputs an electronic pulse when one or more photons are detected.[22–26] The detection events (counts) arrive at random times with a statistical distribution given by the detected state and the response of the detector. The detection rate R^{det} is then a function of the incident rate R; the rate is given in units of counts/s, or simply cps. Sometimes the detector outputs a pulse even when no photon is detected due to various background contributions (dark counts) or as a result of a previous detection (SPAD afterpulses). Furthermore, the detector occasionally fails to detect photons because it is not ready to do so after the previous detection event, such as during dead time or a latched state. Figure 1(b) illustrates the nonlinear behavior of an SPD. A simple expectation is a $\int$-shaped sub-linear dependence, where background count rate and dead time saturation dominate on opposite sides of the power range. However, we found that some detectors exhibit an S-shaped response where the slope gets supralinear in the middle. Supralinearity has been reported for silicon photodiodes for strong classical illumination,[29–31] but current SPAD models and measurements do not predict any such phenomenon. For SNSPDs, supralinear behavior has been observed for very high rates ($\geq$ 10 Mcps) due to AC detector coupling[32] and for short optical pulses (mean photon number $\geq$ 0.1) as a result of two-photon absorption.[33] However, no observation of supralinearity of SNSPDs has been reported under continuous illumination with the detection rate below 1 Mcps.

The particular detection imperfections are specific to SPADs and their quenching circuits,[23,24] or SNSPDs[22,25,26] due to their different operational principles. There is a great number of results in modeling the response of a SPAD with the goal of including all the relevant factors.[34–38] Their accuracy has been limited so far and many counterexamples exist for which the measured SPAD response differs from the theoretical model. Consequently, determining the nonlinearity of the SPAD response and finding the optimum detection rate to access the minimum achievable deviation from the ideal linear behavior represents a significant challenge. This issue is even more pronounced for SNSPDs due to the lack of a precise theoretical model taking into account all physical processes.[33] A semi-empirical model was proposed and tested with accuracy 10^{-2}.[39,40] Detector tomography based on probing with precisely calibrated signals was suggested to thoroughly characterize a detector and obtain the corresponding positive-operator-valued measure.[40–48] However, if the tomography does not include memory effects, the results can be compromised.[49] The approach presented in the rest of the paper does not rely on a theoretical model or detector tomography. Instead, we focus on a direct measurement of the detector nonlinearity as a function of the detection rate.

B. Nonlinearity characterization

The nonlinearity of various photodetectors, mainly photodiodes, has been explored in great depth using relative and absolute measurement methods.[50] Relative methods require a calibrated reference detector, calibrated attenuators, or time-resolved probe signals and detection. Absolute measurements, which are generally preferred, are based on a superposition method where the response of the detector is evaluated separately for two signals and their total.[29,51–56] The individual signals have to be incoherent to prevent optical interference. Often

two independent optical sources are used, for this reason,[56] preferably exhibiting short coherence lengths.[55] Optionally, the two signals are derived from the same source and superimposed in a slightly misaligned interferometer.[53,54] The direct absolute measurement of nonlinearity of a SPAD under continuous illumination was reported by Kauten *et al.*, and no statistically significant deviation from the standard model [Eq. (2)] was found.[56] We use a single-source approach to directly characterize major single-photon technologies, namely, passively and actively quenched SPADs, and SNSPDs for various values of bias current. Our measurements reveal deviations from standard models, as well as anomalous supralinearity.

II. DIRECT NONLINEARITY MEASUREMENT

We employ the absolute measurement strategy based on the single-source two-beam superposition method. For a constant intensity level of the optical source, we perform a series of three measurements of the detector response, see Fig. 2(a). The beam is split into two paths A and B that can be individually blocked. The detection rates R_A^{det} and R_B^{det} are recorded for each path, and R_{AB}^{det} is acquired with both paths open. Nonlinearity is defined as a deviation from the ideal linear response,

$$\Delta = \frac{R_A^{det} + R_B^{det}}{R_{AB}^{det}} - 1. \quad (1)$$

Our study considers two types of nonlinearity: sub- ($\Delta > 0$) and supralinearity ($\Delta < 0$). The splitting ratio R_A^{det}: R_B^{det} is chosen to be 50:50. For the ultimate nonlinearity characterization, all splitting ratios would need to be measured, but without calibration, 50:50 splitting is the only one that can be set with certainty, as $R_A^{det} = R_B^{det}$. Furthermore, it harnesses the method's invariance—a 10% deviation from the 50% splitting leads to a $\leq 4\%$ relative error in Δ. This is in contrast to using a 3 dB attenuator, where a 1% error in its transmission can change Δ by an order of magnitude (see the discussion in Appendix D).

Our approach to the nonlinearity measurement does not require an absolute calibration of a light source. The measured nonlinearity (1) is a function of the detected count rate (in counts per second). This is well-justified, as nonlinearity affects relative measurements of photon flux, where two detection rates are compared, and the corresponding flux ratio needs to be accurately inferred. In order to correct for the measured value of nonlinearity, we only need to know at which detection rate it occurs.

The experimental setup is shown in Fig. 2(b). A stabilized superluminescence diode is the source, with a central wavelength of 810 nm and the spectral width exceeding 20 nm. The signal is attenuated to scan the input power over the full dynamic range of the detector. To avoid phase interference, the beam is split and joined using polarization beam splitters (PBS), with a 35 mm path difference that is several orders of magnitude larger than the coherence length. The output of the Mach–Zehnder is coupled into a single-mode optical fiber connected to the active area of the SPD under test. The fiber coupling guarantees the same spatial profile of the signals at the detector, but it can decrease the overall stability. Therefore, extra effort was made to reach 10^{-5} power stability of the setup including the coupling stages (see Appendix B).

In the case of SNSPD measurements, the source spectrum was filtered to 12 nm around 800 nm, for which the detector is optimized. The reduced spectral width corresponds to a coherence length shorter than 30 μm. The coupling fiber was equipped with a diagonal polarizer and a polarization controller to set the optimal polarization for the maximization of the SNSPD efficiency. In this case, the incoherence relied solely on the Mach–Zehnder path difference.

The acquisition time of each individual rate measurement was set to 20 s. The complete characterization of a single detector typically took 20 hours with each measurement repeated 30 times for 40 different rate levels. The precision was found to be limited mostly by fundamental Poissonian variance (see Appendix C).

We measured three actively quenched SPADs (SPAD-1–3), a passively quenched SPAD (SPAD-P), and an SNSPD for various values of bias current. The output of the detector under test is processed by an FPGA-based counter with a 2.2 ns pulse-pair resolution that is well below the dead time of any detector. For technical information, see Appendix A.

III. RESULTS AND DISCUSSION

A. Single-photon avalanche diodes

Figure 3 shows the measured nonlinearity as a function of the detection rate R_{AB}^{det} for various SPADs. The typical log-log plot of the SPAD nonlinearity Δ is V-shaped due to dark counts (left slope) and

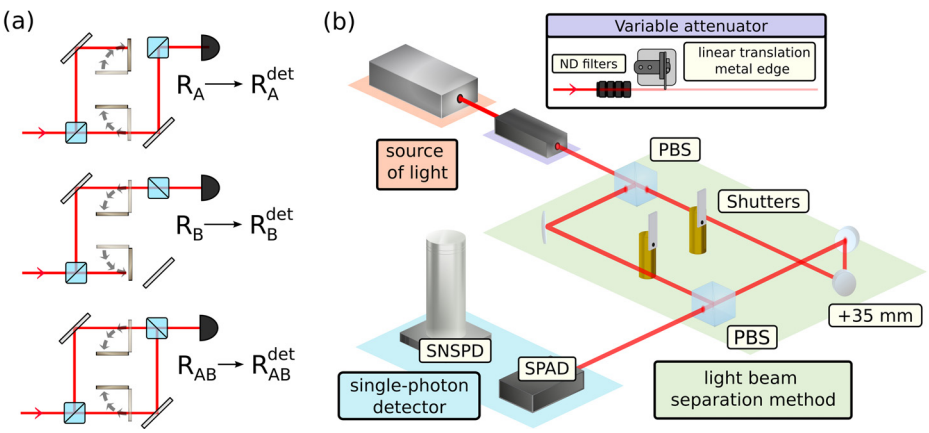


FIG. 2. The single-source two-beam method for the absolute measurement of nonlinearity consisting of three measurements of the detection rate (a) and the simplified experimental scheme of the method (b).

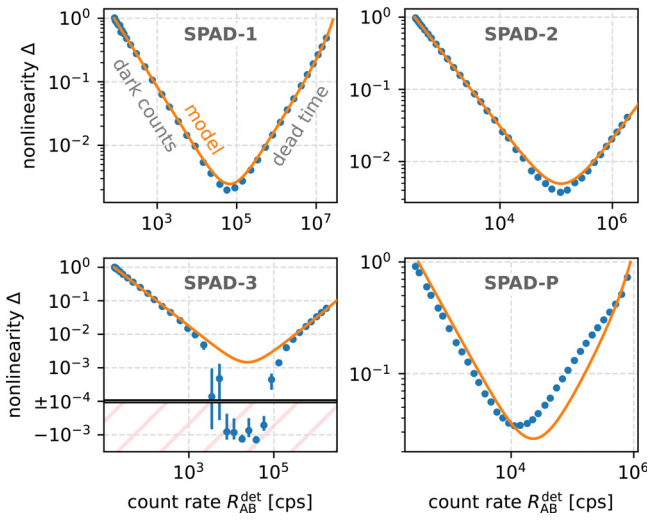


FIG. 3. Nonlinear response of the tested SPADs. The solid line represents theoretical model (3). Each point was measured 30 times, and the error bars show the corresponding standard error of the mean.

dead time saturation (right slope). The nonlinearity reaches its minimum value (10^{-3}–10^{-2}) for detection rates between 10^4 and 10^5 cps.

The established model of the detection rate R^{det} takes into account the dark count rate R_0 and the non-paralyzable dead time τ,[38,56] reading

$$R^{det} = f(R) = \frac{R + R_0}{1 + (R + R_0)\tau}, \tag{2}$$

where R is the incident rate. The nonlinearity is measured in a balanced configuration so that the incident rates are $R_A = R_B = R_{AB}/2$. If we substitute the detector model $R^{det} = f(R)$ given in Eq. (2), we obtain the detection rates $R_A^{det} = R_B^{det} = f(f^{-1}(R_{AB}^{det})/2)$. The expected nonlinearity as a function of the detection rate is

$$\Delta\left(R_{AB}^{det}\right) = \frac{2f\left(\frac{1}{2}f^{-1}\left(R_{AB}^{det}\right)\right)}{R_{AB}^{det}} - 1. \tag{3}$$

The parameters of this model are the dark count rate R_0 and dead time τ. More elaborate models of actively quenched SPADs are discussed in Appendix F, although none provide a better fit than (2).

The data were fit with Eq. (3) in Fig. 3, and they are in significant disagreement with the model, affirming the need for direct nonlinearity measurement. The most prominent feature is the supralinear behavior ($\Delta < 0$) of SPAD-3 in Fig. 3, which is a hitherto unreported phenomenon for SPADs. In fact, all actively quenched SPADs 1-3 exhibit systematically lower nonlinearities than expected in their minima, and supralinearity appears after correcting for R_0 and τ (see Fig. 11 in Appendix F). Additionally, the fit parameters R_0, τ do not agree with values that were obtained from independent time-resolved measurements.

A comparison of dead times and directly observed recovery times is given in Table I. For actively quenched detectors (SPAD-1–3), the measured recovery time τ_R also includes a brief detector reset time, so must always be $\tau_R \geq \tau$.[57] Reset effects like twilight pulsing could, therefore, explain the difference for SPAD-2 but not for the other detectors SPAD-1 and SPAD-3.[38] SPAD-P is passively quenched and so the recovery and dead time values differ significantly. Such detectors exhibit gradual efficiency recovery after each detection, which would require a complex empirical model. As shown in Appendix F, no other few-parameter models offer a better fit than (2).

TABLE I. Comparison of directly measured dark counts R_0 and recovery times τ_R that were measured directly using time-resolved detection techniques[38] and the dead time values τ_{fit} and dark counts $R_{0,fit}$ that were the best fit of model (2).

	SPAD-1	SPAD-2	SPAD-3	SPAD-P
R_0 (cps)	88(3)	314(6)	20(2)	264(1)
$R_{0,fit}$ (cps)	83(4)	304(2)	17.5(7)	300(200)
τ_R (ns)	29.5(5)	47.0(5)	56.6(4)	517(6)
τ_{fit} (ns)	36.7(1)	40.2(4)	61(1)	1130(20)

Because not all systematic errors can be seen in Fig. 3, we evaluated the χ^2 with the highest p-value being $p_{SPAD-2} \sim 10^{-70}$. It means that under the assumption that model (2) is valid and only the statistical error takes effect, the probability of obtaining the measured data or any worse data would be equal to the p-value. This means that model (2) already deviates from the measurement with high statistical significance for 20 s integration times. Additionally, we applied standard rate corrections on the data [inversion of (2)] to evaluate any residual nonlinearity, which reveals a significant mismatch even for $R_{AB}^{det} \gtrsim 10^6$ cps. A detailed statistical analysis can be found in Appendix F.

We have also tested more elaborate response models such as Ref. 38 (including dead time, afterpulses, and twilight pulses), and a combination of paralyzable and non-paralyzable dead time[58–63] developed originally for Geiger–Müller detectors. None of these can reproduce the measured results or explain supralinearity, which is shown in Appendix F. We also ruled out potential changes in dead time, dark counts, efficiency, or afterpulsing as hypothetical explanations for the observed supralinearity (see Appendix F).

Current evidence shows that SPADs do not always exhibit constant dark counts,[64] efficiencies,[65] or recovery time (dead time + reset time).[38] This means that state-of-the-art models are insufficient, leaving the experimentalist with two options: (1) a detailed characterization of the SPAD involving time-resolved measurements and circuit analysis; and (2) direct nonlinearity measurement.

B. Superconducting nanowire single-photon detectors

Figure 4 shows the nonlinearity of the SNSPD with respect to the bias current. Inset 4(a) shows the dependence of dead time and dark counts on the bias. The property affected the most is detection efficiency, shown in inset 4(b). The detection efficiencies were calculated relatively to the manufacturer's specification of 86±3% for 25 μA. Inset 4(c) shows the measured nonlinearity Δ as a function of the total detected rate R_{AB}^{det}. Each data set corresponds to a different bias current, where the plot markers match the respective points in inset 4(b).

The nonlinearity of the SNSPD is a combination of several phenomena. For lower rates, the effect of dark counts is easily recognizable. Around 10^4 cps, all sub-critical regimes ($I_{bias} \leq 25$ μA) begin to exhibit supralinearity ($\Delta < 0$). For higher count rates, two scenarios are observed. Bias currents corresponding to the efficiency plateau

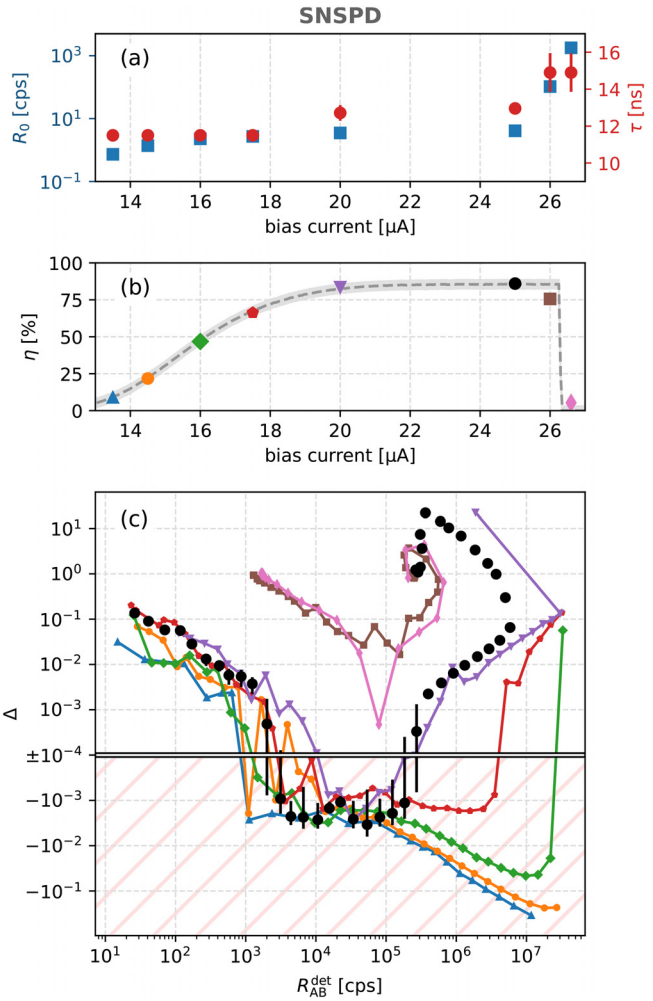


FIG. 4. (a) Dark counts R_0, dead time τ, and (b) detection efficiency η as a function of bias current. The dashed line is the manufacturer's specification. (c) The measured nonlinearity of the SNSPD for the values of bias current given by the respective markers in (b).

(20–25 μA) lead to dead time saturation, while lower biases maintain the supralinearity. Finally, for $R_{AB}^{det} > 10^6$ cps, latching rapidly increases for all but the lowest biases. This introduces strong saturation and eventually results in an inverse detection response, where the count rate decreases with increasing illumination.

Let us address the supralinear behavior, linked previously to two effects. With decreasing bias current, two-photon and higher-order detection efficiencies rise as indicated by detector tomography.[40] So far, the two-photon absorption of SNSPDs was observed for ultra-short optical pulses in a picosecond regime with mean photon numbers ≥ 0.1.[33] Under continuous illumination, the multi-photon absorption becomes challenging to model due to hotspot relaxation dynamics.[33] The second factor is the AC coupling of the readout circuit—the settling of the bias current after each detection results in rate-dependent efficiency. The supralinear behavior due to the AC coupling was observed only for very high rates ≥ 10 Mcps.[32]

Our results represent the first observation of the SNSPDs supralinearity under continuous illumination with the detection rates as low as 10^3 cps.

IV. CONCLUSION AND OUTLOOK

We designed and realized a direct measurement of single-photon detection absolute nonlinearity with high accuracy over several orders of magnitude of light intensity. We performed the measurement for SPADs and SNSPDs. The measurement technique does not require a calibrated single-photon source, a calibrated reference detector, or time-resolved detection. Due to the method's robustness, the nonlinearity is not affected by technical issues, but rather reflects the detector properties.

For all SPADs, we found significant disagreement of established theoretical models with the measured data. We also discovered anomalous supralinear behavior of a SPAD. This phenomenon has been neither predicted nor observed yet and cannot be explained in terms of the known SPAD parameters. The SPADs in question were based on silicon, but SPADs based on III–V materials (InGaAs/InP) possess the same principle of operation and, therefore, exhibit similar nonlinear phenomena as silicon SPADs, albeit of different magnitudes. We expect that sublinear behavior would be generally stronger for the InGaAs/InP SPADs than silicon SPADs due to higher values of dark counts and dead time. Consequently, the chance of observing subtle deviations, such as supralinearity, is expected to be lower.

For the SNSPD, we performed its detailed nonlinearity analysis over 7 orders of magnitude of incident illumination and the full range of relevant values of bias current, which goes beyond any SNSPD characterization reported so far. We detected SNSPD supralinearity under continuous illumination at unprecedentedly low detection rates down to 10^3 cps, which has not been observed before. The results for both SPADs and SNSPDs show that nonlinearity in single-photon detection is a complex mixture of non-trivial phenomena, which have so far eluded accurate theoretical description.

The direct SPD characterization can also be applied to emerging detection techniques, such as those based on low-dimensional materials. The results in this field suggest susceptibility to nonlinear response as well.

Some of these detectors operate in a linear gain regime, such as p-n junction arrays,[66] 2D-layer Schottky barriers,[67] or quantum dots.[68,69] The results in Refs. 66–68 show that the overall nonlinearity can be quite significant and dependent on the bias voltage. Until there are direct nonlinearity characterizations performed for such detectors, it is hard to speculate whether both sub- and supra-linearity could be expected, but the published results suggest that both can be present.

Low-dimensional detectors based on photogating operate in a counting regime.[70,71] That means individual single-photon absorptions are distinguished as quantized jumps in photocurrent. Due to the long-lived traps responsible for the photogating effect, the trapped carriers can accumulate even at low rates below 1 cps. Additionally, trap decays cause reverse jumps in photocurrent that can, in principle, interfere with detection readouts.[70] Electronic noise contributes to crosstalk in photon-number distinction. All these effects may contribute to nonlinearity, and the results in Ref. 70 suggest both strong sub- and supra-linearity.

A Josephson-junction-based detector reported in Ref. 72 also operates in a counting regime. The results therein only show

nonlinearity due to dark counts, but a dedicated measurement may reveal additional effects.

There are other photosensitive nanostructures[73] that have not yet been sufficiently characterized in terms of their response to incident power. With all such emerging single-photon detection techniques, the ability to perform direct self-contained characterizations is valuable due to the lack of accurate theoretical or empirical models.

The findings presented in this work can be applied to radiometric, spectroscopic, imaging, and optical communication methods that rely on precise assessment of illumination or transmission levels. We have shown that correcting for nonlinear aspects based on state-of-the-art response models is generally not sufficient, so the direct measurement of nonlinearity becomes a necessity. Accurate detector calibration is particularly critical for measuring photon statistics and reaching the quantum advantage in quantum metrology. The presented results open the way for ultra-precise identification and mitigation of nonlinear effects in SPDs.

ACKNOWLEDGMENTS

We acknowledge the funding from the Czech Science Foundation (Project No. 21-18545S), MEYS, and European Union's Horizon 2020 (2014–2020) research and innovation framework programme (Project HYPER-U-P-S, Grant No. 8C18002). Project HYPER-U-P-S has received funding from the QuantERA ERA-NET cofund in Quantum Technologies implemented within the European Union's Horizon 2020 Programme. J.H. acknowledges the Palacký University under Project Nos. IGA-PrF-2021-006 and IGA-PrF-2022-005.

AUTHOR DECLARATIONS

Conflict of Interest

Yes, The custom-made counter used in the presented measurements is commercially available and sold by the Palacký University. Miroslav Ježek is one of the developers and thus eligible for a percentage of the income. However, the presented measurements are not affected by the choice of the counter, as the necessary parameters can be met by other devices available on the market. All other authors declare they have no competing interest.

Author Contributions

Josef Hloušek: Data curation (equal); Formal analysis (equal); Software (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **Ivo Straka:** Formal analysis (equal); Methodology (equal); Software (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **Miroslav Jezek:** Conceptualization (equal); Supervision (equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are openly available in GitHub at https://github.com/PepaHlousek/Nonlinearity, Ref. 74.

APPENDIX A: EXPERIMENTAL SETUP

Figure 5 depicts the experimental setup for the absolute measurement of nonlinearity using the single-source two-beam superposition method. The light source is a temperature-stabilized superluminescence diode (QPhotonics QSDM-810-2) in the constant current mode with the central wavelength of 810 nm and the spectral width exceeding 20 nm. Angled physical contacts (APC) between optical fibers reduce back-reflections approximately by 60 dB and improve the stability of the source. Attenuation by several orders of magnitude is needed to generate an optical signal in the dynamic range of the tested detector. For this purpose, the optical beam is strongly attenuated by a series of neutral density (ND) filters (Thorlabs NE530, NE520, NE513, and NE504) down to $\sim 10^7$ counts per second. Continuous adjustment is realized by a metal edge fixed to a linear translation motorized stage (LTME, Newport MFA-CC) driven by a motion controller (SMC-100CC). The moving edge along with single-mode fiber coupling serves as variable attenuation.

Since the splitting ratio of the interferometer is sensitive to input polarization, the input beam is well polarized and coupled in the slow axis of a polarization-maintaining fiber. The unbalanced Mach–Zehnder interferometer consists of two polarizing beam splitters (PBS) with extinction ratio $T_P : T_S > 1000 : 1$. The Mach–Zehnder setup is used to split the attenuated beam into two spatially separated beams that can be individually blocked and superimposed again at the output. The combination of perpendicular polarizations and the path difference $l = 35.4$ mm between the individual arms ensure incoherent mixing at the second polarization beam splitter. Mechanical blocking of the two optical signals is done by compact home-made optical shutters using a digital RC

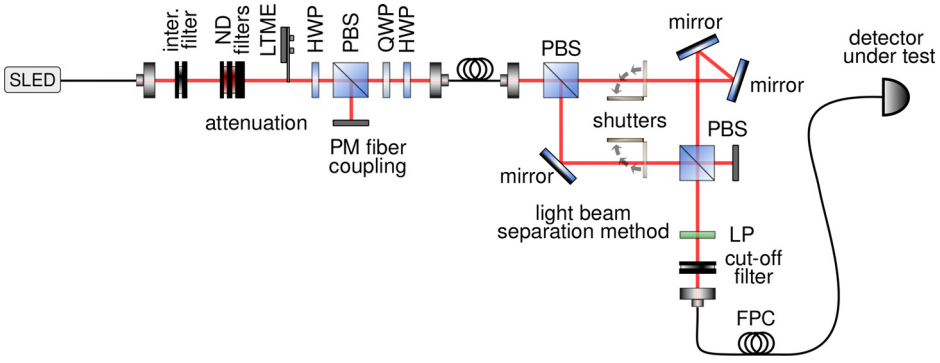


FIG. 5. Experimental setup of the single-source two-beam superposition method for the absolute nonlinearity measurement: the scheme includes the preparation of a stabilized and attenuated optical signal, its polarization control, beam separation and switching, and detector under test.

TABLE II. Manufacturer specifications for tested detectors. Parameters shown: active area diameter A, photon detection efficiency η, dark count rate R_0, dead time τ, and afterpulsing probability p_a.

	A (μm)	η (%)	R_0 (cps)	τ (ns)	p_a (%)
SPAD-1	180	50	<100	28.7	0.1
SPAD-2	180	47	<500	50	0.3
SPAD-3	100	57	<20	55	0.31
SPAD-P	500	80	200	1000	negligible
SNSPD	–	86	≤10	≤10	none

servo (Savox SH-0262MG). A thin (0.5 mm) metal flag with dimensions of 15×70 mm attached to the servo shaft blocks the collimated Gaussian beam that has a radius $w_0 = 1.03$ mm. To perform the operation of opening or blocking the beam, it is necessary to turn the servo by 15° with a total latency of 50 ms. The RC servo is controlled by a pulse width modulation signal generated by Arduino Uno with a microprocessor AT-mega328. For more information about the employed shutters, see Ref. 75. After the second PBS, the output light is fed into a single-mode fiber and coupled to the tested detector. Stray light is eliminated by a cutoff filter (Semrock BLP01-635R-25).

We investigated the response of five SPDs: actively (SPAD-1-3) and passively quenched (SPAD-P) thick-junction silicon single-photon avalanche diodes (SPADs) and a NbTiN superconducting (2.7 K) nanowire single-photon detector (SNSPD) with an AC readout and room-temperature amplifier. Namely, SPAD-1—Excelitas SPCM-AQRH CD3432H, SPAD-2—Perkin Elmer SPCM-AQ4C module s.n. 167, SPAD-3—Laser Components COUNT-20C-FC D4967, SPAD-P—ID Quantique ID120 s.n. 1518006, and SNSPD—Single Quantum Eos CS SNSPD system s.n. SQ071. The detectors differ in many technical parameters and physical effects, such as active area, photon detection efficiency, dark count rate, dead time, and afterpulsing probability. The manufacturer's specifications are shown in Table II. The detector SPAD-P (ID Quantique ID120) differs from SPAD-1-3 not only by the passive quenching mechanism but also by being a free-space module with tunable temperature and bias voltage. All measurements were done at the temperature −40 °C and bias voltage 180 V.

The SNSPD possesses a limited spectral region of the maximum efficiency. For this reason, the wide emission of SLED was reduced using a 12 nm bandpass interference filter (Semrock FF01-800/12–25). Furthermore, the SNSPD is naturally sensitive to polarization due to its nanowire structure. The optical beams of the unbalanced Mach–Zehnder interferometer have orthogonal H/V polarizations, so it is necessary to add a diagonally oriented linear polarizer (LP) at the output to make them indistinguishable for the detector. A fiber polarization controller (FPC) is used to optimize the detection efficiency with respect to polarization.

Electronic output signals from the tested detectors were processed by an electronic counter. The most critical parameter is its pulse-pair resolution, which has to be better than the recovery time of all tested detectors. Initially, we were using a commercial 100 MHz counter (ORTEC 974 C), which was used to measure the SPADs. Later, we developed an FPGA-based 230 MHz counter with a 2.2 ns pulse-pair resolution and digitally tunable threshold voltages (QOLO Countex[76]) This counter was used to perform measurements on the SNSPD, as well as additional measurements on SPAD-1 and SPAD-3. Detectors SPAD-1 and SPAD-3 were measured using both counters with the same results.

APPENDIX B: STABILITY OF THE SOURCE AND MEASUREMENT SETUP

Generally, intensity fluctuations of the light source can affect the measurement accuracy. We investigated the stability of the SLED and also the intensity stability at the output of the setup. We performed a long-term measurement of optical intensity and evaluated the Allan deviation[77] to determine the integration time T for which the measurement is least affected by intensity fluctuations. Figure 6 shows the relative Allan deviation σ_{Allan} as a function of the integration time T. The deviation is better than 10^{-5} for integration times from 10 s to 10^3 s. For the nonlinearity measurement setup, optimal integration times are shifted toward shorter times

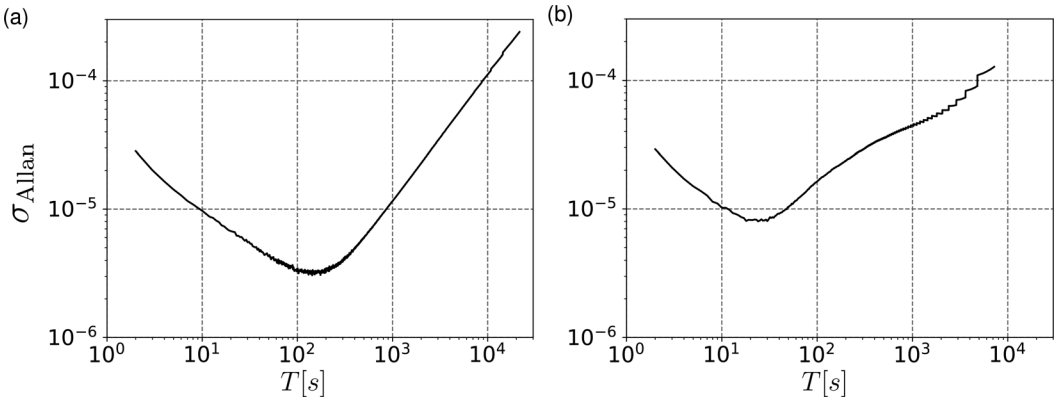


FIG. 6. Optical intensity fluctuations represented by the relative Allan deviation. Shown are (a) temperature-stabilized superluminescence diode (QPhotonics QSDM-810-2) in the constant current mode and (b) signal from the same source transmitted through the experimental setup for the nonlinearity measurement.

and the minimum relative Allan deviation gets worse, but still below 10^{-4} for integration times up to 6×10^3 s. The raw data and their processing are available on GitHub.[74]

APPENDIX C: MEASUREMENT UNCERTAINTY

1. Shot noise limitation

We performed an analysis of nonlinearity measurement uncertainty $\sigma(\Delta)$ to find its ultimate physical limits and find out whether the data approach the fundamental nonlinearity resolution. The statistical uncertainty was calculated for the case of a balanced experimental setup $R_A^{det} = R_B^{det}$, measurement time $T = 20$ s for each rate, and $N = 30$ repeated measurements.

The fundamental limit is represented by the shot noise of the Poisson counting process, where the measured number of events $(R^{det}T)$ exhibits variance equal to the mean value. This implies the standard deviation of the detection rate $\sigma(R^{det}) = \sqrt{R^{det}/T}$. If we introduce this standard deviation to each detection rate in the nonlinearity formula,

$$\Delta = \frac{R_A^{det} + R_B^{det}}{R_{AB}^{det}} - 1, \tag{C1}$$

through standard error propagation $\sigma^2(\Delta) = \sum_i [\sigma(R_i^{det}) \times \partial\Delta/\partial R_i^{det}]^2$, we obtain the standard deviation,

$$\sigma(\Delta) = \sqrt{\frac{(1+\Delta)(2+\Delta)}{R_{AB}^{det}T}} \approx \sqrt{\frac{2}{R_{AB}^{det}T}}. \tag{C2}$$

The measured nonlinearity $\bar{\Delta}$ is averaged from N measurements, so the standard deviation of the result is

$$\sigma(\bar{\Delta}) = \sigma(\Delta)/\sqrt{N}. \tag{C3}$$

This limit is compared to the experimental values in Fig. 7. The measured standard errors of the nonlinearity for SPADs meet their lower bound 10^{-4} for rates higher than 2×10^5 cps [Figs. 7(a)–7(d)]. For higher rates, excess noise limits the measurement uncertainty; it, nevertheless, remains below $\sigma(\bar{\Delta}) < 10^{-3}$. In this region, the mean nonlinearity value is larger than the uncertainty by at least one order of magnitude. Insets 7(e)–7(l) show the nonlinearity measurement uncertainty of the SNSPD for several different values of bias current. Latching rapidly increases for bias current $I_{bias} > 25$ μA and brings extra uncertainty to the measurement [Figs. 7(k) and 7(l)].

The shot noise limit is well applicable to Poisson count rates. However, dead time saturation will further affect the measurement, because the measured count rates become sub-Poissonian as they approach the limit $R^{det} \to 1/\tau_{NP}$. Because the standard deviation of accumulated counts is complicated to calculate analytically,[38] we are going to assume a sufficiently long integration time so that the

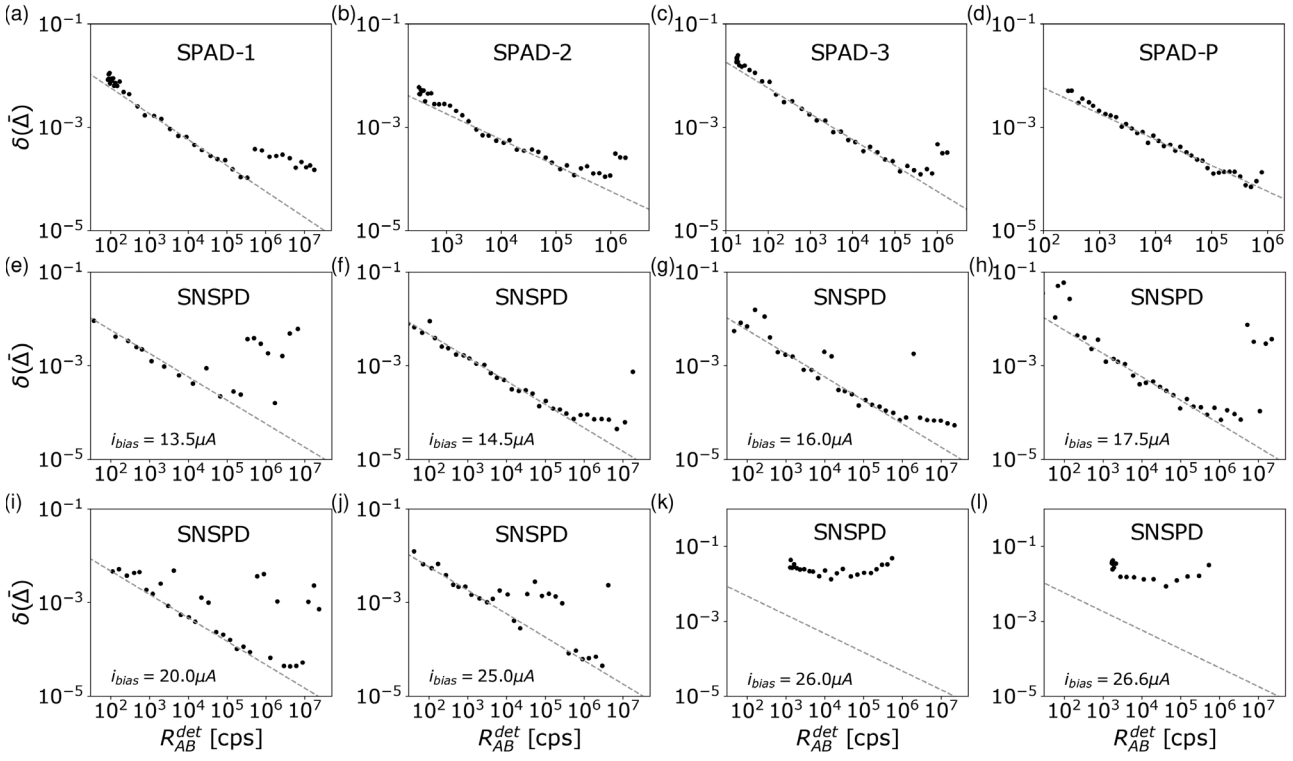


FIG. 7. The nonlinearity measurement standard error $\sigma(\bar{\Delta})$ of tested detectors (a) SPAD-1, (b) SPAD-2, (c) SPAD-3, (d) SPAD-P, and (e) and (f) SNSPD as a function of detection rate R_{AB}^{det}. Insets (e) and (f) show $\sigma(\bar{\Delta})$ of an SNSPD with respect to the bias current. The measured standard errors are plotted as black dots, and the lower bound on the standard deviation is represented as a gray dashed line.

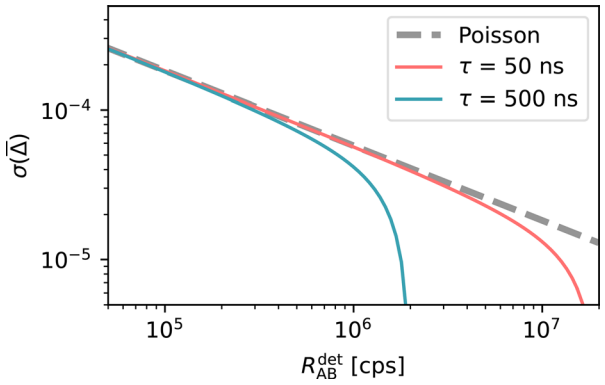


FIG. 8. The effect of dead time saturation on the fundamental limit of uncertainty resolution. The integration time is $T = 20$ s and the number of measurements $N = 30$.

number of accumulated counts is large: $\langle n \rangle = R^{det}T \gg 1$. This is a reasonable assumption for high rates, where saturation occurs. Then, we can make the following derivation.

Upon measuring n detections during a time T, we obtain the average time between detections, $\overline{\Delta t} = T/n$. Because each realization of Δt is independent, the average time $\overline{\Delta t}$ is normally distributed with a standard deviation $\sigma(\overline{\Delta t}) = \sigma(\Delta t)/\sqrt{n}$ according to the central limit theorem. We assume Δt to be a dead time τ_{NP} plus an exponentially distributed delay with a rate parameter λ, which makes $\langle \Delta t \rangle = \tau_{NP} + 1/\lambda$ and $\sigma(\Delta t) = 1/\lambda$. Because $n \gg 1$, the standard deviation is small relative to the mean; $\sigma(\overline{\Delta t}) \ll \langle \Delta t \rangle$. As a result, the mean rate $R^{det} = 1/\overline{\Delta t}$ is also normally distributed and the standard deviation is scaled based on the slope around the mean value,

$$\sigma(R^{det}) = \left| \frac{\partial R^{det}}{\partial \overline{\Delta t}} \right| \sigma(\overline{\Delta t}) = (R^{det})^2 \sigma(\overline{\Delta t}). \tag{C4}$$

Due to the large number of detections, on the right side, we can simplify $\overline{\Delta t} \approx \langle \Delta t \rangle \approx 1/R^{det}$, so

$$\sigma(R^{det}) = (1 - \tau_{NP}R^{det})\sqrt{R^{det}/T}. \tag{C5}$$

The resulting sub-Poissonian bounds are shown in Fig. 8 for two dead time values to illustrate the cases of an actively and passively quenched detectors. The bounds do not make a difference within the data range of Fig. 7.

2. Optimal measurement times

Here, we address the optimal allocation of measurement time to minimize the uncertainty. Given the overall measurement time T_O per one sample of Δ, there exists an optimal distribution between measuring R_A^{det}, R_B^{det}, and R_{AB}^{det}. Assuming a balanced splitting $R_A^{det} \approx R_B^{det}$, the minimum uncertainty of Δ is reached for

$$T_{AB} = \frac{T_O}{1 + \sqrt{2/(1+\Delta)}}, \tag{C6}$$

$$T_A = T_B = \frac{1}{2}(T_O - T_{AB}). \tag{C7}$$

In the regime where $|\Delta| \ll 1$, the approximate ratios are $T_A : T_B : T_{AB} = 0.3 : 0.3 : 0.4$. The assumption that we made on the way is that the measurement time is long enough so that all the accumulated counts C_i (i = A, B, AB) are approximately normally distributed. This holds if $\langle C_i \rangle = R_i^{det}T_i \gg 1$.

APPENDIX D: TWO-BEAM METHOD COMPARED TO AN ATTENUATOR

Here, we illustrate the advantage of the two-beam method compared to single-beam measurements. An analogous way of characterizing nonlinearity is comparing the detection rates of two intensity levels by employing an attenuator of power transmittance η. The nonlinearity Δ essentially tests two quantities that should be equal if the detector was linear; $R_A^{det} + R_B^{det} \stackrel{?}{=} R_{AB}^{det}$. With an attenuator, we can only measure the transmitted power R_η^{det} and total power R_1^{det}, whereas the lost power $R_{1-\eta}^{det}$ is inaccessible. We can, however, analogously test $R_{1-\eta}^{det} \stackrel{?}{=} \frac{1-\eta}{\eta} R_\eta^{det}$. Thus, we construct nonlinearity,

$$\Delta^{att} = \frac{R_\eta^{det} + \frac{1-\eta}{\eta} R_\eta^{det}}{R^{det}} - 1 = \frac{1}{\eta}\frac{R_\eta^{det}}{R^{det}} - 1. \tag{D1}$$

For $\eta = 1/2$, this is equal to nonlinearity Δ (1).

Let us examine a scenario where we attempt to measure nonlinearity for a balanced scheme, but there is a slight error in calibration,

$$\eta = 0.5 + \delta\eta. \tag{D2}$$

Because the value of nonlinearity may span many orders of magnitude, let us examine its *relative* error, meaning the difference proportional to the ideal value of Δ for $\eta = 0.5$,

$$\delta\Delta = \left| \frac{\Delta(0.5 + \delta\eta) - \Delta(0.5)}{\Delta(0.5)} \right|. \tag{D3}$$

This will generally depend on the model of the detector and the incident rate, so let us illustrate the basic behavior for a simple nonparalyzing model with background R_0 and dead time τ,

$$R_\eta^{det} := R^{det}(\eta R) = \frac{1}{\frac{1}{\eta R + R_0} + \tau}. \tag{D4}$$

The nonlinearity for the two-beam measurement is then

$$\Delta(\eta) = \frac{R_\eta^{det} + R_{1-\eta}^{det}}{R_1^{det}} - 1, \tag{D5}$$

while, for the attenuator, an unknown error manifests only in the measured intensity,

$$\Delta^{att}(\eta) = \frac{1}{0.5}\frac{R_\eta^{det}}{R_1^{det}} - 1. \tag{D6}$$

Apart from η, these quantities depend on R, R_0, and τ. When examining the error, we look for the maximum over all incident rates R. Thus, we obtain the maximum relative error (for both nonlinearities),

$$\delta\Delta_{\max}(\delta\eta) = \max_R \left| \frac{\Delta(0.5+\delta\eta)}{\Delta(0.5)} - 1 \right|. \tag{D7}$$

Now, let us assume a small error in calibration $\delta\eta \ll 1$, so we can consider polynomial contributions of $\delta\eta$. Let us also examine the quantities for a certain realistic range of parameters,

$$0 < R_0 < 10^3 \text{ cps}, \tag{D8}$$

$$0 < \tau < 10^{-6} \text{ s}. \tag{D9}$$

For $\delta\Delta^{\text{att}}_{\max}$, we make certain approximations, chiefly $R_0\tau \ll 1$, that yield the maximal rate $R \approx \sqrt{2R_0/\tau}$ and

$$\delta\Delta^{\text{att}}_{\max} \approx \sqrt{\frac{2}{R_0\tau}} \times \delta\eta. \tag{D10}$$

On the other hand, the two-beam nonlinearity error scales more favorably, and with a constant factor,

$$\delta\Delta_{\max} \approx 4 \times \delta\eta^2. \tag{D11}$$

In practical terms, for a 1% deviation in beam splitter transmittance, $\delta\eta = 0.01$, the value of nonlinearity measured by the two-beam method will change by $\delta\Delta \lesssim 0.04\%$ for any realistic dead time and dark counts.

If the same 1% error is applied to an attenuator, the error depends on the detector parameters. For an extreme case of $R_0 = 10^3$ cps and $\tau = 1$ μs, the error is $\delta\Delta^{\text{att}} \lesssim 44\%$, whereas for values corresponding more with modern SPADs, $R_0 = 50$ cps and $\tau = 30$ ns, the value can change drastically by $\delta\Delta^{\text{att}} \lesssim 1200\%$.

By accessing the complementary optical power, the two-beam nonlinearity offers a much more reliable measurement. The dependence on the splitting ratio is quite weak, with the value of Δ changing by $\lesssim$15% within the range $\eta \in (0.3, 0.7)$ relative to $\eta = 0.5$. Furthermore, the balanced splitting ratio is the only one that can be calibrated with arbitrary precision, so it easily becomes the most suitable measurement setting to use, as it represents the nonlinear properties of the detector much more than any technical parameters of the measurement.

APPENDIX E: REPEATABILITY OF THE MEASUREMENT

In 2014, we built the experimental setup for absolute measurement of nonlinearity to characterize the detector response and since then, we have performed nonlinearity measurements repeatedly. As an example representing the consistency of the results obtained from the presented method, we chose the actively quenched SPAD from Laser Components (SPAD-3). This detector exhibits anomalous supralinear behavior ($\Delta < 0$) for all measurements. Figure 9 shows individual nonlinearity measurements separated by years during which the detector was used in many other experimental setups. The first measurement of nonlinearity was performed in November 2014 using the Ortec counter (Fig. 9—blue). In January 2016, we repeated the measurement—again using the Ortec counter—in order to confirm the supralinear response of the tested detector (Fig. 9—green). This was the most accurate measurement, with acquisition time 120 s for each individual rate measurement, repeated 20 times. The last measurement was performed in February 2021 using the Countex counter (Fig. 9—black). Throughout the years, we have obtained consistent results.

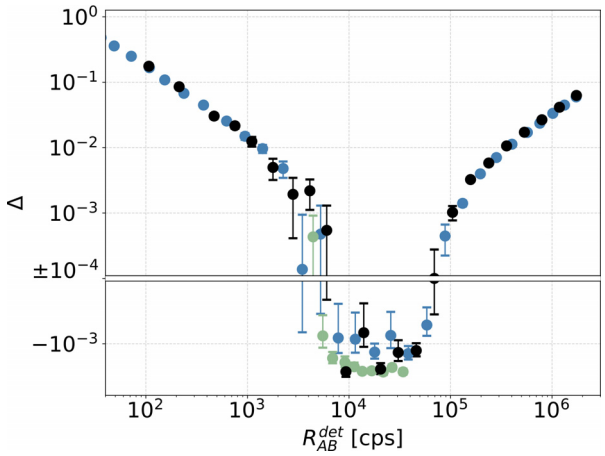


FIG. 9. The repeatability of the nonlinearity measurement of the detector SPAD-3 throughout the years. The nonlinearity Δ is a function of the detection rate $R^{\text{det}}_{\text{AB}}$, shown at different dates: black—February 2021, green—January 2016, and blue—November 2014.

APPENDIX F: THEORETICAL MODELS OF THE SPAD RESPONSE FUNCTION

1. Correcting the data for dead time and background

There have been many SPAD response models proposed that model saturation effects and noise.[34–38,57,64,78] The model used to fit the presented results considers a non-paralyzable dead time τ_{NP}, dark count rate R_0, and neglects afterpulsing. The rate of detection events as a function of the incident rate R then follows a standard formula,

$$f_{\text{NP}}(R) = \frac{R + R_0}{1 + (R + R_0)\tau_{\text{NP}}}. \tag{F1}$$

In Fig. 10, we illustrate the effects of dead time and dark counts, when this model is applied to nonlinearity (1). The left slope represents the background noise characterized by dark counts and the right slope represents dead time saturation. Other theoretical models shown below yield similarly V-shaped nonlinearity Δ with analogous scaling.

To correct for nonlinearities under this model, one can apply the formula,

$$R_i^{\text{corr}} = \frac{R_i^{\text{det}}}{1 - R_i^{\text{det}}\tau_{\text{NP}}} - R_0 \tag{F2}$$

with $i \in \{\text{A, B, AB}\}$ to the measured detection rates. Figure 11 shows the nonlinearity values calculated from corrected rates. The same data as presented in the main text are used, and the values of R_0 and τ_{NP} are taken from the least squares fits (see Fig. 12 and Table III).

If model (F1) holds, then the corrected nonlinearity should be close to zero. However, due to the Poissonian uncertainty (C5) and the nonlinear nature of the formula (C1), there is a nonzero bias (solid line in Fig. 11). The exact probability distribution of Δ^{corr} needs to be computed numerically for $R^{\text{corr}}_{\text{AB}} < 10^2$ cps, because a normal approximation no longer holds.[74]

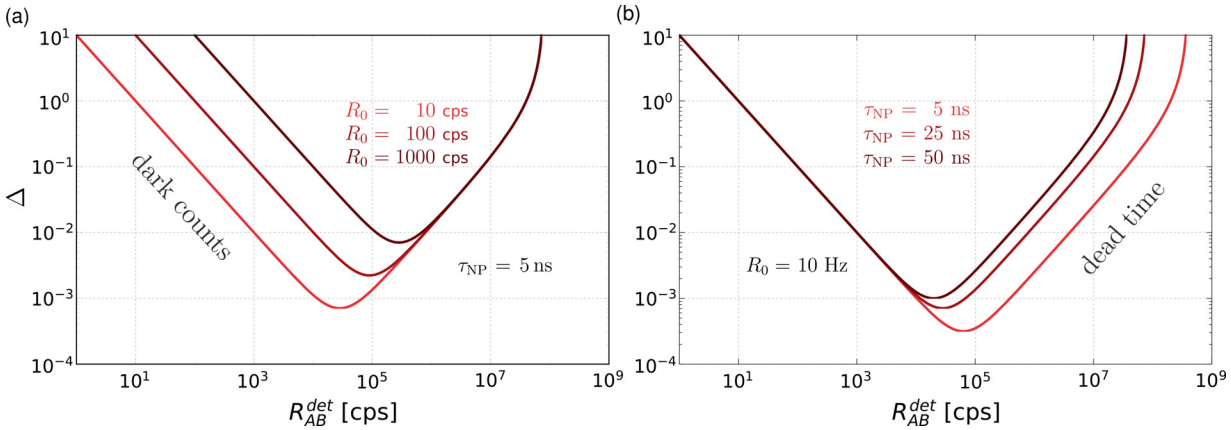


FIG. 10. Nonlinearity Δ as a function of the detection rate R_{AB}^{det} for (a) a several different values of dark counts: $R_0 = 10, 100, 1000$ cps ($\tau =$ const.) and (b) dead time: $\tau = 5, 25, 50$ ns ($R_0 =$ const.).

Like in the main text, each point represents the average from 30 measurements. To test the compatibility of the corrected data and model (F1), we test both the average value and the distribution of the 30 samples for each point. The average nonlinearity is compared to a 95% confidence interval, and the Kolmogorov–Smirnov (KS) p-value of the distribution is denoted by the point color. The p-value is a probability that, assuming the model holds, we would obtain "worse" data than those measured, as parameterized by the KS statistic (see Appendix G below). Very low p-values show that statistical errors alone are unlikely to explain the data.

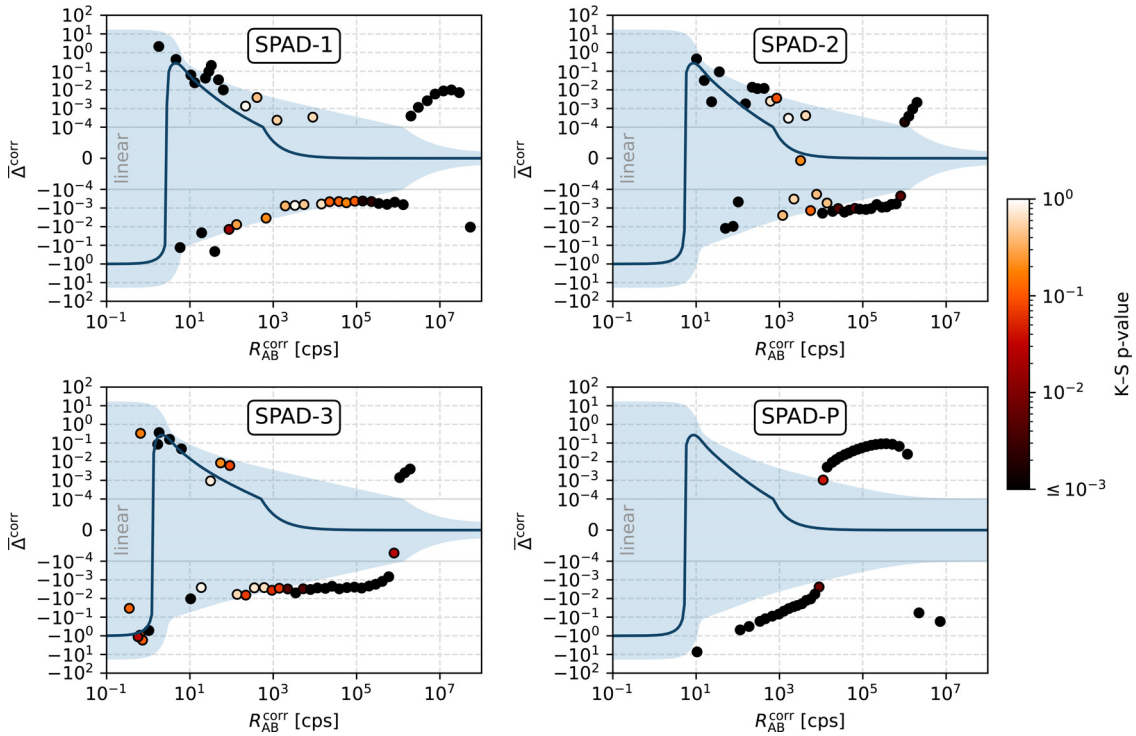


FIG. 11. Measured average nonlinearities after all detection rates are corrected for dead time and dark-count effects (F2) based on the best-fit values given in Table III. The points represent 30-sample averages; the solid lines are the theoretical mean values assuming model (F1) holds; the blue area is a 95% confidence interval for the average values. Point color shows the Kolmogorov–Smirnov p-value, which tests whether the 30 constituent samples come from the expected distribution. Dark points are very unlikely to occur under the presumed model.

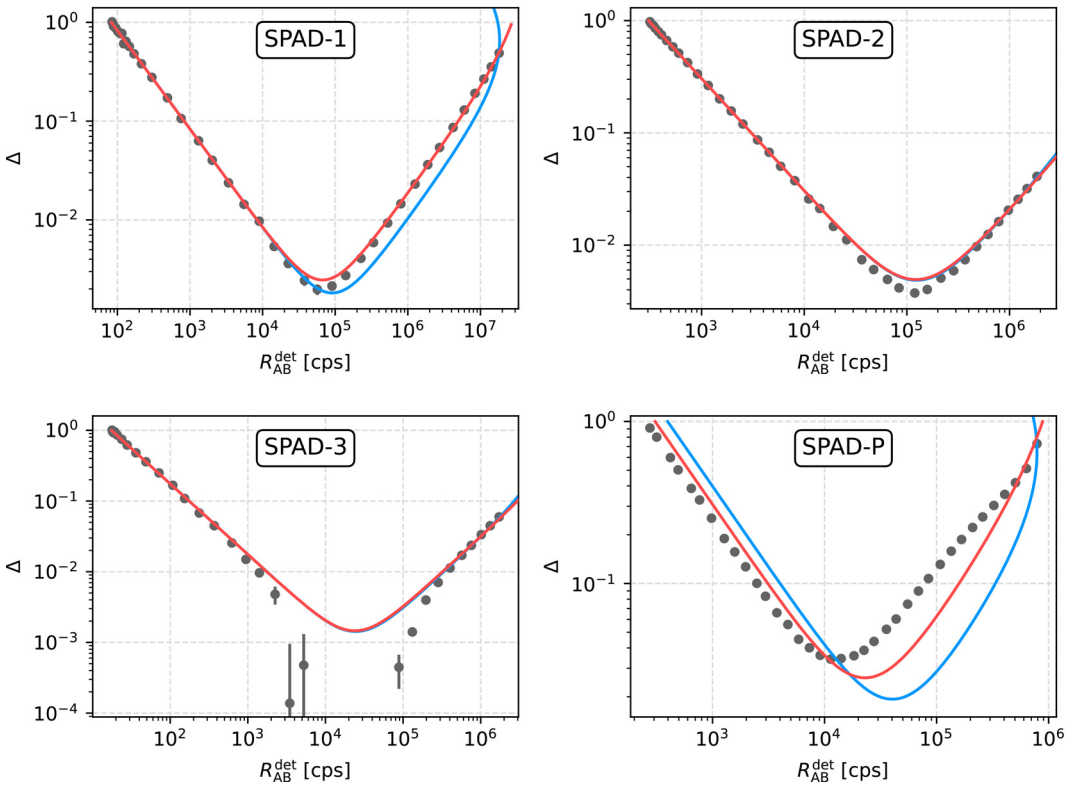


FIG. 12. Nonlinearity data fits using two models: the τ_{NP} model (red) and the τ_P model (blue). Note the right-side turning point of the paralyzable model (blue), which is a consequence of non-monotonous response to illumination.

Many points show disagreement with the theoretical model, with the difference standing out more in comparison to uncorrected data (Fig. 12), especially for $R_{AB}^{det} \gtrsim 10^6$ cps.

Accordingly, basic model (F1) does not provide a satisfactory explanation of the measured data, so we should justify its use as opposed to other dead time models and quantify the effect of afterpulsing.

2. Dead time models

The effect of dead time in Geiger-mode SPADs is akin to Geiger–Müller (GM) counters.[58–62,78–82] Two basic types of idealized models for dead time have been defined.[59,60,79] Namely, it is the paralyzable dead time τ_P model,

$$f_P(R) = (R + R_0)\exp\left(-(R + R_0)\tau_P\right), \tag{F3}$$

and the non-paralyzable dead time τ_{NP} model (F1). For the non-paralyzable case, each *registered* detection is followed by dead time, during which no further events are registered. In the paralyzable case, dead time follows every photon absorption, even those that occur within a previous dead time and are not otherwise recorded. This case covers the fact that secondary detections in GM tubes still require quenching but are not recognized due to low voltage output.

TABLE III. Comparison of directly measured dark counts R_0 and recovery times τ_R, and dark counts and dead times τ_{NP}, τ_P, that were the best fits of individual response models. The last two columns show which scheme the hybrid models converge to.

	Measured		τ_{NP} model		τ_P model			
	R_0 (cps)	τ_R (ns)	R_0 (cps)	τ_{NP} (ns)	R_0 (cps)	τ_P (ns)	τ_{NP}-τ_P model	τ_P-τ_{NP} model
SPAD-1	88(3)	29.5(5)	83(4)	36.7(1)	80(40)	20.0(1)	τ_{NP} model	τ_{NP} model
SPAD-2	314(5)	47.0(5)	304(2)	40.2(4)	304(1)	38.9(3)	τ_P model	τ_P model
SPAD-3	20(2)	56.6(6)	17.5(7)	61(1)	17.5(6)	58.2(9)	τ_P model	τ_P model
SPAD-P	264(1)	517(6)	300(200)	1130(20)	400(500)	467	τ_{NP} model	τ_{NP} model

TABLE IV. Evaluated χ^2/ν values of the fitted theoretical response models for tested SPADs.

χ^2/ν	τ_{NP} model	τ_P model
SPAD-1	250	27 000
SPAD-2	12	8.4
SPAD-3	40	31
SPAD-P	9.8×10^4	1.1×10^6

In the case of GM counters, single-parameter dead time models are just an approximation. Hybrid models were proposed by combining paralyzable and non-paralyzable dead times.[63] There are two variants; the NP-P model,

$$f_{NP-P}(R) = \frac{(R+R_0)\exp\left(-(R+R_0)\tau_P\right)}{1+(R+R_0)\tau_{NP}}, \tag{F4}$$

and the P-NP model,[61,83]

$$f_{P-NP}(R) = \frac{(R+R_0)\exp\left(-(R+R_0)\tau_P\right)}{1+(R+R_0)\tau_{NP}\exp\left(-(R+R_0)\tau_P\right)}. \tag{F5}$$

We have tested whether these theoretical models can be used empirically to fit the measured nonlinearity data. We found that all hybrid-model fits converge to either the paralyzable or non-paralyzable case. Figure 12 shows both of these response models applied to the measured nonlinearity of all detectors. A complete list of the measured parameters and best-fit parameters is given in Table III.

As in the main text, we evaluated χ^2 to demonstrate that investigated models significantly deviate from the measured nonlinearity. In Table IV, we show the chi-squared per one degree of freedom, χ^2/ν, where ν is the number of data points minus the number of fitting parameters.

Both the χ^2 values and the plots in Fig. 12 show that the paralyzable and hybrid models do not offer a better explanation of the data nor can they satisfactorily match the trend in the middle of the graphs of SPAD-1–3, where all nonlinearities seem to be systematically lower.

3. Afterpulsing

Actively quenched SPADs exhibit afterpulsing and twilight pulsing that affect the mean detection rate.[38] Both effects can be evaluated numerically,[84] or—if we neglect the temporal distribution of afterpulses—an approximate rate formula can be used,[38]

$$f_{AP}(R) = \left(\left[\frac{1}{(R+R_0)} - \alpha\right]e^{-\langle n_{AP}\rangle} + \tau_{NP}\right)^{-1}. \tag{F6}$$

The new parameters are the mean number of afterpulses per detection $\langle n_{AP}\rangle$, and the twilight-pulse proportionality constant α that introduces rate dependence. As one would expect, when both of these parameters are zero, formula (F6) is reduced to basic non-paralyzable model (F1).

The key observation here is parameter degeneracy. In formula (F6), the effect of afterpulsing can be substituted by adjusting the dead time and dark count variables in (F1). If we put the model parameters in square brackets, the equivalence can be expressed as

$$f_{AP}(R)[R_0,\tau_{NP},\langle n_{AP}\rangle,\alpha] \equiv f_{NP}\left(R/e^{-\langle n_{AP}\rangle}\right)\left[R_0/e^{-\langle n_{AP}\rangle},\tau_{NP}-\alpha e^{-\langle n_{AP}\rangle}\right]. \tag{F7}$$

Let us now substitute both rate models into nonlinearity (1), assuming balanced splitting $R_A^{det} = R_B^{det}$ and the parameter equivalence (F7),

$$\Delta = \frac{2R_A^{det}}{R_{AB}^{det}} - 1 = \frac{2f\left[\frac{1}{2}f^{-1}(R_{AB}^{det})\right]}{R_{AB}^{det}} - 1. \tag{F8}$$

One finds that $\Delta_{AP} \equiv \Delta_{NP}$, which follows from the equivalence (F7) and the scaling of R by a linear factor there. The principle is that both f_{AP}^{-1} and f_{NP}^{-1} yield incident rates that differ by the same linear factor, and their ratio does not change upon the multiplication by 1/2. A subsequent application of f, therefore, gives identical rates for both models. As a result, all fitted models are identical and yield the same χ^2 either with or without afterpulsing.

A further consideration would be a more complex afterpulsing model that considers the temporal distribution of afterpulses.[38] This model is based on computationally demanding Monte Carlo simulations and is, therefore, unsuitable for least squares fitting. However, the difference between this full model Δ_{full} and simple model (F6) can be evaluated for an example case. In Fig. 13, the difference is shown for SPAD-1 based on the afterpulsing data measured in Ref. 38. Note that we have already established the equivalency of (F1) and (F6), so we can conclude that the full model introduces a very small correction to the basic model that we used to fit the data.

4. Discussion of the models

An example case of an actively quenched SPAD, represented by SPAD-1, shows that an afterpulsing correction to our nonlinearity model would yield a relative difference of ≈0.1% in Δ. A slightly simplified afterpulsing correction that can be used for data fitting yields no difference from the basic model. For this reason, we can

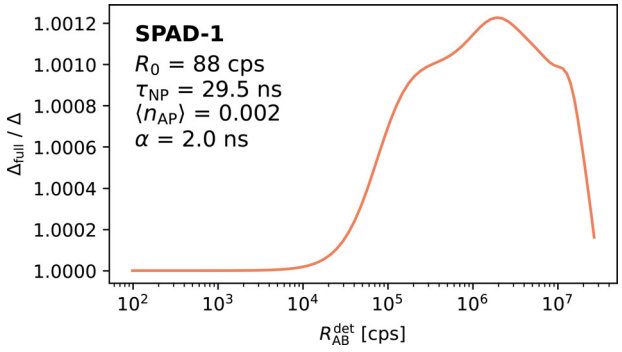


FIG. 13. The full afterpulsing model is compared to simple model (F6) and, thus, to an equivalent basic model (F1) using (F7). The difference is shown for SPAD-1. It corresponds to a relative change of the fits depicted in Fig. 12 if a full afterpulsing correction was made.

conclude that afterpulses do not play a role in fitting the nonlinearity data of actively quenched SPADs.

The other detectors exhibit a much more complex behavior. Passively quenched SPADs exhibit paralyzable dead time, but also a rising efficiency curve after each detection,[85] which makes the model too complex to be parameterized by a few measurable quantities.

SNSPDs show no afterpulsing, but, to our knowledge, their (non-)paralyzability has not been confirmed. Additionally, both SNSPDs and passively quenched SPADs exhibit changing efficiency after each detection, an effect that is both non-negligible and complex. In the case of SNSPDs, the single-photon and two-photon efficiencies also depend on the bias current. For these reasons, no standard rate model has been formulated that would be sufficiently accurate.

Upon exploration of certain empirical models combining non-paralyzable and paralyzable dead time, we found that none of them offer a significant advantage over the basic non-paralyzable model. SPAD-1–3 are non-paralyzable. SPAD-2 and SPAD-3 are better fitted with the paralyzable model, but the difference is very small, as can be seen from Fig. 12. SPAD-P is significantly better fitted by the non-paralyzable model. Consequently, we use single model (F1) for all SPADs in the manuscript.

5. *Ad hoc* hypothesis

The most significant supralinearity ($\Delta < 0$) among SPADs is exhibited by SPAD-3. Here, we use an *ad hoc* model to fit the data and assess the feasibility of various factors as possible explanations of the supralinearity.

The data can be fit with a modified version of formula (F1),

$$f(\Phi) = \frac{1}{(a_0 + a_1\Phi + a_2(\Phi/\Phi_1)^b)^{-1} + \tau}, \tag{F9}$$

where the term a_0 corresponds to a constant dark count rate, a_1 corresponds to a constant detection efficiency, and Φ is the incident photon flux. The last term (with $b = 1.004$ and $R_1 = 10^6$ cps) is empirical and can be attributed to either of these parameters as a hypothetical rate-dependent term.

Figure 14 depicts the hypothesis (a), and the corresponding dependence of dark counts (b) and efficiency (c) that would be required to match the hypothesis. The dark count rate dependence is arbitrary up to a linear factor $a\Phi$, but we can calculate the absolute minimum dark count rate directly from the data [green points in Fig. 14(b)].

Dependence on the dark count rate was ruled out by an independent pulsed measurement, where the count rate can be inferred from detections registered outside the pulse window. The dark count rate was found to be below 30 cps for a signal of 10^5 cps mean detection rate, whereas the nonlinearity data would require at least $R_0(10^5\ \text{cps}) > 1000$ cps. This means that rate-dependent dark count rate is not the cause behind the supralinearity of SPAD-3.

If we attribute the supralinearity to rate-dependent efficiency, we are left with relative changes depicted in Fig. 14(c). While efficiency decreasing with rate has recently been reported in Ref. 65 for InGaAs SPADs subjected to pulsed signals, here the efficiency would need to increase. For a SPAD, the efficiency

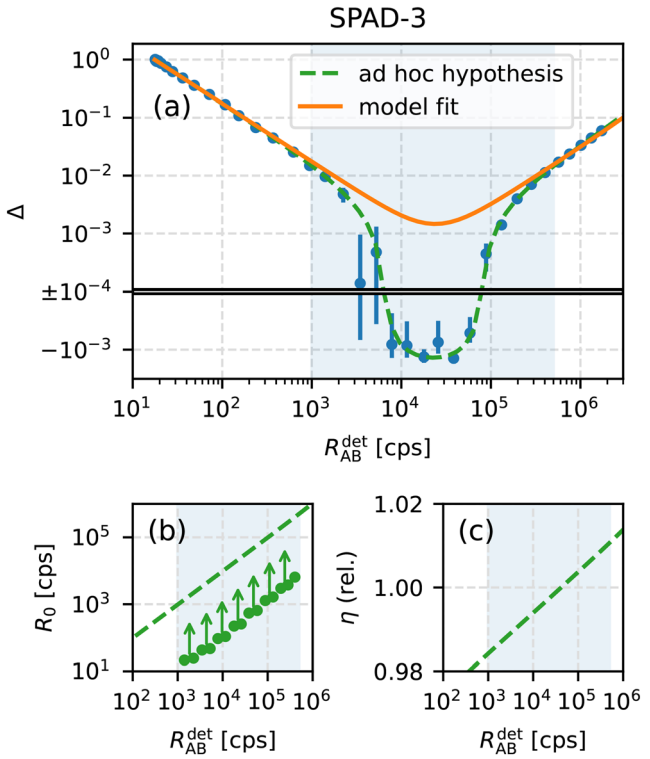


FIG. 14. Depiction of the hypothetical rate dependencies of dark counts and detection efficiency that would be necessary to account for the suprealinearity of SPAD-3 in the highlighted region. (a) depiction of an *ad hoc* model; followed by respective attribution of the extra empirical term to dark counts (b) and efficiency (c). The points in (b) represent the minimal values calculated from the data points.

increases with bias voltage, which should be fixed in its steady state by a DC voltage supply. However, we need to rule out any temporal dependence of the bias voltage, which is addressed in Subsection 6 of Appendix F.

6. Efficiency settling

A possible explanation for the supra-linear response of a SPAD would be time-dependent efficiency after each detection. When an actively quenched detector resets, the overall response may fluctuate.[57] In order for a detector to exhibit supra-linearity, the dominant effect would have to include decreasing efficiency over time after each detection. The slope of this decrease needs to be significant on time scales close to inverse mean count rates where supra-linearity occurs (∼10–100 μs). This could, in principle, be caused by a settling bias voltage, which affects detection efficiency.

In an attempt to roughly estimate the necessary magnitude of such an effect, we modeled the efficiency with an exponential. The results are shown in Fig. 15. To achieve supralinearity roughly similar to the data observed on SPAD-3, the relative decrease in efficiency would have to be 2.5% over more than 0.1 ms. However, as the quenching electronics operate on a GHz timescale, such a slow and significant effect is extremely unlikely.

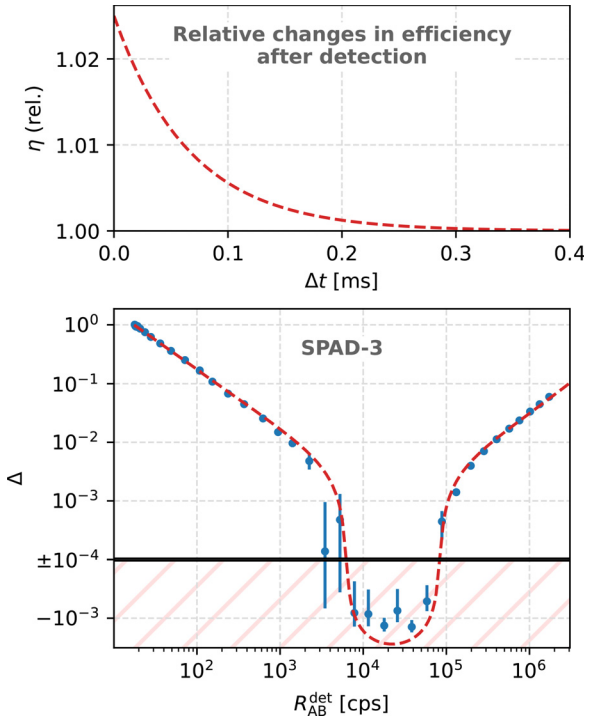


FIG. 15. Estimation of the necessary magnitude of the efficiency settling effect. If the detection efficiency η settled as depicted in the top graph, the measured supralinearity could be approximately achieved.

This effect can be ruled out based on interarrival histograms of detections under continuous illumination. For constant efficiency, interarrival times follow an exponential distribution, save for dead time (here 57 ns) and afterpulsing effects that take place on time scales <10 μs. For time-dependent efficiency, the interarrival time Δt is distributed with the probability density,

$$p(\Delta t) \approx R\eta(\Delta t) \exp\left(-R\int_0^{\Delta t} \eta(t')\mathrm{d}t'\right), \tag{F10}$$

where R is the incident rate and η is a relative term describing the changes in detection efficiency. Interarrival histograms then sample this distribution.

Let us assume the hypothesis shown in Fig. 15 that $\eta(\Delta t) = 1 + 0.025\exp(-15\Delta t \times \mathrm{ms}^{-1})$. Then, for $\Delta t \gtrsim 0.2$ ms, $\eta(\Delta t) \approx 1$ and the distribution will scale as an exponential—$p(\Delta t) \propto \exp(-R\Delta t)$—which is the same scaling as for constant efficiency. On shorter timescale, $\Delta t \lesssim 0.2$ ms, the distribution will be more complex.

Data in Fig. 16 show the measured interarrival histograms relative to the two models—constant and settling efficiency. Because the difference is small, the histogram values $p_{\mathrm{meas}}(\Delta t)$ are divided by the expected values $p(\Delta t)$ to visually distinguish the two cases more clearly. A horizontal trend means that the scaling corresponds to the expectation up to a constant proportionality factor. It can be seen that the data follow a negative exponential for constant efficiency, whereas the comparison to settling efficiency does not result

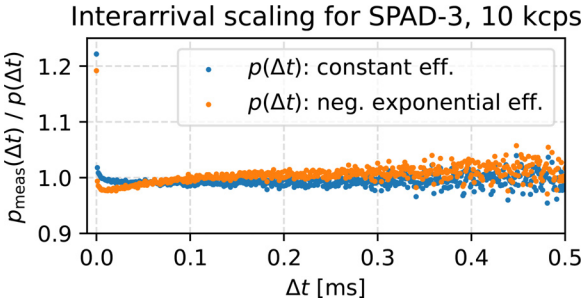


FIG. 16. Scaling of interarrival times for CW illumination at 10 kcps for SPAD-3. The data are taken relative to expected histogram values for constant efficiency and settling efficiency as depicted in Fig. 15. One point corresponds to a 1 μs window. The peak near zero is due to afterpulsing. The model of constant efficiency shows a consistent ratio of its expected values relative to the data, as opposed to settling efficiency, where the scaling is different below $\Delta t < 0.2$ ms.

in a horizontal data trend for $\Delta t < 0.2$ ms. This provides good evidence that the efficiency settling effect is not the reason for supralinearity of SPAD-3.

7. Suprealinearity summary

Here, we summarize the factors that we rule out as causes of supralinearity, as demonstrated on SPAD-3.

a. Dead time

Hybrid dead time models do not satisfactorily fit the data. The non-paralyzable nature of dead time (more strictly, recovery time) is directly measurable and has been observed to slightly increase with the count rate,[38] which actually works against supralinearity. To match the data, dead time values would have to reach unphysical negative values.

b. Dark counts

Dark counts are known to fluctuate,[64] but for nonlinearity, systematic dependence of the mean dark count rate on the incident rate is needed. This dependence would need to span orders of magnitude, which was ruled out by a pulsed measurement.

c. Efficiency

Detection efficiency depends on the SPAD bias voltage, which is held constant by an active quenching circuit. There could be, in theory, a temporal dependence after each detection. However, the settling effect necessary to explain supralinearity between 10^4 and 10^5 cps would have to be both significant and low-frequency, which was ruled out by time-resolved measurements.

d. Afterpulsing

The probability of an afterpulse has a constant and linear contribution for CW signals.[37,38] This was shown to have no effect on nonlinearity, and the effect of the afterpulse temporal distribution is negligible, as we demonstrated.

The above considerations are based on current knowledge of SPAD operation, and none of them offer an explanation for the supralinearity that we observed.

APPENDIX G: STATISTICAL METHODS

This section elaborates on the statistical methods used to analyze and fit the SPAD data.

Due to the integration time $T = 20$ s, all the measured numbers of counts are $R^{\mathrm{det}}T > 300$, which is sufficient for the normal approximation of the Poisson distribution. We, therefore, assume that all acquired count rates are normally distributed $\mathcal{N}(\theta, \sigma^2)$ with variance σ^2 (C5) as a function of a mean-value parameter θ,

$$R^{\mathrm{det}} \sim \mathcal{N}\big(\theta, (1-\theta\tau_{\mathrm{NP}})^2\theta/T\big). \tag{G1}$$

We also assume that nonlinearity Δ, as a function of these rates, is also normally distributed, which requires the standard deviations to be small enough with respect to the mean values. The estimation of Δ is then carried out based on $N = 30$ samples. Due to the normality of the data, a least squares fit of the model can be performed (Fig. 12).

However, when applying the correction (F2) on the data, the normality of Δ ceases to hold. This happens when the detection rate approaches the background rate, as then the corrected value fluctuates significantly compared to the mean value. For such low rates, dead time correction can be neglected and $R^{\mathrm{corr}} \approx R^{\mathrm{det}} - R_0$,

$$R^{\mathrm{corr}} \sim \mathcal{N}(R, (R+R_0)/T) \quad \text{for} \quad R \ll 1/\tau_{\mathrm{NP}}, \tag{G2}$$

where R denotes the incident rate. Due to Poissonian fluctuations of the dark counts, the corrected values can also end up negative. The probability density of

$$\Delta^{\mathrm{corr}} = \frac{R_{\mathrm{A}}^{\mathrm{corr}} + R_{\mathrm{B}}^{\mathrm{corr}}}{R_{\mathrm{AB}}^{\mathrm{corr}}} - 1 \tag{G3}$$

can be computed from (G2) by using the probability-density transformations,

$$p_{(\mathrm{Z=X+Y})}(z) = \int_{-\infty}^{\infty} p_{\mathrm{X}}(t)p_{\mathrm{Y}}(z-t)\mathrm{d}t, \tag{G4}$$

$$p_{(\mathrm{Z=X/Y})}(z) = \int_{-\infty}^{\infty} p_{\mathrm{X}}(zt)p_{\mathrm{Y}}(t)|t|\mathrm{d}t. \tag{G5}$$

For $R \to 0$, the distribution of Δ^{corr} tends toward the Lorentz distribution with undefined variance, so that one cannot employ the central limit theorem to arrive at estimations of N-sample averages. As per (G4), the probability density of $\sum_{i=1}^{N} \Delta_i^{\mathrm{corr}}$ is computed numerically by $(N-1)$ successive convolutions. This allows plotting the mean value and confidence intervals shown in Fig. 11, as well as evaluating the Kolmogorov–Smirnov statistic.

For $R \gg 0$, the distribution of Δ^{corr} can be regarded as normal, with its standard deviation obtained by locally linear propagation of (G1). However, as we are concerned with small values of Δ, we need to account for a bias in the mean value.

We note that all the measured rates in (G3) are independent, and the scheme is balanced, $\langle R_{\mathrm{A}}^{\mathrm{corr}}\rangle = \langle R_{\mathrm{B}}^{\mathrm{corr}}\rangle$, and so

$$\langle \Delta^{\mathrm{corr}}\rangle = 2\langle R_{\mathrm{A}}^{\mathrm{corr}}\rangle \left\langle \frac{1}{R_{\mathrm{AB}}^{\mathrm{corr}}} \right\rangle - 1. \tag{G6}$$

We employ the function (F2), where R_i^{det} is a normal-distributed variable described by (G1). The mean values θ_i for $i = \mathrm{A}, \mathrm{AB}$ are given by the function (F1), and the incident rate is R for $i = \mathrm{AB}$ and $R/2$ for $i = \mathrm{A}$. The main source of the bias is nonlinear dependence of the averaging terms on the random variables. We expand the averaging terms in a Taylor series,

$$f(R_i^{\mathrm{det}}) \approx \sum_{n=0}^{2} \left(\frac{\partial^n f}{\partial (R_i^{\mathrm{det}})^n} \right)(\theta_i) \frac{(R_i^{\mathrm{det}} - \theta_i)^n}{n!}, \tag{G7}$$

where it is sufficient to stop at the quadratic term. Averaging over $R_i^{\mathrm{det}} \sim \mathcal{N}(\theta_i, \sigma_i^2)$ yields non-zero only for even $n = 2k$,

$$\left\langle \frac{(R_i^{\mathrm{det}} - \theta_i)^{2k}}{(2k)!} \right\rangle = \frac{(\sigma_i^2/2)^k}{k!}. \tag{G8}$$

We can also safely assume a short dead time and low background—$\tau \ll T$ and $\tau R_0 \ll 1$. This all results in

$$\langle \Delta^{\mathrm{corr}}\rangle \approx \frac{R+R_0}{TR^2}, \tag{G9}$$

which holds well for $R \gtrsim R_0$.

The above methods allow us to establish the probability distribution of $\bar{\Delta}^{\mathrm{corr}} = \frac{1}{N}\sum_{i=1}^{N} \Delta_i^{\mathrm{corr}}$ for the whole range of incident rates R. Then, one way of testing the validity of the data is the confidence interval of $\bar{\Delta}^{\mathrm{corr}}$. A second test evaluates whether the samples Δ^{corr} conform to the expected distribution. However, due to non-normality, it would be difficult to evaluate measures based on multivariate probability density akin to χ^2. Instead, we compare cumulative distributions.

The Kolmogorov–Smirnov statistic D_N is defined as the maximum difference between the empirical and theoretical cumulative distributions. For a certain $R_{\mathrm{AB}}^{\mathrm{det}}$, let the number of samples of Δ^{corr} lower than Δ be $n(\Delta)$. Then, the empirical cumulative distribution is $S_N(\Delta) = n(\Delta)/N$ with $S_N \in [0, 1]$. Let the theoretical cumulative distribution be $S(\Delta) = \Pr[\Delta^{\mathrm{corr}} < \Delta]$. We compare them by defining the statistic

$$D_N := \sup_{\Delta} |S_N(\Delta) - S(\Delta)|. \tag{G10}$$

Then, for many samples, as $N \to \infty$, the quantity $(\sqrt{N}D_N)$ follows a known distribution assuming $S(\Delta)$ holds.[86] This quantity can, therefore, be evaluated for each set of N measurements, as depicted in Fig. 11. The p-value is then the probability $\Pr[D_N > D_N^{\mathrm{meas}}]$, which tests the null hypothesis that the data conform to the theoretical model.

REFERENCES

[1] V. Giovannetti, S. Lloyd, and L. Maccone, "Advances in quantum metrology," Nat. Photonics **5**, 222–229 (2011).

[2] K. R. Motes, J. P. Olson, E. J. Rabeaux, J. P. Dowling, S. J. Olson, and P. P. Rohde, "Linear optical quantum metrology with single photons: Exploiting spontaneously generated entanglement to beat the shot-noise limit," Phys. Rev. Lett. **114**, 170802 (2015).

[3] The LIGO Scientific Collaboration, "Enhanced sensitivity of the LIGO gravitational wave detector by using squeezed states of light," Nat. Photonics **7**, 613–619 (2013).

[4] S. Slussarenko, M. M. Weston, H. M. Chrzanowski, L. K. Shalm, V. B. Verma, S. W. Nam, and G. J. Pryde, "Unconditional violation of the shot-noise limit in photonic quantum metrology," Nat. Photonics **11**, 700–703 (2017).

[5]G. Brida, M. Genovese, and I. Ruo Berchera, "Experimental realization of sub-shot-noise quantum imaging," Nat. Photonics **4**, 227–230 (2010).
[6]M. A. Taylor, J. Janousek, V. Daria, J. Knittel, B. Hage, H.-A. Bachor, and W. P. Bowen, "Biological measurement beyond the quantum limit," Nat. Photonics **7**, 229–233 (2013).
[7]T. Ono, R. Okamoto, and S. Takeuchi, "An entanglement-enhanced microscope," Nat. Commun. **4**, 2426 (2013).
[8]Y. Israel, R. Tenne, D. Oron, and Y. Silberberg, "Quantum correlation enhanced super-resolution localization microscopy enabled by a fibre bundle camera," Nat Commun. **8**, 14786 (2017).
[9]P.-A. Moreau, E. Toninelli, T. Gregory, and M. J. Padgett, "Imaging with quantum states of light," Nat. Rev. Phys. **1**, 367–380 (2019).
[10]D. A. Kalashnikov, A. V. Paterova, S. P. Kulik, and L. A. Krivitsky, "Infrared spectroscopy with visible light," Nat. Photonics **10**, 98–101 (2016).
[11]R. Whittaker, C. Erven, A. Neville, M. Berry, J. L. O'Brien, H. Cable, and J. C. F. Matthews, "Absorption spectroscopy at the ultimate quantum limit from single-photon states," New J. Phys. **19**, 023013 (2017).
[12]E. Jakeman and J. Rarity, "The use of pair production processes to reduce quantum noise in transmission measurements," Opt. Commun. **59**, 219–223 (1986).
[13]J. Sabines-Chesterking, R. Whittaker, S. K. Joshi, P. M. Birchall, P. A. Moreau, A. McMillan, H. V. Cable, J. L. O'Brien, J. G. Rarity, and J. C. F. Matthews, "Sub-shot-noise transmission measurement enabled by active feed-forward of heralded single photons," Phys. Rev. Appl. **8**, 014016 (2017).
[14]E. Losero, I. Ruo-Berchera, A. Meda, A. Avella, and M. Genovese, "Unbiased estimation of an optical loss at the ultimate quantum limit with twin-beams," Sci. Rep. **8**, 7431 (2018).
[15]J. Sabines-Chesterking, A. R. McMillan, P. A. Moreau, S. K. Joshi, S. Knauer, E. Johnston, J. G. Rarity, and J. C. F. Matthews, "Twin-beam sub-shot-noise raster-scanning microscope," Opt. Express **27**, 30810–30818 (2019).
[16]U. Sinha, C. Couteau, T. Jennewein, R. Laflamme, and G. Weihs, "Ruling out multi-order interference in quantum mechanics," Science **329**, 418–421 (2010).
[17]I. Söllner, B. Gschösser, P. Mai, B. Pressl, Z. Vörös, and G. Weihs, "Testing Born's rule in quantum mechanics for three mutually exclusive events," Found. Phys. **42**, 742–751 (2012).
[18]O. S. Magaña-Loaiza, I. De Leon, M. Mirhosseini, R. Fickler, A. Safari, U. Mick, B. McIntyre, P. Banzer, B. Rodenburg, G. Leuchs, and R. W. Boyd, "Exotic looped trajectories of photons in three-slit interference," Nat Commun. **7**, 13987 (2016).
[19]T. Kauten, R. Keil, T. Kaufmann, B. Pressl, Č. Brukner, and G. Weihs, "Obtaining tight bounds on higher-order interferences with a 5-path interferometer," New J. Phys. **19**, 033017 (2017).
[20]J. P. Cotter, C. Brand, C. Knobloch, Y. Lilach, O. Cheshnovsky, and M. Arndt, "In search of multipath interference using large molecules," Sci. Adv. **3**, e1602478 (2017).
[21]G. Rengaraj, U. Prathwiraj, S. N. Sahoo, R. Somashekhar, and U. Sinha, "Measuring the deviation from the superposition principle in interference experiments," New J. Phys. **20**, 063049 (2018).
[22]C. M. Natarajan, M. G. Tanner, and R. H. Hadfield, "Superconducting nanowire single-photon detectors: Physics and applications," Supercond. Sci. Technol. **25**, 063001 (2012).
[23]A. Migdall, S. V. Polyakov, J. Fan, and J. C. Bienfang, *Single-Photon Generation and Detection* (Academic Press, 2013).
[24]C. J. Chunnilall, I. P. Degiovanni, S. Kück, I. Müller, and A. G. Sinclair, "Metrology of single-photon sources and detectors: A review," Opt. Eng. **53**, 081910 (2014).
[25]*Superconducting Devices in Quantum Optics*, edited by R. H. Hadfield and G. Johansson (Springer International Publishing, 2016).
[26]I. Esmaeil Zadeh, J. Chang, J. W. N. Los, S. Gyger, A. W. Elshaari, S. Steinhauer, S. N. Dorenbos, and V. Zwiller, "Superconducting nanowire single-photon detectors: A perspective on evolution, state-of-the-art, future developments, and applications," Appl. Phys. Lett. **118**, 190502 (2021).
[27]A. Rogalski, M. Kopytko, and P. Martyniuk, "Two-dimensional infrared and terahertz detectors: Outlook and status," Appl. Phys. Rev. **6**, 021316 (2019).
[28]H. Wang, J. Guo, J. Miao, W. Luo, Y. Gu, R. Xie, F. Wang, L. Zhang, P. Wang, and W. Hu, "Emerging single-photon detectors based on low-dimensional materials," Small **18**, 2103963 (2021).
[29]A. R. Schaefer, E. F. Zalewski, and J. Geist, "Silicon detector nonlinearity and related effects," Appl. Opt. **22**, 1232–1236 (1983).
[30]M. Tanabe, K. Amemiya, T. Numata, and D. Fukuda, "Spectral supralinearity prediction of silicon photodiodes in the near-infrared range," Appl. Opt. **54**, 10705–10710 (2015).
[31]M. Tanabe, K. Amemiya, T. Numata, and D. Fukuda, "Spectral supralinearity of silicon photodiodes in visible light due to surface recombination," Appl. Opt. **55**, 3084–3089 (2016).
[32]A. J. Kerman, D. Rosenberg, R. J. Molnar, and E. A. Dauler, "Readout of superconducting nanowire single-photon detectors at high count rates," J. Appl. Phys. **113**, 144511 (2013).
[33]F. Marsili, M. J. Stevens, A. Kozorezov, V. B. Verma, C. Lambert, J. A. Stern, R. D. Horansky, S. Dyer, S. Duff, D. P. Pappas, A. E. Lita, M. D. Shaw, R. P. Mirin, and S. W. Nam, "Hotspot relaxation dynamics in a current-carrying superconductor," Phys. Rev. B **93**, 094518 (2016).
[34]M. Stipčević, D. Wang, and R. Ursin, "Characterization of a commercially available large area, high detection efficiency single-photon avalanche diode," J. Lightwave Technol. **31**, 3591–3596 (2013).
[35]V. Kornilov, "Effects of dead time and afterpulses in photon detector on measured statistics of stochastic radiation," J. Opt. Soc. Am. A **31**, 7–15 (2014).
[36]F. Wang, W. Chen, Y. Li, D. He, C. Wang, Y. Han, S. Wang, Z. Yin, and Z. Han, "Non-Markovian property of afterpulsing effect in single-photon avalanche detector," J. Lightave Technol. **34**, 3610–3615 (2016).
[37]M. A. Wayne, J. C. Bienfang, and S. V. Polyakov, "Simple autocorrelation method for thoroughly characterizing single-photon detectors," Opt. Express **25**, 20352 (2017).
[38]I. Straka, J. Grygar, J. Hloušek, and M. Ježek, "Counting statistics of actively quenched SPADs under continuous illumination," J. Lightwave Technol. **38**, 4765–4771 (2020).
[39]M. K. Akhlaghi and A. H. Majedi, "Semiempirical modeling of dark count rate and quantum efficiency of superconducting nanowire single-photon detectors," IEEE Trans. Appl. Supercond. **19**, 361–366 (2009).
[40]M. K. Akhlaghi, A. H. Majedi, and J. S. Lundeen, "Nonlinearity in single photon detection: Modeling and quantum tomography," Opt. Express **19**, 21305–21312 (2011).
[41]A. Luis and L. L. Sánchez-Soto, "Complete characterization of arbitrary quantum measurement processes," Phys. Rev. Lett. **83**, 3573–3576 (1999).
[42]J. Fiurášek, "Maximum-likelihood estimation of quantum measurement," Phys. Rev. A **64**, 024102 (2001).
[43]J. S. Lundeen, A. Feito, H. Coldenstrodt-Ronge, K. L. Pregnell, C. Silberhorn, T. C. Ralph, J. Eisert, M. B. Plenio, and I. A. Walmsley, "Tomography of quantum detectors," Nat. Phys. **5**, 27–30 (2009).
[44]G. Brida, L. Ciavarella, I. P. Degiovanni, M. Genovese, L. Lolli, M. G. Mingolla, F. Piacentini, M. Rajteri, E. Taralli, and M. G. A. Paris, "Quantum characterization of superconducting photon counters," New J. Phys. **14**, 085001 (2012).
[45]C. M. Natarajan, L. Zhang, H. Coldenstrodt-Ronge, G. Donati, S. N. Dorenbos, V. Zwiller, I. A. Walmsley, and R. H. Hadfield, "Quantum detector tomography of a time-multiplexed superconducting nanowire single-photon detector at telecom wavelengths," Opt. Express **21**, 893–902 (2013).
[46]M. Cooper, M. Karpiński, and B. J. Smith, "Local mapping of detector response for reliable quantum state estimation," Nat. Commun. **5**, 4332 (2014).
[47]S. Grandi, A. Zavatta, M. Bellini, and M. G. A. Paris, "Experimental quantum tomography of a homodyne detector," New J. Phys. **19**, 053015 (2017).
[48]S. Izumi, J. S. Neergaard-Nielsen, and U. L. Andersen, "Tomography of a feedback measurement with photon detection," Phys. Rev. Lett. **124**, 070502 (2020).
[49]S. Ferrari, V. Kovalyuk, A. Vetter, C. Lee, C. Rockstuhl, A. Semenov, G. Gol'tsman, and W. Pernice, "Analysis of the detection response of waveguide-integrated superconducting nanowire single-photon detectors at high count rate," Appl. Phys. Lett. **115**, 101104 (2019).
[50]C. L. Sanders, "Accurate measurements of and corrections for nonlinearities in radiometers," J. Res. Natl. Bur. Stand., Sect. A **76A**, 437 (1972).
[51]J. S. Preston and L. H. McDermott, "The illumination-response characteristics of vacuum photoelectric cells of the Elster-Geitel type," Proc. Phys. Soc. **46**, 256–276 (1934).
[52]L. Coslovi and F. Righini, "Fast determination of the nonlinearity of photodetectors," Appl. Opt. **19**, 3200–3203 (1980).

[53]R. D. Saunders and J. B. Shumaker, "Automated radiometric linearity tester," Appl. Opt. **23**, 3504–3506 (1984).
[54]A. Haapalinna, T. Kübarsepp, P. Kärhä, and E. Ikonen, "Measurement of the absolute linearity of photodetectors with a diode laser," Meas. Sci. Technol. **10**, 1075–1078 (1999).
[55]D.-J. Shin, S. Park, K.-L. Jeong, S.-N. Park, and D.-H. Lee, "High-accuracy measurement of linearity of optical detectors based on flux addition of LEDs in an integrating sphere," Metrologia **51**, 25–32 (2013).
[56]T. Kauten, B. Pressl, T. Kaufmann, and G. Weihs, "Measurement and modeling of the nonlinearity of photovoltaic and Geiger-mode photodiodes," Rev. Sci. Instrum. **85**, 063102 (2014).
[57]M. Ware, A. Migdall, J. C. Bienfang, and S. V. Polyakov, "Calibrating photon-counting detectors to high accuracy: Background and deadtime issues," J. Mod. Opt. **54**, 361–372 (2007).
[58]W. Feller, *On Probability Problems in the Theory of Counters* (Springer, New York, 1948), pp. 105–115.
[59]R. D. Evans, *The Atomic Nucleus, Applications of Poisson Statistics to Some Instruments Used in Nuclear Physics* (McGraw-Hill, New York, 1955), Chap. 28, p. 785.
[60]G. F. Knoll, *Radiation Detection and Measurement* (John Wiley & Sons, New York, 1989) p. 120.
[61]J. W. Müller, "Generalized dead times," Nucl. Instrum. Methods Phys. Res. B **301**, 543–551 (1991).
[62]R. P. Gardner and L. Liu, "On extending the accurate and useful counting rate range of GM counter detector systems," Appl. Radiat. Isot. **48**, 1605–1615 (1997).
[63]S. H. Lee and R. P. Gardner, "A new G-M counter dead time model," Appl. Radiat. Isot. **53**, 731–737 (2000).
[64]M. A. Karami, L. Carrara, C. Niclass, M. Fishburn, and E. Charbon, "RTS noise characterization in single-photon avalanche diodes," IEEE Electron Device Lett. **31**, 692–694 (2010).
[65]S. M. F. Raupach, I. P. Degiovanni, H. Georgieva, A. Meda, H. Hofer, M. Gramegna, M. Genovese, S. Kück, and M. López, "Detection rate dependence of the inherent detection efficiency in single-photon detectors based on avalanche diodes," Phys. Rev. A **105**, 042615 (2022).
[66]S. J. Gibson, B. van Kasteren, B. Tekcan, Y. Cui, D. van Dam, J. E. M. Haverkort, E. P. A. M. Bakkers, and M. E. Reimer, "Tapered InP nanowire arrays for efficient broadband high-speed single-photon detection," Nat. Nanotechnol. **14**, 473–479 (2019).
[67]S. Lei, F. Wen, L. Ge, S. Najmaei, A. George, Y. Gong, W. Gao, Z. Jin, B. Li, J. Lou, J. Kono, R. Vajtai, P. Ajayan, and N. J. Halas, "An atomically layered InSe avalanche photodetector," Nano Lett. **15**, 3048–3055 (2015).
[68]G. Bulgarini, M. E. Reimer, M. Hocevar, E. P. A. M. Bakkers, L. P. Kouwenhoven, and V. Zwiller, "Avalanche amplification of a single exciton in a semiconductor nanowire," Nat. Photonics **6**, 455–458 (2012).
[69]E. J. Gansen, M. A. Rowe, M. B. Greene, D. Rosenberg, T. E. Harvey, M. Y. Su, R. H. Hadfield, S. W. Nam, and R. P. Mirin, "Photon-number-discriminating detection using a quantum-dot, optically gated, field-effect transistor," Nat. Photonics **1**, 585–588 (2007).
[70]W. Luo, Q. Weng, M. Long, P. Wang, F. Gong, H. Fang, M. Luo, W. Wang, Z. Wang, D. Zheng, W. Hu, X. Chen, and W. Lu, "Room-temperature single-photon detector based on single nanowire," Nano Lett. **18**, 5439–5445 (2018).
[71]K. Roy, T. Ahmed, H. Dubey, T. P. Sai, R. Kashid, S. Maliakal, K. Hsieh, S. Shamim, and A. Ghosh, "Number-resolved single-photon detection with ultra-low noise van der Waals hybrid," Adv. Mater. **30**, 1704412 (2017).
[72]E. D. Walsh, W. Jung, G.-H. Lee, D. K. Efetov, B.-I. Wu, K.-F. Huang, T. A. Ohki, T. Taniguchi, K. Watanabe, P. Kim, D. Englund, and K. C. Fong, "Josephson junction infrared single-photon detector," Science **372**, 409–412 (2021).
[73]A. Gao, J. Lai, Y. Wang, Z. Zhu, J. Zeng, G. Yu, N. Wang, W. Chen, T. Cao, W. Hu, D. Sun, X. Chen, F. Miao, Y. Shi, and X. Wang, "Observation of ballistic avalanche phenomena in nanoscale vertical InSe/BP heterostructures," Nat. Nanotechnol. **14**, 217–222 (2019).
[74]J. Hloušek, I. Straka, and M. Ježek (2021). "Experimental observation of anomalous supralinear response of single-photon detectors," GitHub, https://github.com/PepaHlousek/Nonlinearity.
[75]See http://quantum.opticsolomouc.org/archives/1177 for QOLO-made 30 ms optical shutter using RC servo, 2012.
[76]See http://quantum.opticsolomouc.org/archives/2412 for Countex—QOLO-made counter and coincidence unit, 2020.
[77]D. Allan, "Statistics of atomic frequency standards," Proc. IEEE **54**, 221–230 (1966).
[78]J. W. Müller, "Dead-time problems," Nucl. Instrum. Methods Phys. Res. B **112**, 47–57 (1973).
[79]C. Levert and W. Scheen, "Probability fluctuations of discharges in a Geiger–Müller counter produced by cosmic radiation," Phys. **10**, 225–238 (1943).
[80]G. E. Albert and L. Nelson, "Contributions to the statistical theory of counter data," Ann. Math. Stat. **24**, 9–22 (1953).
[81]L. Takacs, "On a probability problem in the theory of counters," Ann. Math. Stat. **29**, 1257–1263 (1958).
[82]J. W. Muller, "A simple derivation of the Takács formula," in *Rapport BIPM-88/3* (Bureau International des Poids et Mesures, 1988).
[83]J. Lee, I. Kim, and H. Choi, "On the dead time problem of a GM counter," Appl. Radiat. Isot. **67**, 1094–1098 (2009).
[84]I. Straka, J. Grygar, J. Hloušek, and M. Ježek (2020). "SPAD counting model," GitHub. https://doi.org/10.24433/CO.8487128.v1
[85]S. Cova, M. Ghioni, A. Lacaita, C. Samori, and F. Zappa, "Avalanche photodiodes and quenching circuits for single-photon detection," Appl. Opt. **35**, 1956–1976 (1996).
[86]W. Feller, "On the Kolmogorov–Smirnov limit theorems for empirical distributions," Ann. Math. Stat. **19**, 117–189 (1948).

Accurate Detection of Arbitrary Photon Statistics

Josef Hloušek, Michal Dudka, Ivo Straka, and Miroslav Ježek*
Department of Optics, Palacký University, 17. listopadu 12, 77146 Olomouc, Czechia

(Received 16 December 2018; published 11 October 2019)

We report a measurement workflow free of systematic errors consisting of a reconfigurable photon-number-resolving detector, custom electronic circuitry, and faithful data-processing algorithm. We achieve an unprecedented accurate measurement of various photon-number distributions going beyond the number of detection channels with an average fidelity of 0.998, where the error is primarily caused by the sources themselves. Mean numbers of photons cover values up to 20 and faithful autocorrelation measurements range from $g^{(2)} = 6 \times 10^{-3}$ to 2. We successfully detect chaotic, classical, nonclassical, non-Gaussian, and negative-Wigner-function light. Our results open new paths for optical technologies by providing full access to the photon-number information without the necessity of detector tomography.

DOI: 10.1103/PhysRevLett.123.153604

The probability distribution of the number of photons in an optical mode carries a great deal of information about physical processes that generate or transform the optical signal. Along with modal structure and coherence, the statistics provides a full description of light. Precise characterization of photon statistics is a crucial requirement for many applications in the field of photonic quantum technology [1], such as quantum metrology [2,3], nonclassical light preparation [4,5], quantum secure communication [6], and photonic quantum simulations [7,8]. Measurement of statistical properties and nonclassical features of light also represents enabling technology for many emerging biomedical imaging and particle-tracking techniques [9–11]. Statistical correlations are routinely applied to quantify the nonclassicality of light [12,13]. Obtaining photon statistics requires repeated measurements using a photon-number-resolving detector (PNRD). The important parameters of PNRDs are dynamic range, speed, and accuracy.

The main result of our Letter is a photon-statistics retrieval method based on expectation-maximization entropy and implemented in a PNRD design that is virtually free of systematic errors. Our results show unprecedented accuracy across dozens of tested optical signals, ranging from a highly sub-Poissonian single-photon state to super-Poissonian thermal light with non-negligible multiphoton content up to $n = 30$. The accuracy is achieved despite leaving all systematic errors uncorrected and operating with raw data. The proposed method also provides faithful $g^{(2)}$ values [14] for states, where the commonly used Hanbury Brown–Twiss measurement would fail due to high multiphoton content [15].

We demonstrate the accuracy of the reported PNRD by performing photon-statistics measurements for many different states of light, from which 25 states are shown in Fig. 1, covering various mean photon numbers and $g^{(2)}$ values. Furthermore, the reconfigurability of the presented PNRD also allows for direct measurement of correlation functions and nonclassicality witnesses [5,16].

Contemporary PNRD technologies all rely on multiplexing with the exception of transition-edge detectors [4,17] that require temperatures below 100 mK, offer rates of 10–100 kHz, and suffer from range-versus-crosstalk compromise. Photon-number resolution using a single

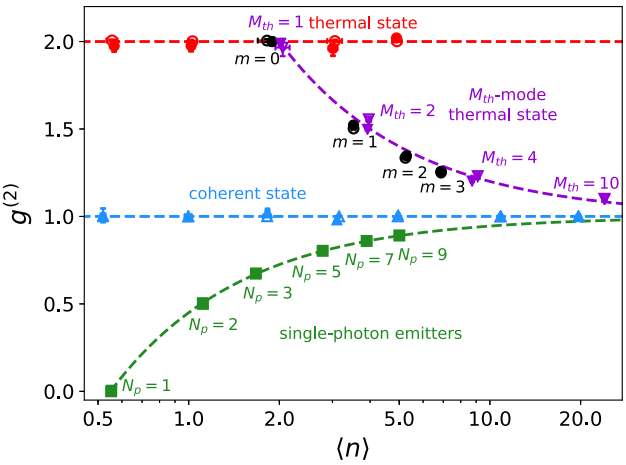


FIG. 1. The autocorrelation $g^{(2)}$ evaluated from the measured photon statistics (solid marker) and the corresponding ideal statistics (empty marker) of various optical signals with mean photon number $\langle n \rangle$ [14]. Shown are the coherent states with $g^{(2)} = 1$ (blue triangle up), thermal states (also termed chaotic light) with $g^{(2)} = 2$ (red circle), M_{th}-mode thermal states with $M_{\mathrm{th}} = 1, 2, 4, 10$ (violet triangle down), and m-photon-subtracted thermal states for $m = 0$, 1, 2, 3 (black circle). The cases of $M_{\mathrm{th}} = 1$ and $m = 0$ coincide with the thermal state. Furthermore, the emission from a cluster of N_p single-photon emitters is shown for $N_p = 1,\ldots,9$ with $g^{(2)} = 1 - 1/N_p$.

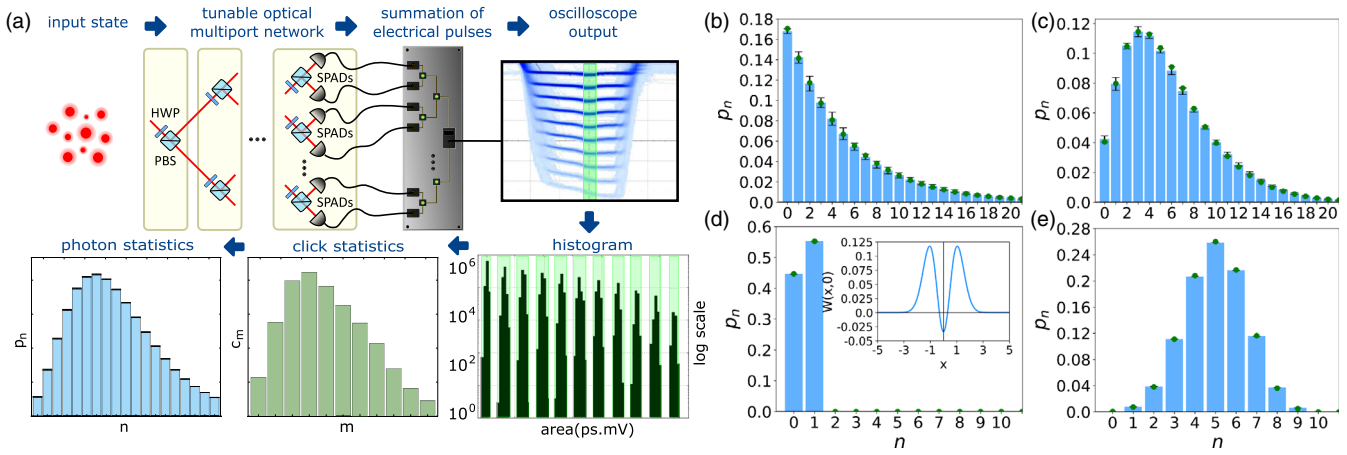


FIG. 2. (a) Experimental setup of the PNRD based on a discrete optical network with full reconfigurability and continuous tunability of splitting ratios, pulse-height spectrum of the analog output of the detector, and scheme of photon-statistics retrieval. Measured (blue bars) and the corresponding theoretical photon statistics (green dots) for (b) thermal state with $\langle n \rangle = 4.93(4)$, (c) two-photon-subtracted thermal state, (d) single-photon, and (e) heralded nine-photon state that emulates emission from a cluster of single-photon emitters. (Inset) Wigner function evaluated from the measured statistics. Note that the data agree with theory even beyond the number of channels of the PNRD (ten).

superconducting nanowire single-photon detector is also possible, but suffers from significant cross talk [18].

The multiplexing approach is based on dividing an input optical signal into multiple on-off detectors [19,20]. Many schemes of temporal and spatial multiplexing have been reported using bulk on-off detectors [21–25], integrated on-off pixels [26–31], or even a few photon-number-resolving detectors [4,32]. Though being economical in respect to the number of on-off detectors employed, the temporal multiplexed scheme trades off a decrease of the detector speed for an increase in a number of the detection channels. Decreasing the losses and the balancing of temporal multiplexors require a great deal of optimization [33] or even active signal switching [25]. On the other hand, multiple-pixel PNRDs typically suffer from strong cross talk effects, which demands an extensive characterization of the detector [34] and advanced numerical postprocessing to correct for the imperfections [28,30]. Also, the multipixel detectors offer very limited reconfigurability and complicate channel balancing. Recent on-chip integration of independent on-off cryogenic detectors represents a promising direction [29,31,35], which has yet to be tested for various classical and, particularly, nonclassical sources.

The reported photon-number-resolving detector is based on spatial multiplexing of the input photonic signal by a reconfigurable optical network as depicted in Fig. 2(a). The multiport network consists of cascaded tunable beam splitters composed of a half-wave plate (HWP) and a polarizing beam splitter (PBS), which allow for accurate balancing of the output ports or, if needed, changing their number so there is no need to physically add or remove detectors. The whole network works as a one-to-M splitter balanced with the absolute error below 0.3%. To measure the multiplexed signal, we use single-photon avalanche photodiodes (SPADs) with efficiency close to 70%, 250 ps timing jitter, and 25 ns recovery time. The electronic outputs of the SPADs are summed by a custom coincidence logic while keeping the individual channels synchronized. Alternatively, the output can be visualized using an oscilloscope, see Fig. 2(a). Each of the resulting $M+1$ distinct voltage levels corresponds to the particular number of m-fold coincidences. Repeated measurements give rise to click statistics. Full technical details are given in the Supplemental Material [36], including a discussion of processing electronic signals from single-photon detectors.

It is important to stress here that the PNRD operates in real time and yields a result for every single input pulse with a latency (input-output delay) lower than 30 ns, including the response of the SPADs, which also allows its application as a communication receiver, quantum discrimination device, or for a feedback operation. The use of independent detectors and well-balanced coincidence circuitry completely removes any cross talk between the histogram channels, see Fig. 2(a). The effects of dark counts and afterpulses are virtually eliminated by operating the detector in the pulse regime with the repetition rate below 5 MHz. The period between individual measurement runs can be ultimately decreased to be only slightly longer than the recovery time of the constituent single-photon detectors, provided that afterpulsing is low enough. Furthermore, differences in SPAD efficiencies and other optical imperfections or imbalances of the PNRD can be arbitrarily compensated by adjusting the splitting ratios of the optical network. The result is a balanced multiplex with an overall efficiency η. This means that all systematic errors are eliminated either by design or by a sufficiently precise adjustment, independent of constituent detectors employed.

For a balanced M-channel PNRD with efficiency η, the probability of m channels clicking upon the arrival of n photons is

$$C_{mn} = \binom{M}{m} \sum_{j=0}^{m} (-1)^j \binom{m}{j} \left((1-\eta) + \frac{(m-j)\eta}{M} \right)^n. \tag{1}$$

The click statistics c_m is then determined by the photon statistics p_n [19,20,23],

$$c_m = \sum_n C_{mn} p_n. \tag{2}$$

Finding the photon statistics p_n, $n = 0, \ldots, \infty$, that satisfies the system of equations (2) for a measured click statistics c_m, $m = 0, \ldots, M$, represents the core problem of photon-statistics retrieval. This generally ill-posed problem suffers from underdetermination and sampling error. Fortunately, we have additional constraints facilitating the retrieval; i.e., the photon-number probabilities are non-negative, normalized, and typically non-negligible only within a finite range.

Here we present a novel approach, termed the expectation-maximization-entropy (EME) method, based on an expectation-maximization iterative algorithm weakly regularized by a maximum-entropy principle. The initial zeroth iteration is uniform; $p_n^{(0)} = 1/(n_{\max} + 1)$ for sufficiently large $n_{\max} \gg \langle n \rangle$. Each subsequent iteration is

$$p_n^{(k+1)} = \Pi_n^{(k)} p_n^{(k)} - \lambda (\ln p_n^{(k)} - S^{(k)}) p_n^{(k)}, \tag{3}$$

$$\Pi_n^{(k)} = \sum_{m=0}^{M} \frac{c_m}{(\sum_j C_{mj} p_j^{(k)})} C_{mn}, \qquad S^{(k)} = \sum_{n=0}^{n_{\max}} p_n^{(k)} \ln p_n^{(k)}. \tag{4}$$

Here the superscript (k) denotes the k th iteration. Each iteration is evaluated for $n = 1, \ldots, n_{\max}$. The $\Pi_n^{(k)}$ is a function of the measured click statistics c_m and the efficiency η determined by a separate measurement. $S^{(k)}$ is a negative von Neumann entropy. The parameter λ scales the entropy regularization relative to the likelihood maximization; we use a fixed value of 10^{-3} for all the photon statistics. The process is stopped when two subsequent iterations are practically identical. The retrieved statistics does not change for different initial iterations. The derivation of the algorithm is given in the Supplemental Material [36].

To show the accuracy and the robustness of the novel EME method, a numerical analysis was performed for 25 various photon statistics with different mean photon numbers. We compared the EME method with other frequently used algorithms—direct inversion and the expectation-maximization (EM) method based on likelihood maximization. EME was found to be a unique estimator that guarantees non-negativity and the absence of numerical artifacts in the retrieved photon statistics. Total variation distance $\Delta = \sum_n |p_n - p_n^{\text{true}}|/2$ between the retrieved distribution and the true one is on the order of $\sim 10^{-3}$, one order of magnitude smaller than in the case of direct inversion and maximum-likelihood approaches. Numerical simulations yield average fidelity values $\bar{F} = 0.9996$ using the EME algorithm and $\bar{F} = 0.997$ using the maximum-likelihood approach. The fidelity, defined as $F = (\sum_{n=0}^{n_{\max}} \sqrt{p_n p_n^{\text{true}}})^2$, cannot be evaluated for direct inversion due to negative values of estimated photon statistics.

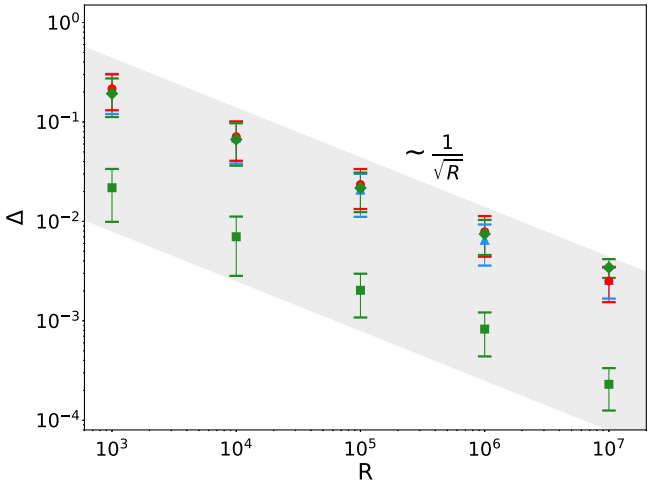


FIG. 3. A numerical analysis of EME total variation distance Δ. With more measurement runs R, the statistical error in the data is lower and the EME result approaches the true photon statistics despite the limited number of channels $M = 10$. Here shown are coherent state with $\langle n \rangle = 10$ (blue triangle up), thermal state with $\langle n \rangle = 5$ (red circle), N_p-photon cluster with $N_p = 1$ (green square), and $N_p = 9$ (green rhombus) for a single value of $\lambda = 10^{-3}$. The gray area illustrates the observed scaling $(0.25/\sqrt{R}–14/\sqrt{R}.)$

In Fig. 3, we show by numerical simulation that the results of the EME algorithm approach the respective theoretical expectations as more data are acquired. This means that, despite a limited number of channels $M = 10$, the chief source of error is the sampling error. We also verified that Δ stays the same if both the mean number of photons and the number of channels are doubled. Therefore, EME scales well to high photon numbers considering limited experimental resources. The precision of the photon-statistics retrieval can be further increased by optimizing over multiple parameters, such as M, λ, $n_{\max}$, or iteration cutoff. Eventually, Δ becomes limited by machine precision and computation time. The analysis of the complex interaction of these parameters will be the subject of further research. We also found that the EME convergence is 10–1000× faster than the plain EM approach (see

TABLE I. The comparison of EM and EME results for the measured data. Coherent state $\langle n\rangle = 4.95(2)$, thermal state $\langle n\rangle = 4.93(4)$, two-photon-subtracted thermal state $\langle n\rangle = 5.77(2)$, single-photon state $N_p = 1$, and nine-photon cluster $N_p = 9$. Both fidelity F and total variation distance Δ are shown. Standard deviations are evaluated by repeating the measurement and data processing ten times. The large distances observed for EM stem from overfitting the ill-posed problem. This is discussed in the Supplemental Material [36].

	Coherent		Thermal		Two-photon-subtracted thermal		Single photon		Nine-photon cluster	
	EM	EME	EM	EME	EM	EME	EM	EME	EM	EME
F	0.6(1)	0.9984(9)	0.69(2)	0.9978(5)	0.90(1)	0.9990(4)	0.993 94(2)	0.999 12(1)	0.5467(2)	0.999 30(2)
Δ	0.50(9)	0.002(9)	0.35(1)	0.033(3)	0.21(1)	0.019(6)	0.074 24(1)	0.000 88(1)	0.1752(5)	0.004 07(1)

the Supplemental Material [36]), while yielding significantly better results.

In our experimental demonstration, we used a balanced ten-channel configuration of the detector. We analyzed coherent states, thermal states, multimode thermal states, single-photon and multiple-photon-subtracted thermal states, and nonclassical multiphoton states. Furthermore, we have varied the mean number of photons, the number of modes, and the number of subtracted or superimposed photons. For each retrieved photon statistics we computed $\langle n\rangle$, $g^{(2)}$, and other quantities presented in detail in the Supplemental Material [36].

The measurements were performed using 1-ns-long optical pulses with the repetition rate of 2 MHz. We prepared the initial coherent signal by using a gain-switched laser diode at 810 nm. The resulting coherent pulses measured by the PNRD show almost perfect Poissonian statistics with $g^{(2)} = 1$ up to $\langle n\rangle = 20$ with average fidelity $\bar{F} = 0.996$ and total variation distance $\bar{\Delta} = 24 \times 10^{-3}$. The thermal state is generated by temporal intensity modulation of the initial coherent light by a rotating ground glass. The scattered light is collected using a single-mode optical fiber. We measured almost ideal Bose-Einstein photon statistics depicted in Fig. 2(b) with $g^{(2)} = 2$ up to $\langle n\rangle = 5$, $\langle \Delta n^2\rangle = 30$ with $\bar{F} = 0.997$ and $\bar{\Delta} = 24 \times 10^{-3}$. We varied the number of the collected thermal modes, which yielded a signal governed by Mandel-Rice statistics, going from Bose-Einstein to Poisson distribution as the number of modes increased. Multiple-photon subtraction from the thermal state was implemented using a beam splitter with a 5% reflectance. When a (multi)coincidence was detected by a multichannel single-photon detector in the reflected port, the heralded optical signal in the transmitted port was analyzed by the reported PNRD. A typical result of two-photon subtraction is shown in Fig. 2(c). Increasing the number of subtracted photons results in a transition from super-Poissonian chaotic light to a Poissonian signal [90,91]. Furthermore, we generated multiphoton states by mixing incoherently several single-photon states from spontaneous parametric down-conversion using time multiplexing. N_p successive time windows, where a single photon was heralded, were merged into a single temporal detection mode. This source emulates the collective emission from identical independent single emitters [5,9,11]. The resulting photon statistics measured for these highly nonclassical multiphoton states corresponds extremely well to the ideal attenuated N_p-photon states, see Figs. 2(d) and 2(e) for $N_p = 1$ and 9 with $\bar{F} = 0.999$ and $\bar{\Delta} = 3 \times 10^{-3}$. Also the $g^{(2)}$ parameter computed from the measured photon statistics perfectly agrees with the theoretical model $1 - 1/N_p$, see Fig. 1.

We utilize fidelity and total variation distance to compare the measured distribution with the ideal one. The worst and the best fidelities $F = 0.985$ and 0.9999 are reached across all the tested sources with average fidelity being $\bar{F} = 0.998$. The average distance is $\bar{\Delta} = 17 \times 10^{-3}$ for all the tested sources. For detailed data and comparison to plain EM, see Table I and the Supplemental Material [36]. The errors of EME are caused by slight imbalances of splitting ratios in the PNRD, variations in PNRD efficiency η, and imperfections of the tested sources, which renders the actual accuracy of the PNRD even higher. Particularly, accurate preparation and characterization of thermal and superchaotic states are highly nontrivial tasks subject to ongoing research [13,92–94].

To conclude, we have reported a fully reconfigurable near-ideal photon-number-resolving detection scheme with custom electronic processing and a novel EME photon-statistics retrieval method. The PNRD design is free of systematic errors, which are either negligible or can be arbitrarily decreased by the user. We have demonstrated exceptional accuracy of detected photon statistics that goes beyond the conventional limit of the number of PNRD channels. We measured dozens of various photonic sources ranging from highly nonclassical quantum states of light to chaotic optical signals. The results were obtained from raw data with no other processing than EME and without any demanding detector characterization. Despite uncorrected systematic errors and significant variability of the input signal, our approach shows superior fidelity across the board with typical values exceeding 99.8% for mean photon numbers up to 20 and the $g^{(2)}$ parameter reaching down to a fraction of a percent. Though having been demonstrated with common single-photon avalanche diodes, the reported measurement workflow is independent of the detection technology and can accommodate any on-off

detectors. Furthermore, the multichannel scheme allows for straightforward on-chip integration. Therefore, further improvements in speed, efficiency, and compactness can be expected using superconducting single-photon detectors [26,29,31] coupled with waveguide technology [95–97].

We acknowledge the support from the Czech Science Foundation under the Project No. 17-26143S. This work has received national funding from the MEYS and the funding from European Union's Horizon 2020 research and innovation framework programme under Grant Agreement No. 731473 (project 8C18002). Project HYPER-U-P-S has received funding from the QuantERA ERA-NET Cofund in Quantum Technologies implemented within the European Union's Horizon 2020 Programme. We also acknowledge the support from Palacký University (Projects No. IGA-PrF-2018-010 and No. IGA-PrF-2019-010).

*jezek@optics.upol.cz

[1] J. L. O'Brien, A. Furusawa, and J. Vučković, Nat. Photonics **3**, 687 (2009).
[2] J. C. F. Matthews, X.-Q. Zhou, H. Cable, P. J. Shadbolt, D. J. Saunders, G. A. Durkin, G. J. Pryde, and J. L. O'Brien, npj Quantum Inf. **2**, 16023 (2016).
[3] S. Slussarenko, M. M. Weston, H. M. Chrzanowski, L. K. Shalm, V. B. Verma, S. W. Nam, and G. J. Pryde, Nat. Photonics **11**, 700 (2017).
[4] G. Harder, T. J. Bartley, A. E. Lita, S. W. Nam, T. Gerrits, and C. Silberhorn, Phys. Rev. Lett. **116**, 143601 (2016).
[5] I. Straka, L. Lachman, J. Hloušek, M. Miková, M. Ježek, and R. Filip, npj Quantum Inf. **4**, 4 (2018).
[6] J. F. Dynes, M. Lucamarini, K. A. Patel, A. W. Sharpe, M. B. Ward, Z. L. Yuan, and A. J. Shields, Opt. Express **26**, 22733 (2018).
[7] M. D. Vidrighin, O. Dahlsten, M. Barbieri, M. S. Kim, V. Vedral, and I. A. Walmsley, Phys. Rev. Lett. **116**, 050401 (2016).
[8] N. Cottet, S. Jezouin, L. Bretheau, P. Campagne-Ibarcq, Q. Ficheux, J. Anders, A. Auffèves, R. Azouit, P. Rouchon, and B. Huard, Proc. Natl. Acad. Sci. U.S.A. **114**, 7561 (2017).
[9] H. Ta, J. Keller, M. Haltmeier, S. K. Saka, J. Schmied, F. Opazo, P. Tinnefeld, A. Munk, and S. W. Hell, Nat. Commun. **6**, 7977 (2015).
[10] Y. Israel, R. Tenne, D. Oron, and Y. Silberberg, Nat. Commun. **8**, 14786 (2017).
[11] S. J. Sahl, S. W. Hell, and S. Jakobs, Nat. Rev. Mol. Cell Biol. **18**, 685 (2017).
[12] H. J. Kimble, M. Dagenais, and L. Mandel, Phys. Rev. Lett. **39**, 691 (1977).
[13] H. E. Kondakci, A. Szameit, A. F. Abouraddy, D. N. Christodoulides, and B. E. A. Saleh, Optica **3**, 477 (2016).
[14] The autocorrelation $g^{(2)}$ stems from the second-order optical coherence evaluated as a cross-correlation of intensity I at zero delay, $g^{(2)}(0) = \langle I(t)^2\rangle/\langle I(t)\rangle^2$. For the measured signals, we calculate its value from the obtained distribution of the number of photons n, $g^{(2)} = \langle n(n-1)\rangle/\langle n\rangle^2$.
[15] K. Stensson and G. Björk, Phys. Rev. A **98**, 033812 (2018).
[16] J. Sperling, W. R. Clements, A. Eckstein, M. Moore, J. J. Renema, W. S. Kolthammer, S. W. Nam, A. Lita, T. Gerrits, W. Vogel, G. S. Agarwal, and I. A. Walmsley, Phys. Rev. Lett. **118**, 163602 (2017).
[17] T. Gerrits, A. Lita, B. Calkins, and S. W. Nam, in *Quantum Science and Technology* (Springer International Publishing, New York, 2016), pp. 31–60.
[18] C. Cahall, K. L. Nicolich, N. T. Islam, G. P. Lafyatis, A. J. Miller, D. J. Gauthier, and J. Kim, Optica **4**, 1534 (2017).
[19] H. Paul, P. Törmä, T. Kiss, and I. Jex, Phys. Rev. Lett. **76**, 2464 (1996).
[20] J. Sperling, W. Vogel, and G. S. Agarwal, Phys. Rev. A **85**, 023820 (2012).
[21] K. Banaszek and I. A. Walmsley, Opt. Lett. **28**, 52 (2003).
[22] D. Achilles, C. Silberhorn, C. Śliwa, K. Banaszek, and I. A. Walmsley, Opt. Lett. **28**, 2387 (2003).
[23] M. J. Fitch, B. C. Jacobs, T. B. Pittman, and J. D. Franson, Phys. Rev. A **68**, 043814 (2003).
[24] J. Řeháček, Z. Hradil, O. Haderka, J. J. Peřina, and M. Hamar, Phys. Rev. A **67**, 061801(R) (2003).
[25] J. Tiedau, E. Meyer-Scott, T. Nitsche, S. Barkhofen, T. J. Bartley, and C. Silberhorn, Opt. Express **27**, 1 (2019).
[26] A. Divochiy, F. Marsili, D. Bitauld, A. Gaggero, R. Leoni, F. Mattioli, A. Korneev, V. Seleznev, N. Kaurova, O. Minaeva, G. Gol'tsman, K. G. Lagoudakis, M. Benkhaoul, F. Lévy, and A. Fiore, Nat. Photonics **2**, 302 (2008).
[27] A. Allevi, M. Bondani, and A. Andreoni, Opt. Lett. **35**, 1707 (2010).
[28] D. A. Kalashnikov, S. H. Tan, M. Chekhova, and L. A. Krivitsky, Opt. Express **19**, 9352 (2011).
[29] F. Mattioli, Z. Zhou, A. Gaggero, R. Gaudio, R. Leoni, and A. Fiore, Opt. Express **24**, 9067 (2016).
[30] J. Krger, T. Ahrens, J. Sperling, W. Vogel, H. Stolz, and B. Hage, J. Phys. B **50**, 214003 (2017).
[31] D. Zhu, Q.-Y. Zhao, H. Choi, T.-J. Lu, A. E. Dane, D. Englund, and K. K. Berggren, Nat. Nanotechnol. **13**, 596 (2018).
[32] B. Calkins, P. L. Mennea, A. E. Lita, B. J. Metcalf, W. S. Kolthammer, A. Lamas-Linares, J. B. Spring, P. C. Humphreys, R. P. Mirin, J. C. Gates, P. G. R. Smith, I. A. Walmsley, T. Gerrits, and S. W. Nam, Opt. Express **21**, 22657 (2013).
[33] M. Mičuda, O. Haderka, and M. Ježek, Phys. Rev. A **78**, 025804 (2008).
[34] J. S. Lundeen, A. Feito, H. Coldenstrodt-Ronge, K. L. Pregnell, Ch. Silberhorn, T. C. Ralph, J. Eisert, M. B. Plenio, and I. A. Walmsley, Nat. Phys. **5**, 27 (2009).
[35] P. Rath, O. Kahl, S. Ferrari, F. Sproll, G. Lewes-Malandrakis, D. Brink, K. Ilin, M. Siegel, C. Nebel, and W. Pernice, Light Sci. Appl. **4**, e338 (2015).
[36] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.123.153604 for full details about the experimental setup, the verification of its operation, and the comparison of photon-statistic retrieval methods, which includes Refs. [37–89].
[37] L. Mandel, Opt. Lett. **4**, 205 (1979).
[38] E. J. O'Reilly and A. Olaya-Castro, Nat. Commun. **5**, 3012 (2014).
[39] P. Grangier, G. Roger, and A. Aspect, Europhys. Lett. **1**, 173 (1986).

[40] L. Lachman and R. Filip, Opt. Express **24**, 27352 (2016).
[41] M. G. Genoni, M. L. Palma, T. Tufarelli, S. Olivares, M. S. Kim, and M. G. A. Paris, Phys. Rev. A **87**, 062104 (2013).
[42] J. Park, J. Zhang, J. Lee, S.-W. Ji, M. Um, D. Lv, K. Kim, and H. Nha, Phys. Rev. Lett. **114**, 190402 (2015).
[43] J. Park, Y. Lu, J. Lee, Y. Shen, K. Zhang, S. Zhang, M. S. Zubairy, K. Kim, and H. Nha, Proc. Natl. Acad. Sci. U.S.A. **114**, 891 (2017).
[44] L. Happ, M. A. Efremov, H. Nha, and W. P. Schleich, New J. Phys. **20**, 023046 (2018).
[45] I. Straka, A. Predojević, T. Huber, L. Lachman, L. Butschek, M. Miková, M. Mičuda, G. S. Solomon, G. Weihs, M. Ježek, and R. Filip, Phys. Rev. Lett. **113**, 223603 (2014).
[46] A. Mandarino, M. Bina, C. Porto, S. Cialdi, S. Olivares, and M. G. A. Paris, Phys. Rev. A **93**, 062118 (2016).
[47] M. A. Wayne, J. C. Bienfang, and S. V. Polyakov, Opt. Express **25**, 20352 (2017).
[48] A. W. Ziarkash, S. K. Joshi, M. Stipčević, and R. Ursin, Sci. Rep. **8**, 5076 (2018).
[49] M. Fujiwara, A. Tanaka, S. Takahashi, K. Yoshino, Y. Nambu, A. Tajima, S. Miki, T. Yamashita, Z. Wang, A. Tomita, and M. Sasaki, Opt. Express **19**, 19562 (2011).
[50] V. Burenkov, H. Xu, B. Qi, R. H. Hadfield, and H.-K. Lo, J. Appl. Phys. **113**, 213102 (2013).
[51] W. Martienssen and E. Spiller, Am. J. Phys. **32**, 919 (1964).
[52] I. Straka, J. Mika, and M. Ježek, Opt. Express **26**, 8998 (2018).
[53] J. Hloušek, M. Ježek, and R. Filip, Sci. Rep. **7**, 13046 (2017).
[54] Yu. I. Bogdanov, K. G. Katamadze, G. V. Avosopiants, L. V. Belinsky, N. A. Bogdanova, A. A. Kalinkin, and S. P. Kulik, Phys. Rev. A **96**, 063803 (2017).
[55] S. M. Barnett, G. Ferenczi, C. R. Gilson, and F. C. Speirits, Phys. Rev. A **98**, 013809 (2018).
[56] I. A. Fedorov, A. E. Ulanov, Y. V. Kurochkin, and A. I. Lvovsky, Optica **2**, 112 (2015).
[57] K. G. Katamadze, G. V. Avosopiants, Y. Bogdanov, and S. P. Kulik, Optica **5**, 723 (2018).
[58] L. Qi, M. Manceau, A. Cavanna, F. Gumpert, L. Carbone, M. de Vittorio, A. Bramati, E. Giacobino, L. Lachman, R. Filip, and M. Chekhova, New J. Phys. **20**, 073013 (2018).
[59] L. Lachman, I. Straka, J. Hloušek, M. Ježek, and R. Filip, Phys. Rev. Lett. **123**, 043601 (2019).
[60] A. Luis and L. L. Sánchez-Soto, Phys. Rev. Lett. **83**, 3573 (1999).
[61] J. Fiurášek, Phys. Rev. A **64**, 024102 (2001).
[62] J. Řeháček, D. Mogilevtsev, and Z. Hradil, Phys. Rev. Lett. **105**, 010402 (2010).
[63] C. M. Natarajan, L. Zhang, H. Coldenstrodt-Ronge, G. Donati, S. N. Dorenbos, V. Zwiller, I. A. Walmsley, and R. H. Hadfield, Opt. Express **21**, 893 (2013).
[64] M. Altorio, M. G. Genoni, F. Somma, and M. Barbieri, Phys. Rev. Lett. **116**, 100802 (2016).
[65] M. Cooper, M. Karpiński, and B. J. Smith, Nat. Commun. **5**, 4332 (2014).
[66] C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and Its Applications* (Wiley, New York, 1971).
[67] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses—Theory and Applications* (Springer-Verlag, Berlin, 2003).
[68] A. N. Tikhonov, Dokl. Akad. Nauk SSSR **39**, 195 (1943).
[69] A. E. Hoerl, Chem. Eng. Prog. **58**, 54 (1962).
[70] A. N. Tikhonov, A. Goncharsky, V. V. Stepanov, and A. G. Yagola, *Numerical Methods for the Solution of Ill-Posed Problems* (Springer-Verlag, Berlin, 1995).
[71] V. N. Starkov, A. A. Semenov, and H. V. Gomonay, Phys. Rev. A **80**, 013813 (2009).
[72] A. P. Dempster, N. M. Laird, and D. B. Rubin, J. R. Stat. Soc. Ser. B **39**, 1 (1977).
[73] Y. Vardi and D. Lee, J. R. Stat. Soc. Ser. B **55**, 569 (1993).
[74] K. Banaszek, Phys. Rev. A **57**, 5013 (1998).
[75] F. Marsili, D. Bitauld, A. Gaggero, S. Jahanmirinejad, R. Leoni, F. Mattioli, and A. Fiore, New J. Phys. **11**, 045022 (2009).
[76] K. Banaszek, Acta Phys. Slovaca **48**, 185 (1998).
[77] G. Zambra, A. Andreoni, M. Bondani, M. Gramegna, M. Genovese, G. Brida, A. Rossi, and M. G. A. Paris, Phys. Rev. Lett. **95**, 063602 (2005).
[78] *Maximum-Entropy and Bayesian Methods in Inverse Problems*, edited by C. R. Smith and W. T. Grandy (Springer, Netherlands, 1985).
[79] R. L. Parker, *Geophysical Inverse Theory* (Princeton University Press, Princeton, NJ, 1994).
[80] Y. Grandvalet and Y. Bengio, in *Advances in Neural Information Processing Systems 17*, edited by L. K. Saul, Y. Weiss, and L. Bottou (MIT Press, Cambridge, MA, 2005), pp. 529–536.
[81] Y. S. Teo, H. Zhu, B.-G. Englert, J. Řeháček, and Z. Hradil, Phys. Rev. Lett. **107**, 020404 (2011).
[82] A. I. Lvovsky and M. G. Raymer, Rev. Mod. Phys. **81**, 299 (2009).
[83] M. A. Golub, Opt. Photonics News **15**, 36 (2004).
[84] A. Hermerschmidt, S. Krüger, and G. Wernicke, Opt. Lett. **32**, 448 (2007).
[85] M. M. Rajadhyaksha and R. H. Webb, Appl. Opt. **34**, 8066 (1995).
[86] J. C. F. Matthews, A. Politi, A. Stefanov, and J. L. O'Brien, Nat. Photonics **3**, 346 (2009).
[87] B. J. Smith, D. Kundys, N. Thomas-Peter, P. G. R. Smith, and I. A. Walmsley, Opt. Express **17**, 13516 (2009).
[88] F. Flamini, L. Magrini, A. S. Rab, N. Spagnolo, V. D'Ambrosio, P. Mataloni, F. Sciarrino, T. Zandrini, A. Crespi, R. Ramponi, and R. Osellame, Light Sci. Appl. **4**, e354 (2015).
[89] J. Carolan, C. Harrold, C. Sparrow, E. Martin-Lopez, N. J. Russell, J. W. Silverstone, P. J. Shadbolt, N. Matsuda, M. Oguma, M. Itoh, G. D. Marshall, M. G. Thompson, J. C. F. Matthews, T. Hashimoto, J. L. O'Brien, and A. Laing, Science **349**, 711 (2015).
[90] A. Allevi, A. Andreoni, M. Bondani, M. G. Genoni, and S. Olivares, Phys. Rev. A **82**, 013816 (2010).
[91] Y. Zhai, F. E. Becerra, B. L. Glebov, J. Wen, A. E. Lita, B. Calkins, T. Gerrits, J. Fan, S. W. Nam, and A. Migdall, Opt. Lett. **38**, 2171 (2013).
[92] A. Allevi and M. Bondani, Opt. Lett. **40**, 3089 (2015).
[93] J. Mika, L. Podhora, L. Lachman, P. Obšil, J. Hloušek, M. Ježek, R. Filip, and L. Slodička, New J. Phys. **20**, 093002 (2018).

[94] M. Manceau, K. Y. Spasibko, G. Leuchs, R. Filip, and M. Chekhova, Phys. Rev. Lett. **123**, 123606 (2019).
[95] D. Sahin, A. Gaggero, Z. Zhou, S. Jahanmirinejad, F. Mattioli, R. Leoni, J. Beetz, M. Lermer, M. Kamp, S. Höfling, and A. Fiore, Appl. Phys. Lett. **103**, 111116 (2013).
[96] R. Heilmann, J. Sperling, A. Perez-Leija, M. Gräfe, M. Heinrich, S. Nolte, W. Vogel, and A. Szameit, Sci. Rep. **6**, 19489 (2016).
[97] J. Münzberg, A. Vetter, F. Beutel, W. Hartmann, S. Ferrari, W. H. P. Pernice, and C. Rockstuhl, Optica **5**, 658 (2018).

High-resolution coincidence counting system for large-scale photonics applications

Josef Hloušek,* Jan Grygar, Michal Dudka, and Miroslav Ježek†
Department of Optics, Faculty of Science, Palacký University, 17 listopadu 12, Olomouc 77900, Czech Republic

(Received 2 November 2023; revised 12 January 2024; accepted 16 January 2024; published 12 February 2024)

The increasing complexity of recent photonic experiments challenges the development of efficient multichannel coincidence counting systems with high-level functionality. Here, we report a coincidence unit able to count detection events ranging from single to 16-fold coincidences with full channel-number resolution. The device operates within sub-100-ps coincidence time windows, with a maximum input frequency of 1.5 GHz and an overall jitter of less than 10 ps. The unit high-level timing performance renders it suitable for quantum photonic experiments employing low-timing-jitter single-photon detectors. Additionally, the unit can be used in complex photonic systems to drive feed-forward loops. We demonstrate the developed coincidence counting unit in photon-number-resolving detection to directly quantify the statistical properties of light, specifically coherent and thermal states, with a fidelity exceeding 0.999 up to 60 photons.

DOI: 10.1103/PhysRevApplied.21.024023

I. INTRODUCTION

A coincidence counting unit (CCU) is an essential tool widely employed in all applications requiring the detection of a large number of photons (or other particles) and processing of the detected signals. Modern quantum experiments employ increasingly complex systems with a growing number of input and output channels [1–3]. These large-scale photonics systems hinge on the ability to generate, control, and analyze the multiphoton quantum states [1,4–6] frequently used in quantum communications [7,8], quantum computation, and simulations [9]. Particularly, complex coincidence processing has become an integral part of measuring unknown optical states by photon-number-resolving detectors based on multiplexing [10–12]. Detected statistical properties of light are routinely applied to quantify nonclassicality and quantum non-Gaussianity [13–17]. Furthermore, the advanced functionality of these devices is of considerable interest in on-the-fly multifold coincidence analysis to control large-scale quantum systems via feed-forward operation and photon-number-resolving postselection [18,19].

Conventional approaches to detect coincidences are (1) time-to-amplitude converter (TAC) together with a single- or multichannel analyzer, (2) time-to-digital converter (TDC) followed by postprocessing, and (3) overlap logic coincidence systems realized with discrete components or using a field-programmable gate array (FPGA). TACs and TDCs both typically offer tens-of-picosecond resolution. TACs are not easily scaled up for multicoincidence systems and possess a considerable dead time limiting rate throughput to tens of thousands of events per second [20,21]. TDC-based solutions stream time tags to a computer for further processing. Therefore, a large amount of data are processed offline. Alternatively, TDCs are combined with an FPGA for subsequent processing [22]. Pulse overlap coincidence systems use fast logic gates and multiplexers to capture detection events and detect coincidences [21]. Coincidence counting and histogramming could be programmed into a microcontroller [23] or FPGA [24]. Functional blocks such as internal delay lines, coincidence counters, and a processor can all be integrated within a single FPGA chip [25–28]. Lately, multichannel TDC-based coincidence counter architecture in the same FPGA chip was introduced [29].

In this paper, we report an ultrafast electronic multichannel CCU, producing a histogram of all possible coincidence events for up to 16 constituent detectors. The device performs a real-time classification of all possible detection events in a 2^{16}-element histogram with a rate of up to 3 million events per second. We have conducted a comprehensive characterization of the presented CCU, revealing excellent performance parameters, including a sub-100-ps coincidence window, sub-10-ps jitter, and an ultralow coincidence error probability. Furthermore, we demonstrate the CCU versatility by implementing a multiphoton counting experiment to fully characterize the statistical properties of incident light.

*hlousek@optics.upol.cz
†jezek@optics.upol.cz

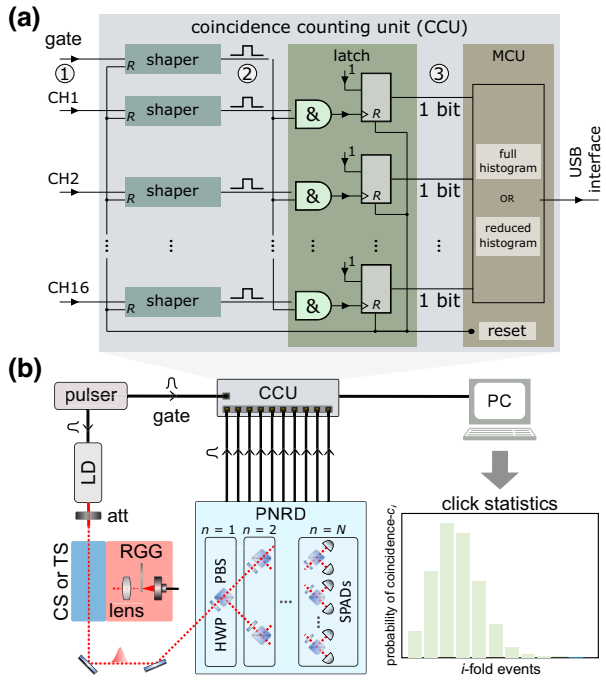


FIG. 1. (a) A block diagram of the CCU: the shapers prepare square pulses; coincidences are evaluated as an overlap (&) of the pulses and can be directly used or counted and sorted in histograms by the microcontroller unit (MCU). (b) For photon-number-resolved detection (PNRD), the CCU is connected to a spatial-multiplexed optical network consisting of tunable beam splitters and SPADs. The resulting coincidence histogram allows for statistical analysis of the incident signal.

II. COINCIDENCE COUNTING UNIT

Figure 1(a) shows an overview of the CCU architecture based on fast positive emitter-coupled logic (ECL) offering high-resolution coincidence counting. The CCU consists of input signal overlap logic and data processing unit. The device accepts 16 data inputs and a single gate input and yields the complete histogram of 2^{16} multicoincidences of the inputs within the gate signal. Each input channel contains a shaping circuit with high-resolution programmable delay lines to detect an input signal edge and provides the output ECL pulse of a given width and delay. The pulse width corresponds to half of the coincidence window, which can be tuned independently for each channel. After the signal shaping, the signal and gate pulse enable inputs of ECL latch circuitry. The latch records the overlap of individual signal pulses and the gate signal and stores information about coincidence events and single-channel pulses. Data from the latch are transferred in the form of bits. A detailed description of the pulse shaping and processing is presented in Appendix A.

The outputs of the latch can be directly employed for implementing real-time feed-forward control. For example, one can switch between states of an electro-optic modulator to modify the performance of the following experimental setup. The maximum input frequency is limited to 1.5 GHz by the maximum operating frequency of the delay lines. When a higher voltage is required to control the subsequent devices, it is necessary to utilize a logic-level translator. In this work, we use a microcontroller unit for controlling and monitoring overlap logic and data processing and storage. Consequently, an ECL-CMOS translator is required, reducing the maximum input frequency to 800 MHz.

The presented CCU allows the operation of two specific regimes depending on the complexity of the measurement. One can store the complete information about all possible coincidence events ($2^{16} = 65\,536$ kinds of coincidence events in total), termed the full histogram regime. In the second regime, the CCU counts only coincidences of the same order without information about the channel number (the reduced histogram regime). For our specific technical solution, the maximum processing rate is about 3 million events per second for the reduced histogram and 2 million events for the full histogram. This rate can be further enhanced by utilizing an FPGA.

Another purpose of the microcontroller unit is to configure the delay lines in the shapers to synchronize the input signals and set the length of the coincidence window. It also enables self-calibration of the CCU and monitoring of the operating temperature and other parameters. The minimum coincidence window can be set below 100 ps, and the maximum coincidence window length can reach 20 ns. The detailed electronic characterization of the CCU is presented in Appendix B. A comparison with the state-of-the-art approaches is given in Appendix C.

III. PHOTON STATISTICS MEASUREMENT OF LARGE OPTICAL STATES

Numerous technological approaches have been developed and experimentally verified to achieve photon-number resolution, falling into two categories: inherent photon-number resolution (superconducting nanowire single-photon detectors [30–32] or transition edge sensors [33–38]) and multiplex and/or multipixel detection schemes [10–17,39–51]. In our case, the multiplexed detector consists of tunable beam splitters composed of a half-wave plate and a polarizing beam splitter; see Fig. 1(b). This beam-splitting approach allows accurate adjustments of the splitting ratio with an absolute error below 0.3%. Each channel is coupled to a multimode fiber and brought to a single-photon avalanche diode (SPAD) with efficiency ranging from 55% to 70% at 0.8 μm, 200–300 ps timing jitter, and 20–30 ns dead time. The total detection efficiency η is defined as the ratio of the total number of detected photons to the total number of incident photons. Based on the measured transmittance of the multiport optical network and SPAD efficiencies,

we experimentally determined η as 50(1)%. Furthermore, utilizing low-loss optics and superconducting nanowire single-photon detectors [52] can enhance global efficiency to over 85%.

As a light source, we use a gain-switched semiconductor laser diode to generate a coherent nanosecond pulsed light with a central wavelength of 0.8 μm. The laser diode is driven by nanosecond electronic pulses at a repetition rate of 1 MHz. We use the temporal intensity modulation of the initial coherent light by rotating ground glass to generate pseudothermal light with Bose-Einstein distribution [53]. The optical signal is collected by a single-mode fiber to produce a single-mode thermal state.

We quantify the statistical properties of light, focusing on coherent and thermal states, across a wide range of mean photon numbers. These states of light are detected by a ten-channel multiplexed detector and processed by the CCU in the reduced histogram regime. The coincidence window width is set to be significantly larger than both the optical pulse width and the detector jitter. Due to nonunity detection efficiency, noise, and a finite number of single-photon detection channels, we observe the probability distribution of the coincidence events (i.e., click statistics) c_m instead of the photon statistics p_n. However, the click statistics still carry information about the character of the initial state of light. The parameter quantifying this phenomenon is called the binomial parameter [11,54] defined as $Q_b = \frac{\langle(\Delta c)^2\rangle}{\frac{\langle c\rangle}{N}\left(1-\frac{\langle c\rangle}{N}\right)} - 1$, where $\langle c\rangle = \sum_{i=0}^{N} i c_i$, $\langle(\Delta c)^2\rangle = \sum_{i=0}^{N}(i-\langle c\rangle)^2 c_i$, and N stands for the number of detection channels. States of light with a Poisson distribution result in a binomial parameter of $Q_b = 0$, whereas non-negative values occur for super-Poissonian states (see Fig. 2). For coherent states, we measure the binomial parameters that vary between $1(5) \times 10^{-4}$ and $1.0(5) \times 10^{-3}$, with a mean number of clicks $\langle c\rangle$ ranging from 0.005 to 5. The values of the binomial parameter do not exactly match zero due to the higher variance in the measured click statistics caused by excess noise in the light source. Measured thermal states cover values of mean number of clicks $\langle c\rangle$ ranging from 0.002 to 3 with Q_b that encompasses the range from $5.1(7) \times 10^{-3}$ to 2.060(1). All measured values exhibit excellent agreement with the theoretical predictions.

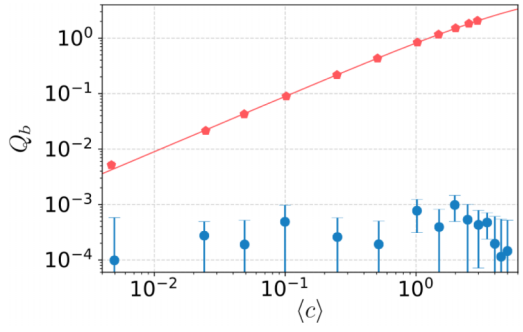


FIG. 2. The binomial parameter Q_b as a function of the mean number of clicks $\langle c\rangle$. Shown are the coherent states (blue circles) and thermal states (red pentagons); the solid line represents the theoretical model for thermal state.

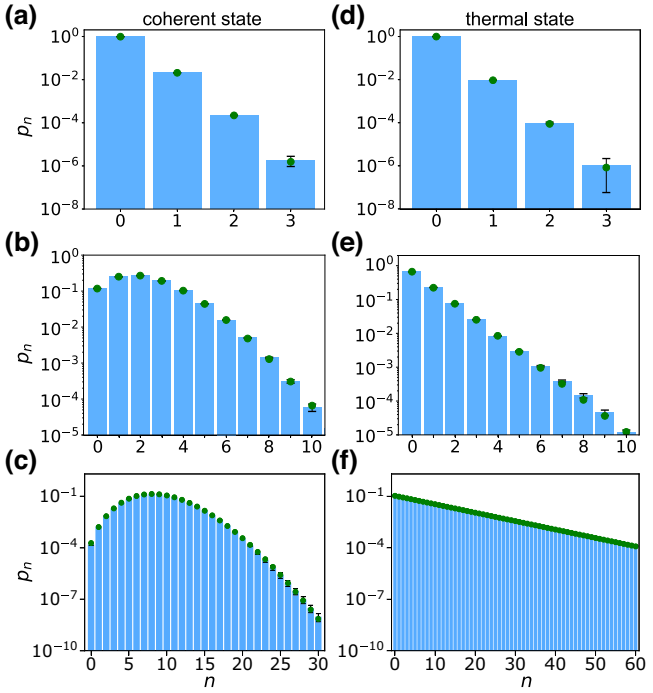


FIG. 3. Retrieved (blue bars) and theoretical (green circles) photon statistics of coherent state (a)–(c) and thermal state (d)–(f) for several mean photon numbers: (a) $\langle n\rangle_{CS} = 0.02119(3)$, (b) $\langle n\rangle_{CS} = 2.144(3)$, (c) $\langle n\rangle_{CS} = 10.21(1)$, (d) $\langle n\rangle_{TS} = 0.00949(5)$, (e) $\langle n\rangle_{TS} = 0.5113(4)$, and (f) $\langle n\rangle_{TS} = 8.41(3)$.

In addition to directly measuring the click statistics, we have also retrieved photon statistics. Here, we demonstrate the faithful photon statistics reconstruction over three orders of magnitude of the mean photon numbers within a dynamic range of up to 60 photons (see Fig. 3). Results show unprecedented agreement between theoretical distributions (green circles) and observed data (blue bars) down to probabilities of 10^{-8}. It is important to stress here that the photon statistics retrieval includes all imperfections, such as slight imbalances of splitting ratios in the photon-number-resolving detector and imperfect light-state preparation. The fidelity of the retrieved photon statistics, defined as $\mathcal{F} = \left(\sum_{n=0}^{n_{max}} \sqrt{p_n p_n^{ideal}}\right)^2$, surpasses 0.999 for all measured states of light [see Fig. 4(a)]. In fact, coherent states exhibit a fidelity exceeding 0.9997. As the mean photon number increases, the retrieval process degrades. This is a result of the increased probability of higher Fock states within the analyzed state, leading to a progressive divergence from the detector dynamic range.

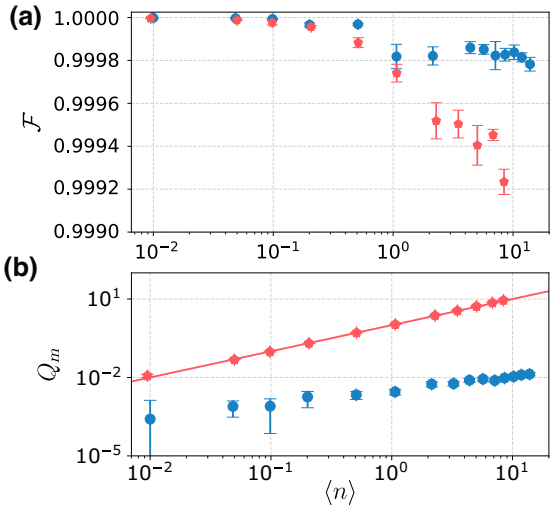


FIG. 4. Quantification of the retrieved photon statistics p_n: (a) fidelity $\mathcal{F}$ and (b) Mandel parameter Q_m as a function of mean photon number $\langle n \rangle$. Shown are the coherent states (blue circles) and thermal states (red pentagons); the solid line represents the theoretical model for thermal state.

The extent of measurement degradation by the reconstruction error naturally depends on the photon statistics of the initial light state.

The presented measurement workflow is scalable with the total number of detection channels. Increasing their quantity can expand the detector dynamic range, enabling the measurement of optical states with even higher intensities. Also, the presented reconfigurable detection network allows for a decrease in the number of employed detectors to measure specific characteristics of the source under test. Specifically, this enables a direct evaluation of the anticorrelation parameter α [55], nonclassicality [14,15,17,54], and quantum non-Gaussianity [16,56].

With a knowledge of the photon statistics, we explore the Mandel parameter [13,57,58], which is a convenient way to characterize deviation from Poisson statistics. The Mandel parameter is defined as the ratio of the second and first moments of photon statistic distribution: $Q_m = (\langle(\Delta n)^2\rangle - \langle n\rangle)/\langle n\rangle$. We calculate the Mandel parameter of the retrieved photon statistics as a function of the mean photon number of the initial state; see Fig. 4(b). We measure coherent states with Poisson statistics with the Mandel parameter Q_m ranging from $3(5) \times 10^{-4}$ to $1.3(2) \times 10^{-2}$ and the mean photon number reaching $\langle n\rangle = 19.84(2)$. The ideal Poisson distribution with photon number variance $\langle(\Delta n)^2\rangle = \langle n\rangle$ reaches the Mandel parameter $Q_m = 0$. The measured data are influenced by excess noise, leading to a slight offset from the anticipated value of the Mandel parameter. For generated pseudothermal states, we obtain the Mandel parameter ranging from $Q_m = 0.012(2)$ to 8.93(3) for $\langle n\rangle$ up to 8.410(3). The ideal chaotic thermal light exhibits photon bunching, with the Mandel parameter Q_m equaling the mean photon number $\langle n\rangle$, as confirmed by the data [see Fig. 4(b)].

IV. CONCLUSION

We designed and developed a high-performance CCU, achieving low propagation delay, a well-defined adjustable coincidence window, 10-ps overall jitter, and high-precision multicoincidence counting. The CCU enables storing the complete histograms of all possible coincidence events (65 536 for 16 channels) with full channel resolution or a reduced histogram carrying information only about summarized n-fold coincidences. Alternatively, the information about the coincidence events can be directly used as an advanced trigger to control complex photonic systems. In terms of scalability, the CCU supports extensions of the total number of input channels, easily exceeding 100. Advancements in coincidence counting and processing to a mesoscopic scale offer groundbreaking applications in quantum communications, simulations, and boson sampling machines.

The presented CCU allows for a wide range of multiple-detector experiments. We performed a photon-number counting experiment employing single-photon detectors to investigate the statistical properties of the initial light. The results show unprecedented accuracy of photon statistics measurement ($\mathcal{F} > 0.999$) with a dynamic range of photon-number resolution up to 60 photons. The presented measurement process is independent of the number of channels and can be scaled up to a level where hundreds of photons can be detected, which is crucial for quantum receivers, quantum tomography, photonic source benchmarking, and other applications.

ACKNOWLEDGMENTS

This work is supported by the Czech Science Foundation (project 21-18545S). J.H. and J.G. also acknowledge support by Palacký University (J.H.: IGA-PrF-2022-005; J.G.: IGA-PrF-2022-001 and IGA-PrF-2023-002).

APPENDIX A: ELECTRONIC DESIGN AND IMPLEMENTATION OF THE COINCIDENCE COUNTING SYSTEM

Here, we present a detailed description of the developed ultrafast electronic multichannel CCU based on ECL circuitry. In the main text, Fig. 1(a) shows an overview of the CCU architecture consisting of a shaper, latch, and microcontroller. A shaper is a crucial circuit building block comprised of a fast comparator (MAX 9600), a pair of delay lines (SY89295U or MC100EP195, depending on market availability), and basic gates (MC100EP05). Figure 5(a) depicts a block diagram of a single shaper. For a detailed pulse shaping and processing description, see Fig. 5(b). The fast comparator processes the input signal to

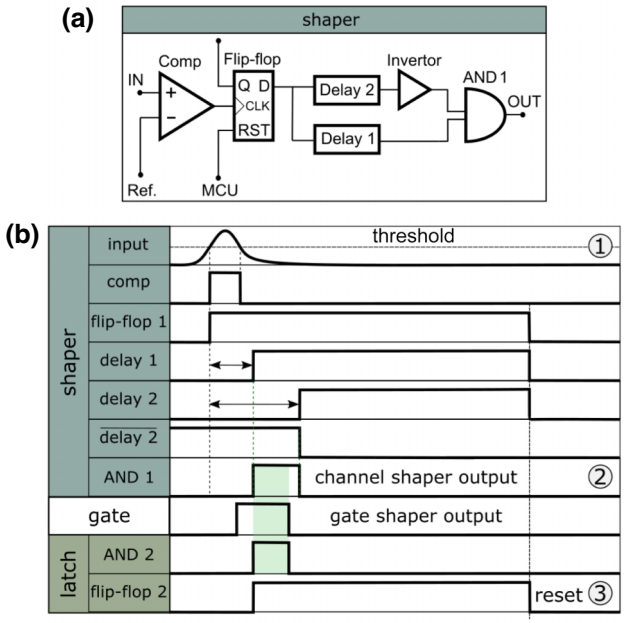


FIG. 5. (a) A block diagram of a shaper and (b) the timing diagram of the signal processing.

convert it from the original waveform to a defined pulse, triggering the first flip-flop circuit. Each input channel has an adjustable threshold from 1 to 4 V. The flip-flop stores state information and creates a time window independent of the input signal. The flip-flop output is split into two pulses and modified by parallel delay lines and inverters.

The pulse emitted by the shaper is defined as an output of the AND gate that implements the logical conjunction of the delayed pulses. The first (second) delay line sets the rising (falling) edge of the corresponding time window. The width of the time window is given as the time difference between these two propagation delays. The gate signal has its shaper, including one extra delay line for additional delay adjustment. The gate pulse is distributed to individual latch inputs utilizing electronic repeaters. Recently, after acquiring data for this work, we updated the CCU design. We distribute the rising edges of the gating signal and subsequently generate individual gate pulses. This approach mitigates spreading of short edges due to long propagation on printed circuit boards and further improves the quality (steepness) of the rising and falling edges.

Each programmable delay line employed in the shaping circuitry has 1024 discrete steps with an average delay of 9 ps (for our setting and operating temperature). The resulting coincidence window is adjustable from the sub-100-ps regime to approximately 20 ns, i.e., two times the maximum delay of a single signal shaper. The delay lines can also compensate for differences in input signal arrival times. The individual channels can be delayed within a range of up to 10 ns, as discussed above. The maximum delay between the gate and the signal channels can be set to 20 ns for a sub-100-ps coincidence window. However, for the maximum 20-ns coincidence window, the entire capacity of the delay lines is used to create the coincidence window, and the input signals need to be externally synchronized.

The latch, composed of 16 AND gates and the flip-flop circuits (MC100EP51), counts all possible coincidences between the rising edge of the gate pulse and 16 rising edges of the signal pulses. The AND gates generate an output pulse only when input and gate pulses are received simultaneously. The microcontroller unit (STM32F429) receives the data from the latch and sorts the results into a histogram in memory. Data, including information about the number of single-channel pulses and coincidences, are transferred to a personal computer via USB. Finally, the microcontroller unit resets all flip-flop circuits and enables the CCU for another detection event. Our CCU design offers unparalleled scalability, accommodating an unlimited number of input channels with the potential to support up to 100 input channels.

APPENDIX B: ELECTRONIC CHARACTERIZATION OF THE COINCIDENCE COUNTING SYSTEM

To characterize the performance and capabilities of the presented device, we evaluate the following figures of merit: (1) the minimum and maximum coincidence window length, (2) timing granularity, (3) timing jitter, (4) higher-order coincidence failure probability, and (5) maximum detection rate. We analyze the coincidence window width and rising and falling edges. The whole measurement is based on changing the mutual position of the gate and channel pulse. As a result, we obtain the number of coincidence events as a function of time delay. The minimum and maximum widths of the coincidence window are sub-100 ps and 20 ns, respectively. The main limitation of the maximum coincidence window width is the maximum range of delay lines. The results show that the coincidence windows are well defined in time, and they are almost perfectly rectangular with sharp edges with a value typically around 18(5) ps for all channels (see Figs. 6 and 7). The employed high-speed ECL components with fast transition times guarantee a short propagation delay below 5 ns. The overall jitter of the presented CCU is less than 10 ps.

The presented CCU meets the conditions for a high-resolution coincidence system with equal probabilities for all orders of coincidence events. We have analyzed the precision of counting coincidences utilizing a home-built multichannel pulse generator with a repetition rate from 0.2 to 10 MHz. The signal is generated by a relaxation oscillator whose output is delayed by a resistor-capacitor low-pass network and fast inverter with Schmitt trigger inputs (74ACT14T). This generator was developed to simulate the typical output signals of state-of-the-art

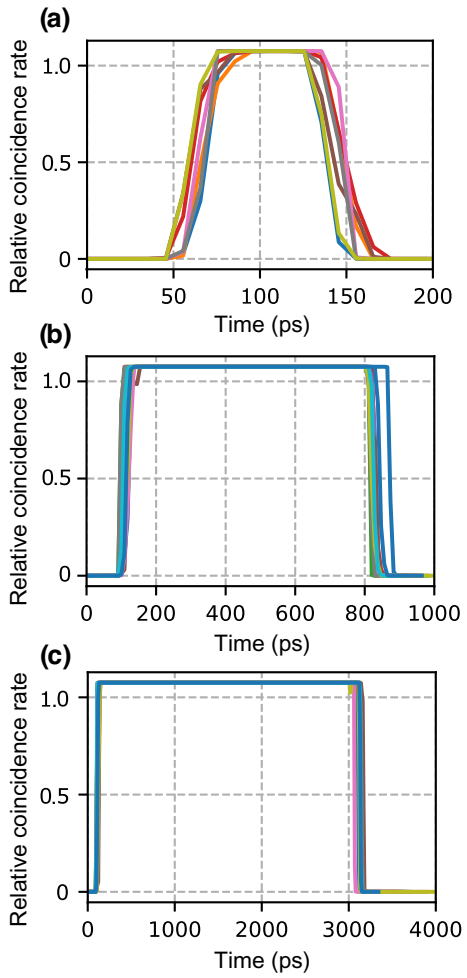


FIG. 6. A full scan of the coincidence window of width (a) sub-100 ps, (b) 700 ps, and (c) 3 ns.

single-photon detectors with a jitter of less than 10 ps. All detection channels are synchronized with the generator output pulse. The result is that all coincidence events

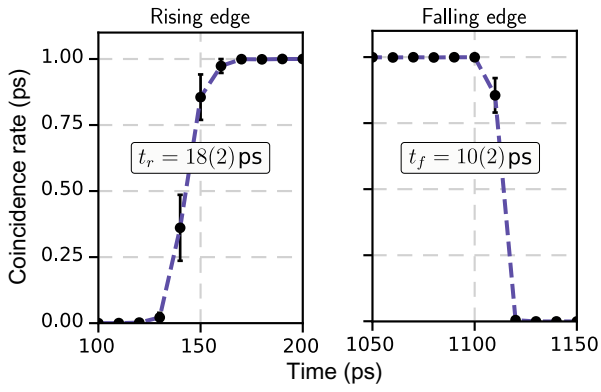


FIG. 7. The typical achieved rising and falling edges (10%–90%) of the coincidence window (width: 1 ns). Presented values are calculated as the average over the 250 measurement runs. The error bars represent one standard deviation.

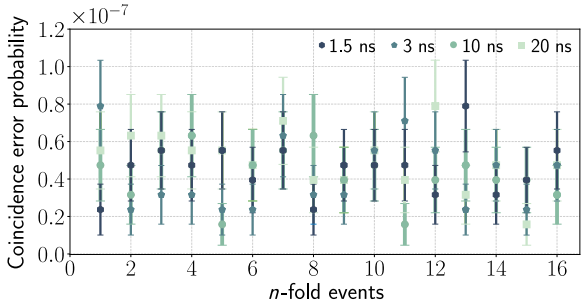


FIG. 8. Analysis of the coincidence error probability of the individual *n*-fold coincidences for several different widths of the coincidence windows: 1.5, 3.0, 10, and 20 ns. Presented values are calculated as the average over the 250 measurement runs. The error bars represent one standard deviation.

should be evaluated as 16-fold coincidences. We compare the measured coincidence events with the number of gating pulses to evaluate the counting error of the number of coincidence events. For measurements ranging from 16-fold coincidences down to single events, we systematically deactivate the detection channels one by one. Figure 8 shows the analysis of coincidence counting precision for several configurations of coincidence window width from 1.5 up to 20 ns. All coincidence errors across all tested coincidence windows are lower than 10^{-7}. This indicates ultralow coincidence errors due to resetting the latches, most probably caused by back reflections via imperfect impedance matching.

TABLE I. Performance parameters comparison of the presented CCU and other coincidence unit approaches.

Ref.	[25]	[26]	[27]	[28]	[29]	[59]	This work
Number of channels	8	48	8	32	8	20	16
Measurable coincidence folds	8	6	8	8	8	20	16
Maximum input frequency (MHz)	163	76	50	80	40[a]	400[b]	1500
Minimum coincidence window (ns)	0.47	0.3	10	0.39	···	0.46	<0.1
Maximum coincidence window (ns)	13.22	1.9	70	···	···	10	20
Channel resolution	No	No	No	No	Yes	No	Yes

[a]Up to 2-fold coincidences only.
[b]Characterized in 2-fold coincidence measurement.

APPENDIX C: CCU PERFORMANCE COMPARISON

In Table I, a summary of the performance of the various CCU approaches is shown. The presented CCU offers high n-fold coincidence counting ($n = 16$) with channel-number resolution, the shortest well-defined sub-100-ps coincidence windows with 10-ps overall jitter, and the highest maximum input frequency of 1.5 GHz.

[1] J. Wang, S. Paesani, Y. Ding, R. Santagati, P. Skrzypczyk, A. Salavrakos, J. Tura, R. Augusiak, L. Mančinska, D. Bacco, D. Bonneau, J. W. Silverstone, Q. Gong, A. Acín, K. Rottwitt, L. K. Oxenløwe, J. L. O'Brien, A. Laing, and M. G. Thompson, Multidimensional quantum entanglement with large-scale integrated optics, Science **360**, 6386 (2018).

[2] H. Wang, J. Qin, X. Ding, M.-C. Chen, S. Chen, X. You, Y.-M. He, X. Jiang, L. You, Z. Wang, C. Schneider, J. J. Renema, S. Höfling, C.-Y. Lu, and J.-W. Pan, Boson sampling with 20 input photons and a 60-mode interferometer in a 10^{14}-dimensional Hilbert space, Phys. Rev. Lett **123**, 25 (2019).

[3] H.-S. Zhong *et al.*, Quantum computational advantage using photons, Science **370**, 6523 (2020).

[4] X.-C. Yao, T.-X. Wang, P. Xu, H. Lu, G.-S. Pan, X.-H. Bao, C.-Z. Peng, C.-Y. Lu, Y.-A. Chen, and J.-W. Pan, Observation of eight-photon entanglement, Nat. Phot. **6**, 225 (2012).

[5] X.-L. Wang, L.-K. Chen, W. Li, H.-L. Huang, C. Liu, C. Chen, Y.-H. Luo, Z.-E. Su, D. Wu, Z.-D. Li, H. Lu, Y. Hu, X. Jiang, C.-Z. Peng, L. Li, N.-L. Liu, Y.-A. Chen, C.-Y. Lu, and J.-W. Pan, Experimental ten-photon entanglement, Phys. Rev. Lett **117**, 21 (2016).

[6] S. Paesani, Y. Ding, R. Santagati, L. Chakhmakhchyan, C. Vigliar, K. Rottwitt, L. K. Oxenløwe, J. Wang, M. G. Thompson, and A. Laing, Generation and sampling of quantum states of light in a silicon chip, Nat. Phys. **15**, 9 (2019).

[7] S. Aaronson and A. Arkhipov, in *Quantum Information and Measurement 2014* (Optica Publishing Group, Messe Berlin, Berlin Germany, 2014).

[8] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani, J. L. Pereira, M. Razavi, J. Shamsul Shaari, M. Tomamichel, V. C. Usenko, G. Vallone, P. Villoresi, and P. Wallden, Advances in quantum cryptography, Adv. Opt. Photonics **12**, 1012 (2020).

[9] A. Aspuru-Guzik and P. Walther, Photonic quantum simulators, Nat. Phys. **8**, 4 (2012).

[10] H. Paul, P. Törmä, T. Kiss, and I. Jex, Photon chopping: New way to measure the quantum state of light, Phys. Rev. Lett **76**, 14 (1996).

[11] J. Sperling, W. Vogel, and G. S. Agarwal, True photocounting statistics of multiple on-off detectors, Phys. Rev. A **85**, 2 (2012).

[12] J. Hloušek, M. Dudka, I. Straka, and M. Ježek, Accurate detection of arbitrary photon statistics, Phys. Rev. Lett. **123**, 15 (2019).

[13] H. J. Kimble, M. Dagenais, and L. Mandel, Photon antibunching in resonance fluorescence, Phys. Rev. Lett. **39**, 11 (1977).

[14] H. Esat Kondakci, A. Szameit, A. F. Abouraddy, D. N. Christodoulides, and B. E. A. Saleh, Sub-thermal to super-thermal light statistics from a disordered lattice via deterministic control of excitation symmetry, Optica **3**, 5 (2016).

[15] J. Sperling, A. Eckstein, W. R. Clements, M. Moore, J. J. Renema, W. S. Kolthammer, S. W. Nam, A. Lita, T. Gerrits, I. A. Walmsley, G. S. Agarwal, and W. Vogel, Identification of nonclassical properties of light with multiplexing layouts, Phys. Rev. A **96**, 1 (2017).

[16] I. Straka, L. Lachman, J. Hloušek, M. Miková, M. Mičuda, M. Ježek, and R. Filip, Quantum non-Gaussian multiphoton light, Npj Quantum Inf. **4**, 1 (2018).

[17] M. Bohmann, L. Qi, W. Vogel, and M. Chekhova, Detection-device-independent verification of nonclassical light, Phys. Rev. Res. **1**, 3 (2019).

[18] V. Švarc, M. Nováková, G. Mazin, and M. Ježek, Fully tunable and switchable coupler for photonic routing in quantum detection and modulation, Opt. Lett. **44**, 5844 (2019).

[19] G. Luiz Zanin, M. J. Jacquet, M. Spagnolo, P. Schiansky, I. A. Calafell, L. A. Rozema, and P. Walther, Fiber-compatible photonic feed-forward with 99% fidelity, Opt. Express **29**, 3425 (2021).

[20] M. Beck, Comparing measurements of $g^{(2)}(0)$ performed with different coincidence detection techniques, J. Opt. Soc. Am. B **24**, 12 (2007).

[21] D. Branning, S. Bhandari, and M. Beck, Low-cost coincidence-counting electronics for undergraduate quantum optics, Am. J. Phys. **77**, 7 (2009).

[22] M. Wahl, T. Röhlicke, H.-J. Rahn, R. Erdmann, G. Kell, A. Ahlrichs, M. Kernbach, A. W. Schell, and O. Benson, Integrated multichannel photon timing instrument with very short dead time and high throughput, Rev. Sci. Instrum. **84**, 4 (2013).

[23] S. Gaertner, H. Weinfurter, and C. Kurtsiefer, Fast and compact multichannel photon coincidence unit for quantum information processing, Rev. Sci. Instrum. **76**, 12 (2005).

[24] D. Branning, S. Khanal, Y. H. Shin, B. Clary, and M. Beck, Note: Scalable multiphoton coincidence-counting electronics, Rev. Sci. Instrum. **82**, 1 (2011).

[25] B. K. Park, Y.-S. Kim, O. Kwon, S.-W. Han, and S. Moon, High-performance reconfigurable coincidence counting unit based on a field programmable gate array, Appl. Opt. **54**, 15 (2015).

[26] C. Zhang, W. Li, Y. Hu, T. Yang, G. Jin, and X. Jiang, 48-channel coincidence counting system for multiphoton experiment, Rev. Sci. Instrum. **87**, 11 (2016).

[27] A. K. Gupta, R. Sankara Prasad, L. Srivani, D. Thirugnana Murthy, B. K. Panigrahi, and G. Raghavan, in *2018 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)* (Institute of Electrical and Electronics Engineers (IEEE), Bangalore, India, 2018), p. 1.

[28] W. Li, Y. Hu, H.-s. Zhong, Y.-f. Wang, X.-l. Wang, C.-z. Peng, and X. Jiang, Time-tagged coincidence counting unit for large-scale photonic quantum computing, Rev. Sci. Instrum. **89**, 10 (2018).

[29] E. Arabul, S. Paesani, S. Tancock, J. Rarity, and N. Dahnoun, A precise high count-rate FGPA based multi-channel coincidence counting system for quantum photonics applications, IEEE Photonics J. **12**, 2 (2020).

[30] C. Cahall, K. L. Nicolich, N. T. Islam, G. P. Lafyatis, A. J. Miller, D. J. Gauthier, and J. Kim, Multi-photon detection using a conventional superconducting nanowire single-photon detector, Optica **4**, 1534 (2017).

[31] D. Zhu, M. Colangelo, C. Chen, B. A. Korzh, F. N. C. Wong, M. D. Shaw, and K. K. Berggren, Resolving photon numbers using a superconducting nanowire with impedance-matching taper, Nano Lett. **20**, 3858 (2020).

[32] X. Tao, H. Hao, X. Li, S. Chen, L. Wang, X. Tu, X. Jia, L. Zhang, Q. Zhao, L. Kang, and P. Wu, Characterize the speed of a photon-number-resolving superconducting nanowire detector, IEEE Photonics J. **12**, 1 (2020).

[33] A. E. Lita, A. J. Miller, and S. W. Nam, Counting near-infrared single-photons with 95% efficiency, Opt. Express **16**, 3032 (2008).

[34] T. Gerrits, B. Calkins, N. Tomlin, A. E. Lita, A. Migdall, R. Mirin, and S. W. Nam, Extending single-photon optimized superconducting transition edge sensors beyond the single-photon counting regime, Opt. Express **20**, 23798 (2012).

[35] L. A. Morais, T. Weinhold, M. P. de Almeida, J. Combes, A. Lita, T. Gerrits, S. W. Nam, A. G. White, and G. Gillett, Precisely determining photon-number in real-time, arXiv:2012.10158 [physics.ins-det] (2022).

[36] M. Eaton, A. Hossameldin, R. J. Birrittella, P. M. Alsing, C. C. Gerry, H. Dong, C. Cuevas, and O. Pfister, Resolution of 100 photons and quantum generation of unbiased random numbers, Nat. Photonics **17**, 106 (2022).

[37] R. Cheng, Y. Zhou, S. Wang, M. Shen, T. Taher, and H. X. Tang, A 100-pixel photon-number-resolving detector unveiling photon statistics, Nat. Photonics **17**, 112 (2022).

[38] P. Li, J. Zhong, W. Zhang, Z. Wang, Q. Ma, Z. Feng, W. Miao, Y. Ren, J. Li, Q. Yao, and S. Shi, High-performance Ti transition-edge sensor-based photon-number resolving detectors, J. Low. Temp. Phys. (2023).

[39] P. Kok and S. L. Braunstein, Detection devices in entanglement-based optical state preparation, Phys. Rev. A **63**, 3 (2001).

[40] J. Řeháček, Z. Hradil, O. Haderka, J. Peřina, and M. Hamar, Multiple-photon resolving fiber-loop detector, Phys. Rev. A **67**, 6 (2003).

[41] K. Banaszek and I. A. Walmsley, Photon counting with a loop detector, Opt. Lett. **28**, 1 (2003).

[42] M. J. Fitch, B. C. Jacobs, T. B. Pittman, and J. D. Franson, Photon-number resolution using time-multiplexed single-photon detectors, Phys. Rev. A **68**, 4 (2003).

[43] D. Achilles, C. Silberhorn, C. Śliwa, K. Banaszek, and I. A. Walmsley, Fiber-assisted detection with photon number resolution, Opt. Lett. **28**, 23 (2003).

[44] D. Achilles, C. Silberhorn, C. Sliwa, K. Banaszek, I. A. Walmsley, M. J. Fitch, B. C. Jacobs, T. B. Pittman, and J. D. Franson, Photon-number-resolving detection using time-multiplexing, J. Mod. Opt. **51**, 9 (2004).

[45] G. A. P. Thé and R. Viana Ramos, Multiple-photon number resolving detector using fibre ring and single-photon detector, J. Mod. Opt. **54**, 8 (2007).

[46] R. Kruse, J. Tiedau, T. J. Bartley, S. Barkhofen, and C. Silberhorn, Limits of the time-multiplexed photon-counting method, Phys. Rev. A **95**, 2 (2017).

[47] M. Mičuda, O. Haderka, and M. Ježek, High-efficiency photon-number-resolving multichannel detector, Phys. Rev. A **78**, 2 (2008).

[48] E. A. Dauler, A. J. Kerman, B. S. Robinson, J. K. Yang, B. Voronov, G. Goltsman, S. A. Hamilton, and K. K. Berggren, Photon-number-resolution with sub-30-ps timing using multi-element superconducting nanowire single photon detectors, J. Mod. Opt. **56**, 2 (2009).

[49] T. J. Bartley, G. Donati, X.-M. Jin, A. Datta, M. Barbieri, and I. A. Walmsley, Direct observation of sub-binomial light, Phys. Rev. Lett **110**, 17 (2013).

[50] X. Chen, C. Ding, H. Pan, K. Huang, J. Laurat, G. Wu, and E. Wu, Temporal and spatial multiplexed infrared single-photon counter based on high-speed avalanche photodiode, Sci. Rep. **7**, 1 (2017).

[51] J. Tiedau, E. Meyer-Scott, T. Nitsche, S. Barkhofen, T. J. Bartley, and C. Silberhorn, A high dynamic range optical detector for measuring single photons and bright light, Opt. Express **27**, 1 (2019).

[52] I. Esmaeil Zadeh, J. Chang, J. W. N. Los, S. Gyger, A. W. Elshaari, S. Steinhauer, S. N. Dorenbos, and V. Zwiller, Superconducting nanowire single-photon detectors: A perspective on evolution, state-of-the-art, future developments, and applications, Appl. Phys. Lett. **118**, 190502 (2021).

[53] W. Martienssen and E. Spiller, Coherence and fluctuations in light beams, Am. J. Phys. **32**, 12 (1964).

[54] J. Sperling, W. Vogel, and G. S. Agarwal, Sub-binomial light, Phys. Rev. Lett. **109**, 9 (2012).

[55] P. Grangier, G. Roger, and A. Aspect, Experimental evidence for a photon anticorrelation effect on a beam splitter: A new light on single-photon interferences, Europhys. Lett. **1**, 4 (1986).

[56] R. Filip and L. Mišta, Detecting quantum states with a positive Wigner function beyond mixtures of Gaussian states, Phys. Rev. Lett. **106**, 200401 (2011).

[57] L. Mandel, Sub-Poissonian photon statistics in resonance fluorescence, Opt. Lett. **4**, 205 (1979).

[58] E. J. O'Reilly and A. Olaya-Castro, Non-classicality of the molecular vibrations assisting exciton energy transfer at room temperature, Nat. Commun. **5**, 1 (2014).

[59] B. K. Park, Y.-S. Kim, Y.-W. Cho, S. Moon, and S.-W. Han, Arbitrary configurable 20-channel coincidence counting unit for multi-qubit quantum experiment, Electronics **10**, 5 (2021).

Research Article | Vol. 29, No. 21 / 11 Oct 2021 / Optics Express 33037
Optics EXPRESS

Accurate polarization preparation and measurement using twisted nematic liquid crystals

MARTIN BIELAK,[1] ROBERT STÁREK,[1] VOJTĚCH KRČMARSKÝ,[1,2,3] MICHAL MIČUDA,[1] AND MIROSLAV JEŽEK[1,*]

[1]Department of Optics, Faculty of Science, Palacky University, 17. listopadu 12, 77146 Olomouc, Czech Republic
[2]Institut für Quantenoptik und Quanteninformation, Österreichische Akademie der Wissenschaften, Technikerstr. 21A, 6020 Innsbruck, Austria
[3]Institut für Experimentalphysik, Universität Innsbruck, Technikerstr. 25, 6020 Innsbruck, Austria
*jezek@optics.upol.cz

Abstract: Generation of particular polarization states of light, encoding information in polarization degree of freedom, and efficient measurement of unknown polarization are the key tasks in optical metrology, optical communications, polarization-sensitive imaging, and photonic information processing. Liquid crystal devices have proved to be indispensable for these tasks, though their limited precision and the requirement of a custom design impose a limit of practical applicability. Here we report fast preparation and detection of polarization states with unprecedented accuracy using liquid-crystal cells extracted from common twisted nematic liquid-crystal displays. To verify the performance of the device we use it to prepare dozens of polarization states with average fidelity 0.999(1) and average angle deviation 0.5(3) deg. Using four-projection minimum tomography as well as six-projection Pauli measurement, we measure polarization states employing the reported device with the average fidelity of 0.999(1). Polarization measurement data are processed by the maximum likelihood method to reach a valid estimate of the polarization state. In addition to the application in classical polarimetry, we also employ the reported liquid-crystal device for full tomographic characterization of a three-mode Greenberger–Horne–Zeilinger entangled state produced by a photonic quantum processor.

1. Introduction

Direction, phase, and coherence of electromagnetic wave oscillation, i.e. its polarization state, represents an important feature determining the interaction of light and matter. Light reflection and scattering are fundamentally affected by incident polarization [1]. Manipulating the state of polarization can significantly improve optical resolution [2] and visualization of anisotropic structures in biomedical imaging [3,4]. Polarization modulation enhances information capacity in optical communication [5] and represents a building block of quantum communication and quantum information processing [6,7]. Precise generation and detection of the polarization state represent the crucial tasks in a vast number of applications. Electrically tunable birefringent elements like liquid crystals, Pockels cells, and integrated electrooptical modulators are often employed to control polarization in such applications. Pockels based polarization modulators were successfully applied in optical switches [8] and loop-based photonic routing, where high transmittance is of paramount importance [9–11]. They offer high-speed operation and an acceptable extinction ratio, but their application is rather cumbersome due to their size and the necessity of high-voltage driving. Integrated devices achieve even wider bandwidth enabling ultra-fast polarimetry [12]. They possess significant losses and their calibration might be challenging, though, mainly due to unavoidable waveguide coupling. Liquid crystals allow

#388675 https://doi.org/10.1364/OE.388675
Journal © 2021 Received 20 Jan 2020; revised 20 Aug 2021; accepted 23 Aug 2021; published 27 Sep 2021

Optics EXPRESS

low-loss free-space polarization addressing using low-voltage control signals. Their operation is faster than mechanically manipulated birefringent elements, which has proven beneficial in polarimetry [13–16] and spectrometry [17,18]. Liquid crystal devices have also been successfully employed in polarization and phase modulation [19–25] and switching [26,27]. Moreover, the absence of moving parts improves the robustness and lifespan of the device [25,28].

In a vast majority of the polarimetric applications reported so far, nematic liquid crystals are used in custom devices acting as variable retarders [13,14,19,21,25]. Incompatibility of these devices with the common *twisted* nematic liquid crystal (TNLC) configuration widely utilized in display technology makes their broad application difficult. Another complication stems from the fact, that additional wave plates or birefringent compensators are often utilized to change the overall polarization transformation. Peinado et al. reported a polarimeter using a quarter-wave plate and a single tilted twisted nematic element in reflective geometry [16]. The device was optimized to project an unknown polarization state to four polarization projections symmetrically placed on the Bloch sphere (Poincaré sphere), i.e. the vertices of a regular tetrahedron inscribed into the sphere. This minimum information measurement, also termed the minimum tomography [29,30], was shown to be efficient but prone to noise and measurement errors [31–35]. Furthermore, a TNLC device applies a complex combination of rotation and retardation to incident polarization state, and the exact theoretical description of the device operation and its calibration remains a significant challenge [36,37].

Here we present a TNLC device based on a commercially available TNLC display with minimal modifications. The construction of the *TNLC device* is described in Section 2. In Section 3, we show that the theoretical model of a twisted nematic liquid crystals cell has a limited ability to predict the polarization changes. The novel calibration method, described in Section 4, *does not require* a theoretical model and yields unprecedented accuracy. We show how to calibrate the TNLC device to prepare an arbitrary polarization state and perform arbitrary polarization projection. The calibration stage is boosted by a genetic algorithm and other possible improvements are also discussed. In section 5, we test the TNLC device by preparing more than a hundred polarization states covering the Bloch sphere uniformly with an unprecedentedly high average fidelity of 0.999(1), including six eigenstates of Pauli operators and four states needed for the minimum tomography. Also, we show depolarized state preparation with degree of polarization 0.03(1). The presented device is ready to be used as a polarimeter, performing the polarization state tomography. In Section 6 we briefly review the polarization state tomography, and then we test the performance of the TNLC-based polarimeter. Furthermore, the TNLC device is employed in full quantum tomography of three-qubit Greenberger–Horne–Zeilinger entangled state produced by photonic quantum Toffoli gate.

2. Construction of TNLC device

The presented polarimetric device consists of three TNLC cells extracted from commercially available displays. We tested several commercially available displays, namely TN reflective display *Lumex LCD-S101D14TR*, super-TN (STN) reflective display *RS Pro 5080PHR*, transmissive TNLC display *OCZ Vrchlabi 2027*. We also tested transmissive STN pixel display *OCZ Vrchlabi 2001* and reflective STN pixel display *DOG XL*. Due to their limited fill factor, the transmitted polarization state is spatially modulated and effectively depolarized. We therefore find them inappropriate for the presented application. We selected display *Lumex LCD-S101D14TR*, which has 7 segments (cells) and dimensions 33×50 mm, see Fig. 1(a). This particular model was chosen as a trade-off between the size of the segment and the overall compactness of the whole device. The TNLC reflective display cannot be directly used in transmission geometry or even to perform a unitary transformation. Therefore, we first removed reflective layers, protective films, polarizing sheets, and also other auxiliary layers from the displays. Using a micrometer screw

Optics EXPRESS

gauge, we measured the thickness of the LC layer to be $6 \pm 2\mu$m. The number in parenthesis represents one standard deviation at the least significant digit.

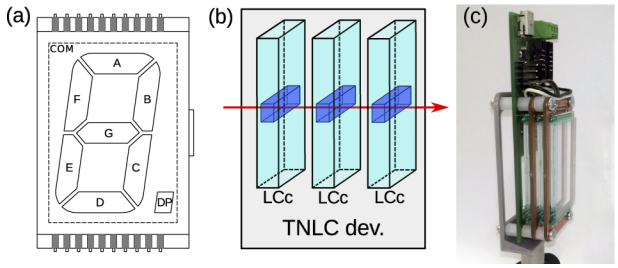


Fig. 1. (a) Seven segment liquid crystal display layout; elements of the digit are labeled from A to G, the decimal point as DP, and the common back electrode as COM. (b) Scheme of the presented TNLC device consisting of three TNLC cells (LCc); dark blue blocks show the active segments and the red arrow shows the optical beam. (c) Photo of the TNLC device mounted on printed circuit board containing the electronic driver and communication interface.

We mounted the TNLC glass cells on printed circuit boards and stacked them, as shown in Fig. 1(b,c). The main circuit board contains a driver consisting of a microcontroller, 16-bit digital-to-analog converter, and a voltage amplifier [38]. The microcontroller receives serial commands from a computer and controls the digital-to-analog converter. The produced control voltage signal is a symmetric square wave with frequency 1 kHz and 50% duty cycle. The square wave amplitude spans from 0 to 10 Vpp and effectively controls the action of the liquid crystals. For brevity, we refer to the control voltage amplitude simply as the *voltage*. The voltage is applied between the central segment electrode, denoted G in Fig. 1(a), and the common back electrode (COM) of the TNLC cell. The central segment has a 15×4.5 mm approximately rectangular clear aperture, which is perfectly sufficient for collimated optical beams with mode field diameter up to 2.5 mm. The device can address the polarization state of several parallel optical beams arranged in a line or a matrix in complex quantum information processing circuits [39–41].

The transmittance of the single cell is 84% and the transmittance of the whole device is 59% at wavelength 810 nm. Limited transmittance is not an obstacle for classical polarimetry. At the single-photon level, the presented transmittance is also sufficient, as we demonstrate in Section 6. Antireflective coating TNLC cells could increase the transmittance of a single cell up to 90%, increasing the overall transmittance to 73%. Only in loss-sensitive scenarios, for example measurements involving squeezed light, the limited transmittance becomes an obstacle.

3. Characterization of a single TNLC cell

Let us start with a brief review of the Dirac notation, used here to describe the polarization, and connect it to the traditional description of polarized light. We describe pure polarization states with a *ket state*, denoted with symbol $|\cdot\rangle$, which is equivalent to the normalized Jones vector. It holds that $|J\rangle = \frac{\mathbf{J}}{||\mathbf{J}||}$, where $\mathbf{J} = J_x\mathbf{x} + J_x\mathbf{y}$ is the Jones vector and $\mathbf{x}$ and $\mathbf{y}$ are orthonormal vectors, both normal to the direction of light propagation. The choice of basis is arbitrary, and the vector can be written as a superposition of any two orthonormal basis states. We choose basis states $|H\rangle$ and $|V\rangle$ corresponding to horizontally and vertically polarized light to keep the close correspondence between Jones vector and ket states. Explicitly, the correspondence is $|J\rangle = \left(J_x|H\rangle + J_y|V\rangle\right)\left(|J_x|^2 + |J_y|^2\right)^{-1/2}$.

Density matrix $\rho = \sum_\lambda p_\lambda |J_\lambda\rangle\langle J_\lambda|$ composed as an incoherent mixture of pure states $|J_\lambda\rangle$ with a probability distribution p_λ also describes polarization of light, but in comparison to ket-vectors, it can also describe partially polarized light. Term $|J_\lambda\rangle\langle J_\lambda|$ is the outer product

Optics EXPRESS

of $|J\rangle$ and its Hermitian conjugate, a 2×2 complex trace-normalized Hermitian matrix. Any polarization density matrix can be expanded using the identity $\mathbb{I}$ and the Pauli matrices, σ_1, σ_2, σ_3, $\rho = (\mathbb{I} + \sum_{j=1}^{3} u_j \sigma_j)/2$, where $\mathbf{u} = (u_1, u_2, u_3)$ is the *Bloch vector*. Compared to the Stokes vector, the Bloch vector does not contain information about the optical intensity, and its size is upper-bounded to one. We represent the polarization states geometrically using the *Bloch sphere*. A Bloch vector contains Cartesian coordinates of a point within the sphere. A point on the surface of the Bloch sphere, described with longitude θ and latitude ϕ, corresponds to a pure state $|J\rangle = \cos(\frac{\theta}{2})|H\rangle + \exp(i\phi)\sin(\frac{\theta}{2})|V\rangle$. In this representation, left- and right-handed circular polarizations lie on the south and north pole of the sphere, respectively. All linear polarizations lie on the equator of the sphere. A point inside the sphere describes a partially polarized state. We quantify its *purity* as $P = \mathrm{Tr}(\rho^2) = \frac{1}{2}\left(1 + ||\mathbf{u}||^2\right)$. The length of Bloch vector, $||\mathbf{u}|| = \sqrt{\sum_{j=1}^{3} u_j^2}$, in the expression for purity is the classical degree of polarization (DoP). *Fidelity* $F = [\mathrm{Tr}\sqrt{\sqrt{\rho_u}\rho_v\sqrt{\rho_u}}]^2$ of density matrices ρ_u and ρ_v quantifies their similarity. It can be expressed in terms of the corresponding Bloch vectors, $\mathbf{u}$ and $\mathbf{v}$, as $F = \frac{1}{2}\left(1 + \mathbf{u}\cdot\mathbf{v} + \sqrt{(1-||\mathbf{u}||)(1-||\mathbf{v}||)}\right)$. The fidelity is connected to the angular deviation of the Bloch vectors $\vartheta = \frac{1}{2}\arccos\left(\frac{\mathbf{u}\cdot\mathbf{v}}{||\mathbf{u}||||\mathbf{v}||}\right)$ via the dot-product term.

The polarization transformation introduced by a device is described with matrix M acting on a density matrix of the input state, $\rho_{\mathrm{out}} = M\rho_{\mathrm{in}}M^{\dagger}$, where $\dagger$ stands for Hermitian conjugation. The transformation matrix M of an ideal TNLC cell can be derived using a sequence of thin wave plates, each introducing a small phase delay and oriented at linearly increasing azimuth angle [42]. The transformation reads

$$M_{\mathrm{TNLC}} = R(\alpha)R(\varphi)\begin{pmatrix} \cos\chi + i\frac{\delta}{\chi}\sin\chi & -\frac{\varphi}{\chi}\sin\chi \\ \frac{\varphi}{\chi}\sin\chi & \cos\chi - i\frac{\delta}{\chi}\sin\chi \end{pmatrix} R(\alpha)^{\dagger}, \tag{1}$$

$$R(\varphi) = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}, \tag{2}$$

where $\chi = \sqrt{\varphi^2 + \delta^2}$, φ is the twist angle, α is the rotation angle of the cell, and δ represents the phase delay dependent on the applied voltage V. The phase delay (*retardance*) is almost constant up to a threshold voltage and decreases for larger voltages. The retardance is usually approximated using an arctan function [43]. We propose to model the retardance more precisely using a logistic function,

$$\delta = A + \frac{1}{B + \exp\left(C - DV^E\right)}, \tag{3}$$

where A, B, C, D, and E are constants and V is the voltage applied on the TNLC cell.

Let us show how a single TNLC cell transforms a polarization state, according to the model, and compare it with experimental data. We prepared state $|H\rangle$ on the input of the tested TNLC cell and changed the control voltage amplitude. For each voltage, we measured the output polarization state using a reference polarimeter. We plot the phase delay introduced by the cell in Fig. 2 as a function of voltage. The least-square fit with the proposed logistic function in Fig. 2 clearly describes the measured data better than the approximation with arctan. This measurement was repeated for four known angular positions β of the TNLC cell, namely $\beta = 0$ deg, 15 deg, 30 deg, and 45 deg. See Fig. 3(a) for a scheme of the experimental setup used in this measurement. We performed a least-square fit of the measured data to obtain seven independent parameters of the TNLC model $A = -7.8$, $B = 0.076$, $C = 176.16$, $D = -2.9$, $E = -2.9$, $\varphi = -5.16$, and $\alpha = -2.6$.

Optics EXPRESS

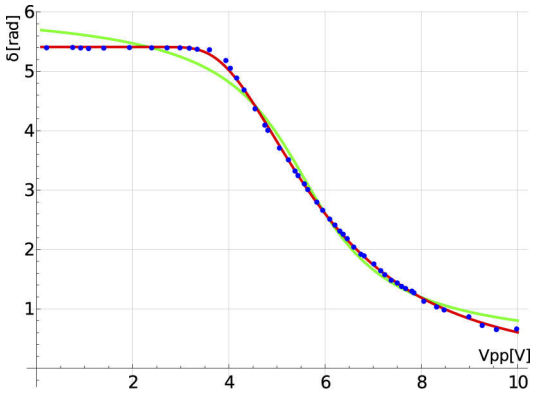


Fig. 2. The measured phase delay of a single TNLC cell obtained from the display Lumex LCD-S101D14TR as a function of the applied voltage at 810 nm (blue markers). The delay is described using an arctan function (green) or a logistic function (red).

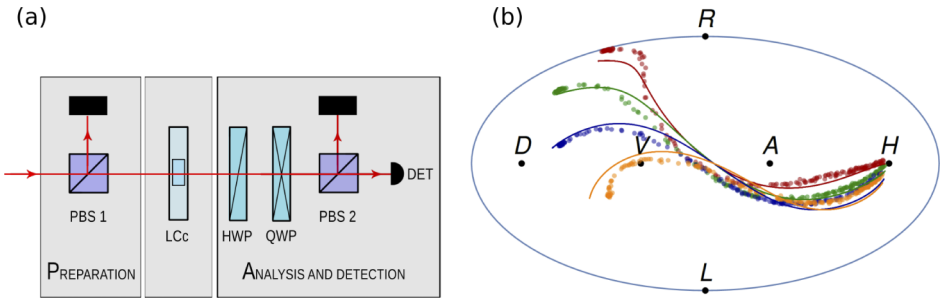


Fig. 3. (a) Scheme of the setup used for TNLC cell characterization. The following components are employed: tested TNLC cell (LCc), a polarizing beam splitter (PBS), quarter-wave plate (QWP), half-wave plate (HWP), detector (DET) – photodiode. The *detection and analysis* block serves as a reference polarimeter. (b) Transformation of the horizontally polarized light by single TNLC cell in Hammer map projection. Dots represent measured states. Solid lines represent fit. Additional display rotation angles 0 deg, 15 deg, 30 deg, and 45 deg are indicated by red, green, blue, and orange colors in this order.

In Fig. 3(b), we compare measured data with the theoretical predictions based on fitted parameters. The prepared and expected states are depicted using the Bloch sphere, which is plotted in the Hammer map projection [44]. Clearly, a single TNLC cell covers only a curve on the Bloch sphere. To transform state $|H\rangle$ into an arbitrary pure polarization state, one needs to introduce additional independent polarization transformations. This is the reason why the reported TNLC device consists of three TNLC cells controlled with independent voltages. Furthermore, the model deviates from the measured data even though it was fitted to the same data. Parameters of the model vary between individual TNLC cells, which represents a complication in modeling multi-cell devices. The accuracy of the model could be further improved by considering edge effects in the vicinity of aligning layers [36,37] and accounting for other imperfections, such as multiple reflections, depolarization effects, and thickness inhomogeneity. However, no model with sufficiently high fidelity with the experimental data covering all possible states has been reported. Also, it is very complicated to make an inverse of such a complex numerical model. In the following section, we introduce a novel calibration method that solves both issues.

Optics EXPRESS

4. Calibration of TNLC device

We aim for the device able to prepare an arbitrary polarization state with high precision. The calibration of the TNLC device consists of finding control voltages V_1, V_2, and V_3, one for each TNLC cell, for which the device transforms the pure input state $|H\rangle$ to the pure particular target state $|\psi\rangle$.

The calibration is thus formulated as *maximization of the fidelity* $F = |\langle\psi|M(V_1, V_2, V_3)|H\rangle|^2$, where M is transformation introduced by the TNLC device, over the three control voltages. The fidelity $|\langle\psi|M(V_1, V_2, V_3)|H\rangle|^2$ is proportional to the intensity of projection onto the target state, which is directly experimentally measurable. The maximization *does not rely* on the theoretical model, and various methods could be used to solve it. Straightforward methods, like brute-force search or random sampling, are time-consuming as they require many measurements. Therefore we choose a genetic algorithm [45] to speed up the calibration.

The initial population is created as three-element random-voltage vectors, and a fitness function is the detected photocurrent. We sequentially apply these voltages settings and directly measure the value of the fitness function. We choose voltages with the largest fitness from this first generation and crossover them to produce "child" voltages, which replace the voltages of lesser fitness. As a crossover method, we chose the arithmetic mean of the corresponding voltages. A random mutation in range of ±500 mVpp is also applied to a randomly chosen component of the voltage vector. These steps are repeated until the fitness changes less than one part per thousand, which is roughly an order of magnitude larger than the uncertainty of the intensity measurement. Further iterations would slightly improve the precision at the expense of the longer time needed for the calibration. Although the TNLC theoretical model (1)–(3) is neither perfect nor required, it can be used to speed up the search by providing a rough initial estimate of the voltage values for the first generation of the genetic algorithm. Even faster calibration might be obtained using black-box optimization methods [46]. As the last step of the calibration, we generate 120 random voltage vectors around the fittest voltage vector using a normal distribution with a half-width of 0.1 Vpp. We prepare states using these voltages and measured them using a reference polarimeter. Finally, we select the optimum voltages that produce the target state with the highest fidelity.

The optimum voltages slightly vary for the TNLC cells extracted from different TNLC displays, even when selected from a single manufacturing batch. The voltages are also affected by offsets of voltage regulators and digital-to-analog converters incorporated in the electronic drivers, see Sec. 2. Consequently, the calibration has to be performed independently for each TNLC device. A detailed calibration sheet with optimum voltages for various target states and several TNLC devices assembled in our lab is available on GitHub [38].

The experimental setup for the calibration is depicted in Fig. 4(a). An 810-nm continuous-wave laser light passes the first horizontal polarizer (calcite crystal with extinction ratio exceeding $1 : 10^4$), and then it passes the TNLC device. The output polarization state is projected onto the desired state $|\psi\rangle$ in a detection block, consisting of half- and quarter-wave plates followed by another horizontal polarizer (another calcite crystal). The photodiode measures the optical intensity, which is proportional to the fitness function. The wave plates are mounted in motorized rotation stages (Newport PR50CC) with the angular speed of 20 deg/s and typical bi-directional repeatability ±30 mdeg. The angular positions of these wave plates control which projection we measure. We also use the waveplates to realize projections to six eigenstates of Pauli matrices, $|H\rangle$, $|V\rangle$, $|D\rangle$, $|A\rangle$, $|R\rangle$, and $|L\rangle$. Section 6 explains how to use these projections to reconstruct a quantum state. For now, we state that the *analysis and detection* block serves also as a reference polarimeter.

In addition to *preparing* arbitrary pure polarization state, the TNLC device can *project* a state onto an arbitrary pure state. Just by simple reversion of the calibration setup, depicted in Fig. 4(b), we could maximize $F = |\langle H|M(V_1, V_2, V_3)|\psi\rangle|^2$ using the same algorithm. The presented calibration can be performed for an arbitrary wavelength within visible and near-infrared regions.

Optics EXPRESS

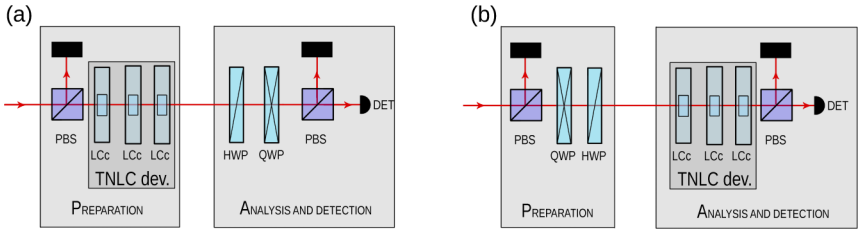


Fig. 4. Schemes of the calibration (a) and the polarimetric measurement using the TNLC device (b). The following components are employed: the TNLC device consisting of three TNLC cells (LCc), polarizing beam splitter (PBS), quarter-wave plate (QWP), half-wave plate (HWP), detector (DET) – either photodiode or single photon detector.

5. Precise preparation of arbitrary polarization

We first test the presented TNLC device calibration on a set of six eigenstates of Pauli operators. We measured the prepared state with the reference polarimeter. The Bloch parameters of the prepared states are shown in Table 1. The prepared states show high purity, $P = 0.999(2)$ on average, and are very close to the ideal target states. The number in parenthesis represents one standard deviation at the last decimal digit [47]. The average fidelity of the observed state to the ideal target states reaches $F = 0.999(1)$ with the corresponding average angular deviation of 1.1(3) deg. We believe that the very small remaining discrepancy is mainly due to the stopping criteria of the calibration search (one part per thousand) and also inaccuracies of the reference polarimeter, such as retardation error of the waveplates or their misalignment.

Table 1. Bloch parameters of H, V, D, A, R, and L polarization states prepared by the TNLC device and measured by the reference polarimeter.

Bloch parameters	H	V	D	A	R	L
u_1	**0.99941(5)**	**−0.99930(3)**	−0.040(2)	−0.0079(6)	−0.0192(5)	0.023(2)
u_2	0.028(1)	−0.0355(9)	**0.99918(7)**	**−0.99723(6)**	−0.0307(2)	0.047(3)
u_3	0.0199(4)	0.0112(4)	−0.0004(7)	−0.0186(1)	**0.997827(9)**	**−0.9938(3)**

We observed perfect stability in time and low sensitivity to temperature changes under common laboratory conditions. The cell's temperature variation within ±3°C did not cause any measurable changes in the calibration. Similarly, central wavelength variation within ±5 nm does not introduce any measurable changes.

We chose the six eigenstates of Pauli operators because they are traditionally used in polarization state tomography, which we will show in Section 6. Another important tomographic set of states forms the vertices of a regular tetrahedron inscribed into the Bloch sphere. We searched for the states that maximize the tetrahedron volume and reached the ratio of the measured volume to the volume of the regular tetrahedron $V_{\mathrm{meas}}/V_{\mathrm{ideal}} = 0.9987(5)$.

Furthermore, we prepare 120 states, quasi-uniformly distributed over the Bloch sphere. Figure 5 visually compares the prepared states with the ideal target states. It clearly shows that the presented device can produce an arbitrary pure state with high precision. The prepared states were again almost pure, with average purity $P = 0.999(2)$. The average fidelity of 0.999(1) and the corresponding average angular deviation of 0.5(3) deg indicate the precise preparation of arbitrary pure state achieved by the presented device.

In addition to the pure polarization states, it is also beneficial to prepare partially mixed states, i.e., states with an arbitrary degree of polarization DoP. Depolarization can be achieved by ensemble averaging of a polarization state in spatial, frequency, or time domain [22]. Often, the maximally depolarized state is prepared and then superimposed to a pure state to form a source

Optics EXPRESS

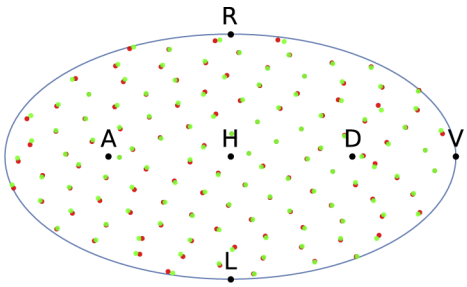


Fig. 5. Demonstration of the Bloch sphere coverage using the reported TNLC device. The target states are represented by red points, green markers show the states prepared by the TNLC device. The sphere is visualized in the Hammer map projection.

with continuously tunable DoP. In the case of a voltage-controlled TNLC cell, a specific voltage applied to the TNLC cell performs corresponding unitary operation. Therefore, time multiplexing of various control voltages within the single acquisition time must be used. We demonstrate the preparation of a maximally mixed state by uniformly cycling through the polarization states H, V, D, A, R, and L during the measurement time. Note that the acquisition is paused during the time needed to switch from one prepared state to another. The resulting prepared state possesses the purity 0.5004(3) and the corresponding DoP 0.03(1).

The reported TNLC device is made from thick TNLC cells and is not optimized with respect to its speed. However, to present the full characterization of the TNLC device, we measured its time response, for which we used the setup shown in Fig. 4(b). The wave plates prepare a pure input state which is then projected onto another pure state using the tested device. The corresponding projection intensity is detected with a photodiode. To study the time response, we switch the projection and sample the transition in the detected photocurrent using an analog-to-digital converter with an 18-bit resolution acquiring approximately 1000 samples per second.

Specifically, we measured the transition times of all the projections onto Pauli operators eigenstates (H, V, D, A, R, L). We prepared a corresponding input state for an initial projection to obtain a maximal optical intensity at the detector. Then we switched the projection. We define the *transition time* as the time from the issued command to the moment when the intensity deviation from its final value stays under 1% of the total intensity (0% to 99% transition time). We chose this value to ensure a stable operation after the transition. The measured transition times are shown in Table 2, and the time traces of all 30 transitions are shown in Fig. 6.

Table 2. Transition times in milliseconds between the initial $|\pi_{in}\rangle\langle\pi_{in}|$ and final polarization projection $|\pi_{fi}\rangle\langle\pi_{fi}|$ of the TNLC device.

		$\|\pi_{fi}\rangle\langle\pi_{fi}\|$					
		H	V	D	A	R	L
$\|\pi_{in}\rangle\langle\pi_{in}\|$	H		49	226	590	251	319
	V	168		182	597	271	329
	D	313	232		200	203	255
	A	444	403	188		312	197
	R	566	151	417	481		141
	L	601	497	299	185	424	

The particular transitions, such as H→V, are rather fast and suitable for rapid polarization switching, but there is a transition exceeding 600 ms, namely L→H. Specially designed nematic

Optics EXPRESS

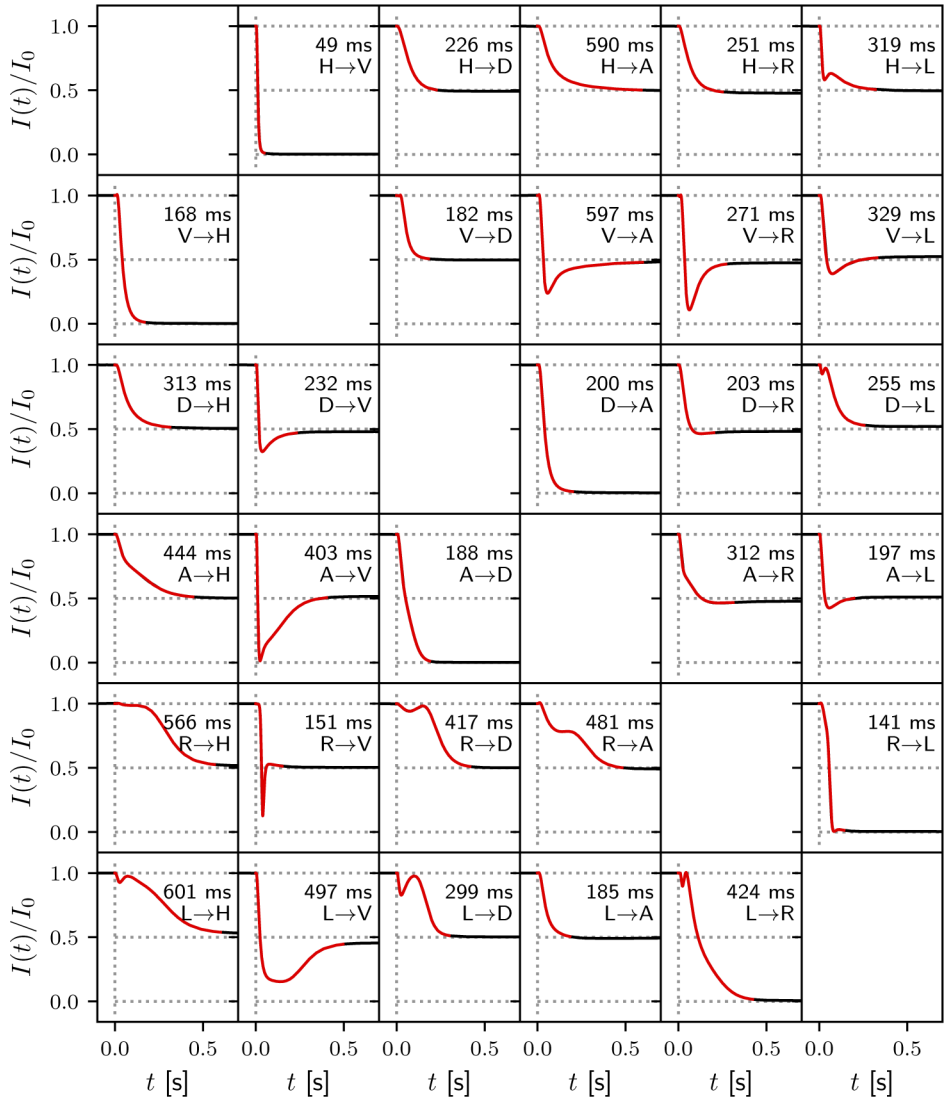


Fig. 6. Time traces of transitions between two projections realized by the TNLC device. Initial projections are arranged in rows and the final projections are arranged in columns. For each transition from initial to final projection, there is a time trace of measured intensity. Red color highlights the transition time. The horizontal axes in all plots represent time in seconds while the vertical axes represent normalized intensity.

Optics EXPRESS

liquid crystal retarders and switches typically specify faster switching, particularly those with a thinner liquid crystal layer and a single optimized polarization transition. Also, their transition times are often specified as 10% to 90% or even 20% to 80%, which is not sufficient for precise settling of the induced polarization transformation in polarimetry.

6. TNLC polarimeter

The presented device is compact, robust, and capable of precise polarization manipulation with less than one second transition times. Here, we utilize the developed TNLC device to create a polarimeter based on *polarization state tomography*. The polarization state tomography finds its applications mainly when we deal with single-photon signals or measure multi-photon polarization states. In these applications, the classical polarimetry techniques are not sufficient. The presented polarimeter outperforms conventional rotating-waveplates-based polarization tomography in speed, precision, and versatility.

Let us first briefly review the principle of polarization state tomography. The goal of the polarization state tomography is to unambiguously reconstruct an unknown state ρ from a set of measurements. Let us start with the strong signal and six-state tomography. We assume that the light has intensity I_0 and unknown polarization. Recall that any density matrix could be decomposed as a sum of Pauli operators multiplied by Bloch vector components $\rho = (\mathbb{I} + \sum_{j=1}^{3} u_j \sigma_j)/2$. The projections onto eigenstates of these operators, $|H\rangle$, $|V\rangle$, $|D\rangle$, $|A\rangle$, $|R\rangle$, and $|L\rangle$ contains information about the polarization. Specifically, the measured intensities of these projections are

$$I_H = I_0 \langle H|\rho|H\rangle = I_0(1 + u_1), \tag{4}$$

$$I_V = I_0 \langle V|\rho|V\rangle = I_0(1 - u_1), \tag{5}$$

$$I_D = I_0 \langle D|\rho|D\rangle = I_0(1 + u_2), \tag{6}$$

$$I_A = I_0 \langle A|\rho|A\rangle = I_0(1 - u_2), \tag{7}$$

$$I_R = I_0 \langle R|\rho|R\rangle = I_0(1 + u_3), \tag{8}$$

$$I_L = I_0 \langle L|\rho|L\rangle = I_0(1 - u_3). \tag{9}$$

These six equations contain three unknown parameters u_i of the polarization state. In the previous section, we showed that our device is capable of realizing these projections. One can obtain the parameters u_i from (4–9) by straightforward inversion,

$$u_1 = \frac{I_H - I_V}{I_H + I_V}, \tag{10}$$

$$u_2 = \frac{I_D - I_A}{I_D + I_A}, \tag{11}$$

$$u_3 = \frac{I_R - I_L}{I_R + I_L}. \tag{12}$$

In the presence of measurement noise, the linear inversion may introduce artifacts. For example, the reconstructed density matrix can have negative eigenvalues, leading to a degree of polarization greater than one. We call such a density matrix *unphysical*.

To avoid these problems, we use a maximum-likelihood reconstruction method [48–51]. The method searches for a valid quantum state that would reproduce the measured data with the maximum likelihood. It guarantees that the reconstructed state is always physical. Instead of optical intensities I_i of projection $\langle \pi_i|\rho|\pi_i\rangle$, we work with relative frequencies f_i, which

Optics EXPRESS

correspond to normalized intensities $f_i = \frac{I_i}{\sum I_i}$. The logarithm of likelihood of observing relative frequencies $\{f_i\}$ given the input polarization state ρ is proportional to

$$\ln \mathcal{L}(\rho) = \sum_i f_i \ln\langle\pi_i|\rho|\pi_i\rangle. \tag{13}$$

The likelihood quantifies the degree of belief in the hypothesis that for the particular set of observations the polarization is in state ρ. We can work with the logarithm of the likelihood with neglected multiplicative factors because if the likelihood is maximized, then any monotonically increasing function of the likelihood is also maximized.

The density matrix ρ, which maximizes the likelihood, satisfies the extremal equation [48,50,51]

$$K\rho = \rho, \tag{14}$$

where

$$K = \sum_i f_i \frac{|\pi_i\rangle\langle\pi_i|}{\langle\pi_i|\rho|\pi_i\rangle}. \tag{15}$$

Equation (14) leads to an iteration that maximizes the likelihood. The iteration reads

$$\rho_{i+1} = \frac{K_i\rho_i K_i}{\mathrm{Tr}[K_i\rho_i K_i]} \tag{16}$$

starting from $\rho_0 = \mathbb{I}/2$ and stopping when the change in one step gets small enough, $\mathrm{Tr}|\rho_{i+1} - \rho_i|^2 < \epsilon$. The derivations of the extremal and iterative equations are covered in Ref. [48,50,51]. So far, we have used projection measurement operators $\Pi_i = |\pi_i\rangle\langle\pi_i|$, namely $|H\rangle\langle H|, \ldots, |L\rangle\langle L|$. Let us note that the polarization tomography is not restricted to these particular projections and *any tomographically complete* set of positive measurement operators (POVM) is applicable. The method also applies to single-photon signals, for which the detected intensity I_i is replaced with the recorded number of detection events in a given projection. We provide a Python script for the maximum likelihood reconstruction on GitHub [38]. Examples of waveplate-based and TNLC-based polarization tomography are also included.

We test the TNLC polarimeter by comparing its readings to the known input polarization states. We prepared the input polarization states using waveplates and measured them using the tested TNLC polarimeter. Namely, we prepared the six eigenstates of Pauli operators. The TNLC polarimeter was calibrated to perform the projections onto Pauli eigenstates as well as the four states required for minimal tomography [29]. As a reference, we replaced the TNLC polarimeter with a pair of waveplates and performed the standard six-state tomography.

The resulting fidelities and angular deviations averaged over the six tested states are summarized in Table 3. The individual sample states analyzed using the presented TNLC device with the six-state tomography and the four-state tomography are visualized in Fig. 7. The six-state tomography outperforms the four-state minimal tomography consistently. The results produced by the TNLC polarimeter agree excellently with the results provided by a reference waveplate-based polarimeter. Note that the used reference polarimeter is not the same polarimeter we used for TNLC device calibration.

Now we show that the presented TNLC polarimeter outperforms the equivalent scheme with waveplates in terms of speed. The ordering of the projections in the tomographic measurement does not affect the result of the measurement but influences its duration [52]. Based on the transition time characterization (Table 2), we can select the optimum sequence ordering of the six-state protocol, which is H, R, L, A, D, V, and back to H to prepare the TNLC device for another tomography cycle. The transition of the optimum sequence takes 1.17 s. For the minimal tomography [29], only four states are required. The transitions times between them are longer than in the six-state scheme. It takes approximately 1.2 s to cycle through these four states. Consequently, we recommend to using six-state tomography with the presented device.

Optics EXPRESS

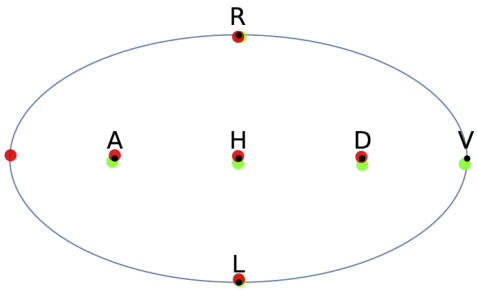


Fig. 7. The Bloch sphere with H, V, D, A, R, and L states prepared by waveplates and analyzed by the TNLC device using six-state tomography (red dots) and four-state minimum tomography (green points). Black points show the ideal position of the states.

Table 3. Achieved average results of the fidelity and angular deviation between the ideal and measured sample polarization states prepared by the reference polarimeter. The reported TNLC polarimeter in six-state and four-state configurations, and the reference wave-plate based polarimeter in six-state regime are compared.

analysis method	TNLC device		wave plates
	six-state conf.	four-states conf.	six-state conf.
fidelity	0.9996(4)	0.998(2)	0.9995(5)
angular deviation [deg]	1.0(2)	1.3(5)	1.0(5)

We compare these values with the total time required by our reference waveplate-based polarimeter. The reference polarimeter uses motorized mounts with a typical speed of 20 deg/s (Newport PR50CC), and the optimum sequence reads H, L, A, R, V, D, and back to H. The sequence takes 12.9 s, which is an order of magnitude slower than using the TNLC device. The optimal sequence ordering can significantly speed up the whole measurement process, especially for multi-mode analysis or quantum circuits with more qubits [52]. Alternatively, all projections can be measured at the same time [30,53] at the expense of a number of detectors and the complexity of the experimental setup.

To show an application of the reported TNLC device at the single-photon level, we used it in quantum tomography of a multi-qubit photonic entangled state. The TNLC device was calibrated in a six-state configuration using the strong laser signal, as described in Sec. 4. The TNLC device was then used to perform tomographic projections on the third output qubit of a three-qubit linear optical controlled-Z gate (CCZ gate). The gate is equivalent to the Toffoli gate up to single-qubit Hadamard transform on a target qubit [39,54–56]. The Toffoli gate is a crucial part of many quantum information processing schemes [54]. The photonic circuit of the gate and its characterization were presented in detail elsewhere [57]. The gate can produce entangled quantum states such as tripartite Greenberger-Horne-Zeilinger (GHZ) state [58,59]. We used the circuit to prepare the GHZ state $(|HHD\rangle + |VVA\rangle)/\sqrt{2}$ and performed its full tomographic characterization using the TNLC device. The resulting density matrix of the three-qubit state is shown in Fig. 8(a,b). The fidelity of the measured state and the ideal GHZ state is 0.9362. For comparison, we characterized the state also using a common waveplate-based tomography and obtained the fidelity to the ideal state 0.9360; the resulting density matrix is shown in Fig. 8(c,d). The respective fidelity of the retrieved GHZ states measured using the TNLC device and the wave-plate polarimeter reaches 0.971. The small resulting discrepancy is caused by the limited repeatability of the experimental setup. When changing the TNLC device for the waveplates, the setup had to be partly reassembled and realigned. Due to the circuit construction, this realignment

Optics EXPRESS

might lead to small changes in polarization-dependent losses. Furthermore, the interferometric phase in the experimental setup was slowly drifting, and the phase drift might result in small changes between two experimental runs.

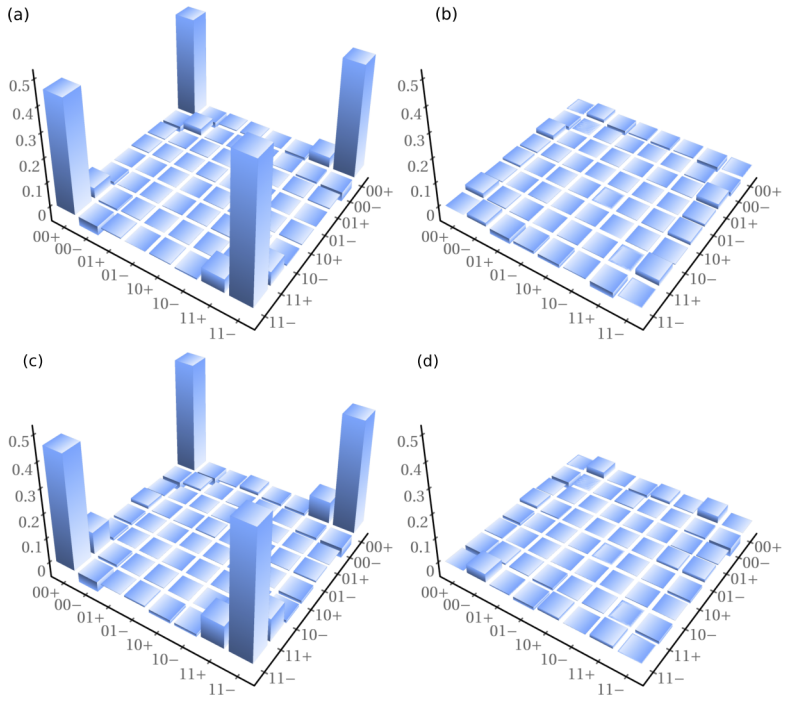


Fig. 8. Real (a) and imaginary (b) part of the GHZ state density matrix measured using the TNLC device. Real (c) and imaginary (d) part of the GHZ state density matrix measured using wave plates. Here 0, 1, +, and − denote H, V, D, and A polarizations, respectively.

7. Conclusion

We have presented a polarimetric device assembled from a commercially available TNLC segment display with minimal modifications. We have demonstrated a universal method of the TNLC device calibration utilizing a genetic algorithm. Various polarization states have been prepared using the TNLC device and characterized with a reference polarimeter with the average fidelity exceeding 0.999(1). Particularly, we have demonstrated a highly accurate generation of four states forming the vertices of a regular tetrahedron inscribed into the Bloch sphere [29], six eigenstates of Pauli operators (horizontal, vertical, diagonal, anti-diagonal, right-handed and left-handed circular polarizations), and 120 states uniformly distributed on the Bloch Sphere. We also have generated a completely depolarized state with the purity 0.5004(3) and degree of polarization 0.03(1). Using the presented calibration procedure, we can generate an arbitrary state of polarization.

Furthermore, we have employed the reported TNLC device as a polarimeter and demonstrated accurate measurement of polarization of light. Physically correct density matrix of the polarization state is retrieved using the maximum likelihood iterative algorithm. We have tested two tomographic schemes, namely four-projection minimal tomography and overdetermined six-projection scheme based on eigenstates of the Pauli matrices. We have characterized and optimized the precision and speed of the developed polarimeter. The TNLC device has been successfully used also for polarimetry of optical signals at the single-photon level. Particularly,

Optics EXPRESS

we have performed the full quantum tomography of the three-qubit Greenberger–Horne–Zeilinger entangled state produced by photonic quantum Toffoli gate. The results agree with those obtained using common wave-plate polarimetry.

To conclude, we have demonstrated the calibration and application in polarimetry of virtually arbitrary *twisted* nematic liquid crystal segment displays. The approach allows highly accurate preparation and measurement of an arbitrary polarization state of light using low-voltage driving and no moving parts. The approach is fully scalable to many optical modes at the single-photon level. It can find applications in a wide range of fields, such as polarization-sensitive biomedical imaging or photonic quantum information processing.

The complete technical documentation of the presented TNLC device, including tomography software, is published on GitHub [38].

Funding. Ministerstvo Školství, Mládeže a Tělovýchovy (8C18002); Univerzita Palackého v Olomouci (IGA-PrF-2020-009, IGA-PrF-2021-006); Grantová Agentura České Republiky (17-26143S); Horizon 2020 Framework Programme (8C18002, HYPER-U-P-S).

Acknowledgments. We thank Michal Dudka for developing electronic driver circuitry. Project HYPER-U-P-S has received funding from the QuantERA ERANET Cofund in Quantum Technologies implemented within the European Union's Horizon 2020 Programme.

Disclosures. The authors declare that there are no conflicts of interest related to this article.

Data availability. Data underlying the results presented in this paper are not publicly available at this time, but they will be available on the GitHub repository [38] and may be obtained from the authors upon request.

References

1. S. Huard, *Polarization of Light* (Wiley, 1997).
2. G. Chen, Z.-Q. Wen, and C.-W. Qiu, "Superoscillation: from physics to optical applications," Light: Sci. Appl. **8**(1), 56 (2019).
3. A. Van Eeckhout, E. Garcia-Caurel, T. Garnatje, M. Durfort, J. C. Escalera, J. Vidal, J. J. Gil, J. Campos, and A. Lizana, "Depolarizing metrics for plant samples imaging," PLoS One **14**(3), e0213909 (2019).
4. K. Zhanghao, X. Chen, W. Liu, M. Li, Y. Liu, Y. Wang, S. Luo, X. Wang, C. Shan, H. Xie, J. Gao, X. Chen, D. Jin, X. Li, Y. Zhang, Q. Dai, and P. Xi, "Super-resolution imaging of fluorescent dipoles via polarized structured illumination microscopy," Nat. Commun. **10**(1), 4694 (2019).
5. Z.-Y. Chen, L.-S. Yan, Y. Pan, L. Jiang, A.-L. Yi, W. Pan, and B. Luo, "Use of polarization freedom beyond polarization-division multiplexing to support high-speed and spectral-efficient data transmission," Light: Sci. Appl. **6**(2), e16207 (2017).
6. F. Flamini, N. Spagnolo, and F. Sciarrino, "Photonic quantum information processing: a review," Rep. Prog. Phys. **82**(1), 016001 (2019).
7. S. Slussarenko and G. J. Pryde, "Photonic quantum information processing: A concise review," Appl. Phys. Rev. **6**(4), 041303 (2019).
8. N. Spagnolo, C. Vitelli, S. Giacomini, F. Sciarrino, and F. D. Martini, "Polarization preserving ultra fast optical shutter for quantum information processing," Opt. Express **16**(22), 17609–17615 (2008).
9. Y. He, X. Ding, Z.-E. Su, H.-L. Huang, J. Qin, C. Wang, S. Unsleber, C. Chen, H. Wang, Y.-M. He, X.-L. Wang, W.-J. Zhang, S.-J. Chen, C. Schneider, M. Kamp, L.-X. You, Z. Wang, S. Höfling, C.-Y. Lu, and J.-W. Pan, "Time-Bin-Encoded Boson Sampling with a Single-Photon Device," Phys. Rev. Lett. **118**(19), 190501 (2017).
10. S. Takeda, K. Takase, and A. Furusawa, "On-demand photonic entanglement synthesizer," Sci. Adv. **5**(5), eaaw4530 (2019).
11. J. Tiedau, E. Meyer-Scott, T. Nitsche, S. Barkhofen, T. J. Bartley, and C. Silberhorn, "A high dynamic range optical detector for measuring single photons and bright light," Opt. Express **27**(1), 1–15 (2019).
12. J. B. Altepeter, N. N. Oza, M. Medić, E. R. Jeffrey, and P. Kumar, "Entangled photon polarimetry," Opt. Express **19**(27), 26011–26016 (2011).
13. J. M. Bueno, "Polarimetry using liquid-crystal variable retarders: theory and calibration," J. Opt. A: Pure Appl. Opt. **2**(3), 216–222 (2000).
14. A. De Martino, Y.-K. Kim, E. Garcia-Caurel, B. Laude, and B. Drévillon, "Optimized Mueller polarimeter with liquid crystals," Opt. Lett. **28**(8), 616–618 (2003).
15. A. Peinado, A. Lizana, J. Vidal, C. Iemmi, and J. Campos, "Optimization and performance criteria of a Stokes polarimeter based on two variable retarders," Opt. Express **18**(10), 9815–9830 (2010).
16. A. Peinado, A. Lizana, J. Vidal, C. Iemmi, and J. Campos, "Optimized Stokes polarimeters based on a single twisted nematic liquid-crystal device for the minimization of noise propagation," Appl. Opt. **50**(28), 5437–5445 (2011).
17. O. Aharon and I. Abdulhalim, "Liquid crystal Lyot tunable filter with extended free spectral range," Opt. Express **17**(14), 11426–11433 (2009).

Optics EXPRESS

18. Y. August and A. Stern, "Compressive sensing spectrometry based on liquid crystal devices," Opt. Lett. **38**(23), 4996–4999 (2013).
19. Z. Zhuang, S.-W. Suh, and J. S. Patel, "Polarization controller using nematic liquid crystals," Opt. Lett. **24**(10), 694–696 (1999).
20. I. Moreno, J. L. Martínez, and J. A. Davis, "Two-dimensional polarization rotator using a twisted-nematic liquid-crystal display," Appl. Opt. **46**(6), 881–887 (2007).
21. A. Safrani and I. Abdulhalim, "Liquid-crystal polarization rotator and a tunable polarizer," Opt. Lett. **34**(12), 1801–1803 (2009).
22. A. Peinado, A. Lizana, and J. Campos, "Use of ferroelectric liquid crystal panels to control state and degree of polarization in light beams," Opt. Lett. **39**(3), 659–662 (2014).
23. A. S. Rab, E. Polino, Z.-X. Man, N. B. An, Y.-J. Xia, N. Spagnolo, R. L. Franco, and F. Sciarrino, "Entanglement of photons in their dual wave-particle nature," Nat. Commun. **8**(1), 915 (2017).
24. A. Lumino, E. Polino, A. S. Rab, G. Milani, N. Spagnolo, N. Wiebe, and F. Sciarrino, "Experimental Phase Estimation Enhanced by Machine Learning," Phys. Rev. Appl. **10**(4), 044033 (2018).
25. A. Lohrmann, C. Perumgatt, and A. Ling, "Manipulation and measurement of quantum states with liquid crystal devices," Opt. Express **27**(10), 13765–13772 (2019).
26. X.-j. Wang, Z.-d. Huang, J. Feng, X.-f. Chen, X. Liang, and Y.-q. Lu, "Liquid crystal modulator with ultra-wide dynamic range and adjustable driving voltage," Opt. Express **16**(17), 13168–13174 (2008).
27. G. Zhu, B. yan Wei, L. yu Shi, X. wen Lin, W. Hu, Z. di Huang, and Y. qing Lu, "A fast response variable optical attenuator based on blue phase liquid crystal," Opt. Express **21**(5), 5332–5337 (2013).
28. C. Perumangatt, T. Vergoossen, A. Lohrmann, S. Sivasankaran, A. Reezwana, A. Anwar, S. Sachidananda, T. Islam, and A. Ling, "Realizing quantum nodes in space for cost-effective, global quantum communication: in-orbit results and next steps," in *Quantum Computing, Communication, and Simulation*, P. R. Hemmer and A. L. Migdall, eds. (SPIE, 2021).
29. J. Řeháček, B.-G. Englert, and D. Kaszlikowski, "Minimal qubit tomography," Phys. Rev. A **70**(5), 052321 (2004).
30. A. Ling, K. P. Soh, A. Lamas-Linares, and C. Kurtsiefer, "Experimental polarization state tomography using optimal polarimeters," Phys. Rev. A **74**(2), 022309 (2006).
31. M. D. de Burgh, N. K. Langford, A. C. Doherty, and A. Gilchrist, "Choice of measurement sets in qubit tomography," Phys. Rev. A **78**(5), 052122 (2008).
32. A. Ling, A. Lamas-Linares, and C. Kurtsiefer, "Accuracy of minimal and optimal qubit tomography for finite-length experiments," arXiv:0807.0991 (2008).
33. Yu. I. Bogdanov, G. Brida, M. Genovese, S. P. Kulik, E. V. Moreva, and A. P. Shurupov, "Statistical Estimation of the Efficiency of Quantum State Tomography Protocols," Phys. Rev. Lett. **105**(1), 010404 (2010).
34. Yu. I. Bogdanov, G. Brida, I. D. Bukeev, M. Genovese, K. S. Kravtsov, S. P. Kulik, E. V. Moreva, A. A. Soloviev, and A. P. Shurupov, "Statistical estimation of the quality of quantum-tomography protocols," Phys. Rev. A **84**(4), 042108 (2011).
35. D. Koutný, Y. S. Teo, Z. Hradil, and J. Řeháček, "Fast universal performance certification of measurement schemes for quantum tomography," Phys. Rev. A **94**(2), 022113 (2016).
36. A. Marquez, J. Campos, M. J. Yzuel, I. S. Moreno, J. A. Davis, C. C. Iemmi, A. Moreno, and A. Robert, "Characterization of edge effects in twisted nematic liquid crystal displays," Opt. Eng. **39**(12), 3301–3307 (2000).
37. M. Yamauchi, "Jones-matrix models for twisted-nematic liquid-crystal devices," Appl. Opt. **44**(21), 4484–4493 (2005).
38. M. Bielak, "Github repository," https://github.com/BielakM/polarimeter (2020).
39. B. P. Lanyon, M. Barbieri, M. P. Almeida, T. Jennewein, T. C. Ralph, K. J. Resch, G. J. Pryde, J. L. O'Brien, A. Gilchrist, and A. G. White, "Simplifying quantum logic using higher-dimensional Hilbert spaces," Nat. Phys. **5**(2), 134–140 (2009).
40. M. Mičuda, M. Sedlák, I. Straka, M. Miková, M. Dušek, M. Ježek, and J. Fiurášek, "Efficient Experimental Estimation of Fidelity of Linear Optical Quantum Toffoli Gate," Phys. Rev. Lett. **111**(16), 160407 (2013).
41. L. Xiao, X. Zhan, Z. H. Bian, K. K. Wang, X. Zhang, X. P. Wang, J. Li, K. Mochizuki, D. Kim, N. Kawakami, W. Yi, H. Obuse, B. C. Sanders, and P. Xue, "Observation of topological edge states in parity–time-symmetric quantum walks," Nat. Phys. **13**(11), 1117–1123 (2017).
42. A. Yariv and A. P. Yeh, *Optical waves in crystals: propagation and control of laser radiation* (Wiley, 1984).
43. J. A. Davis, I. Moreno, and P. Tsai, "Polarization eigenstates for twisted-nematic liquid-crystal displays," Appl. Opt. **37**(5), 937–945 (1998).
44. J. P. P. Snyder, *Flattening the Earth: Two Thousand Years of Map Projections* (University of Chicago Press, 1993).
45. A. E. Eiben and J. E. Smith, *Introduction to Evolutionary Computing* (Springer, 2015).
46. L. M. Rios and N. V. Sahinidis, "Derivative-free optimization: a review of algorithms and comparison of software implementations," J. Global Optim. **56**(3), 1247–1293 (2013).
47. "Standard Uncertainty and Relative Standard Uncertainty," (2020). [Online; accessed 2. Jul. 2020].
48. Z. Hradil, "Quantum-state estimation," Phys. Rev. A **55**(3), R1561–R1564 (1997).
49. D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, "Measurement of qubits," Phys. Rev. A **64**(5), 052312 (2001).
50. M. Ježek, J. Fiurášek, and Z. Hradil, "Quantum inference of states and processes," Phys. Rev. A **68**(1), 012305 (2003).

Optics EXPRESS

51. Z. Hradil, J. Řeháček, J. Fiurášek, and M. Ježek, "3 Maximum-Likelihood Methodsin Quantum Mechanics," in *Quantum State Estimation*, (Springer, 2004), pp. 59–112.
52. R. Hošák, R. Stárek, and M. Ježek, "Optimal reordering of measurements for photonic quantum tomography," Opt. Express **26**(25), 32878–32887 (2018).
53. I. Estévez, V. Sopo, A. Lizana, A. Turpin, and J. Campos, "Complete snapshot Stokes polarimeter based on a single biaxial crystal," Opt. Lett. **41**(19), 4566–4569 (2016).
54. M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).
55. T. Monz, K. Kim, W. Hänsel, M. Riebe, A. S. Villar, P. Schindler, M. Chwalla, M. Hennrich, and R. Blatt, "Realization of the Quantum Toffoli Gate with Trapped Ions," Phys. Rev. Lett. **102**(4), 040501 (2009).
56. A. Fedorov, L. Steffen, M. Baur, M. P. da Silva, and A. Wallraff, "Implementation of a Toffoli gate with superconducting circuits," Nature **481**(7380), 170–172 (2012).
57. R. Stárek, M. Mičuda, M. Miková, I. Straka, M. Dušek, P. Marek, M. Ježek, R. Filip, and J. Fiurášek, "Nondestructive detector for exchange symmetry of photonic qubits," npj Quantum Inf. **4**(1), 35 (2018).
58. D. M. Greenberger, M. A. Horne, and A. Zeilinger, "Going Beyond Bell's Theorem," in *Bell's Theorem, Quantum Theory and Conceptions of the Universe*, (Springer, 1989), pp. 69–72.
59. J.-W. Pan, D. Bouwmeester, M. Daniell, H. Weinfurter, and A. Zeilinger, "Experimental test of quantum nonlocality in three-photon Greenberger–Horne–Zeilinger entanglement," Nature **403**(6769), 515–519 (2000).

Bidirectional Deep Learning of Polarization Transfer in Liquid Crystals with Application to Quantum State Preparation

Dominik Vašinka,* Martin Bielak, Michal Neset, and Miroslav Ježek†

[1] *Department of Optics, Palacký University, 17. listopadu 12, Olomouc 77146, Czechia*

(Received 7 October 2021; revised 10 March 2022; accepted 18 April 2022; published 25 May 2022)

Accurate control of light polarization represents a core building block in polarization metrology, imaging, and optical and quantum communications. Voltage-controlled liquid crystals offer an efficient way of polarization transformation. However, common twisted nematic liquid crystals are notorious for lacking an accurate theoretical model linking control voltages and output polarization. An inverse model, which would predict control voltages required to prepare a target polarization, is even more challenging. Here we report both the direct and inverse models based on deep neural networks, radial basis functions, and linear interpolation. We present an inverse–direct compound model solving the problem of control voltages ambiguity. We demonstrate an order of magnitude improvement in accuracy using deep learning compared to the radial basis function method and 2 orders of magnitude improvement compared to the linear interpolation. Errors of the deep neural network model also decrease faster than the other methods with an increasing number of training data. The best direct and inverse models reach average infidelities of 4×10^{-4} and 2×10^{-4}, respectively, which are accuracy levels not reported yet. Furthermore, we demonstrate local and remote preparation of an arbitrary single-photon polarization state using deep learning models. The results will impact the application of twisted-nematic liquid crystals, increasing their control accuracy across the board. The presented bidirectional learning can be used for optimal classical control of complex photonic devices and quantum circuits beyond interpolation.

DOI: 10.1103/PhysRevApplied.17.054042

I. INTRODUCTION

A vast number of applications require accurate control of the polarization state of light, such as in display technology, ellipsometry, polarization microscopy, and optical communications. The polarization can be manipulated by mechanically adjusting birefringent elements or using electro-optic modulators. The former approach can reach high accuracy but is slow, producing vibrations, and prone to malfunction. The latter approach is ultrafast and vibration-free; however, its accuracy is limited. Free-space electro-optic modulators are bulky and require high-voltage drivers, which tend to fluctuate and decrease the overall accuracy even further. Low-voltage integrated modulators show inherent polarization instabilities due to fiber coupling.

Liquid crystals represent a middle ground between the stability of birefringent elements and the response speed of electro-optic modulators. Voltage-controlled nematic liquid crystals are particularly convenient and widely available, as they are commonly used in the display industry. We distinguish two main types of nematic liquid crystals based on the alignment of the crystals in a device, namely parallel and twisted configurations. The former acts as a polarization retarder and is typically custom-made for specialized applications; the latter is used in displays. Parallel nematic liquid crystals were utilized as polarization retarders for polarization modulation [1], polarization state preparation and tomography [2,3], remote state preparation and imaging [4,5], entangled-photon generation [6,7], and implementation of quantum channels [8,9] and quantum communication protocols [10,11]. Recently, fully reconfigurable topological photonic devices have been proposed employing nematic liquid crystals [12].

Despite the wide utilization of nematic liquid crystal (LC) devices, we lack an accurate theoretical model of LCs. The available models are particularly inaccurate for twisted LCs, which prevents them from entering a more extensive range of applications. The response of LC to control voltage(s) is affected by various imperfections such as alignment layers dragging, multiple reflections, and inhomogeneity-induced depolarization. These effects represent a serious setback to the modeling of the LC response and, particularly, to the inverse task of finding the optimum control voltages to prepare the target polarization state. The theoretical model of twisted LCs [13] was modified to include the boundary effects [14] and yet further adjusted

*vasinkadominik@seznam.cz
†jezek@optics.upol.cz

[15–17] to achieve better results. Unfortunately, even with these improvements, the theoretical model of polarization transformation is not sufficiently accurate. The inaccuracy is especially apparent when considering a device consisting of multiple LC cells [1,18]. When aiming for a discrete set of polarization states, such a complex LC device can be calibrated and even used as a highly accurate polarimeter [18]. However, modeling a continuous polarization response of a LC device to analog control voltages represents an open problem.

Here we model a complex twisted LC device using deep neural networks (DNNs). The model is fitted to a training dataset obtained by measuring the polarization states prepared by a device for different control voltages. We employ the mesh adaptive direct search (MADS) algorithm for black-box optimization of the deep learning model hyperparameters. The model is optimized and tested using separate datasets not involved in the training process. We demonstrate an unprecedented fidelity and repeatability between the polarization state predicted by the model and the measured state. We achieve an average infidelity of 4×10^{-4} of the polarization preparation for a three-cell twisted LC device. The DNN approach outperforms other models based on radial basis functions (RBFs) and linear interpolation. We study the effect of the training dataset size over several orders of magnitude and find that errors of the DNN model decrease faster than the other numerical methods with an increasing number of training data. In other words, the DNN approach is more efficient with respect to the dataset size than other methods. Furthermore, we analyze the DNN model size and its overparametrization and scalability.

Our main result is solving the inverse task of finding control voltages optimal for preparing a target polarization state. We utilize the trained DNN direct model as part of the compound autoencoderlike network to find the inverse model. This allows the use of physics metrics, e.g., the fidelity, consistently in the whole framework and also avoids ambiguous mapping from polarization state to control voltages. The compound model outperforms other approaches by orders of magnitude. We verify the predictive strength of the DNN model by preparing over a thousand single-photon polarization states and performing their independent characterization using full quantum tomography. Finally, we demonstrate a remote preparation of quantum states using entangled photons. The reported local and remote preparation of polarization-encoded quantum bits (qubits) certifies the use of twisted LC devices in quantum technology.

Besides the imminent application of our approach to accurate polarization qubit manipulation, we may think of it as a use case of a more general problem of optimal control of quantum devices; see Fig. 1(a). Various implementations of quantum devices share the common aspect of being controlled by classical analog signals, related

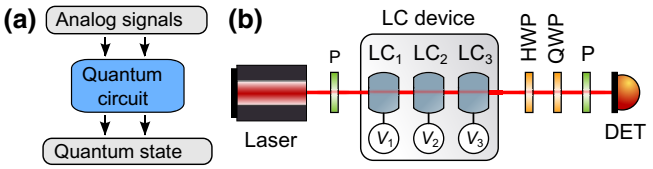


FIG. 1. (a) Classical control of a quantum device: classical analog signals modify the quantum circuit operation and affect the transformation or preparation of quantum states. (b) Experimental setup used to create a dataset of control voltages with corresponding polarization states: the polarizer (P) prepares a horizontal state H, and the liquid cells ($LC_{1,2,3}$) induce the transformation based on applied voltages V_1, V_2, V_3. The prepared polarization state is analyzed using a reference polarimeter consisting of a half-wave plate (HWP), a quarter-wave plate (QWP), a polarizer, and a detector (DET).

nontrivially to the device operation [19–21]. The control signals need to be optimally adjusted to provide high-fidelity operation of the device [22–28]. Photonic circuits on optical chips include voltage-controlled phase modulators with a complex response and crosstalk [29–31]. Superconducting circuits are controlled by radio-frequency signals with variable amplitudes and complex timing, which are subject to optimization [32,33]. Also, semiconductor quantum dots need to be optimally tuned to produce target states [34–40]. The approach developed in this work can be directly applied to the learning of the steady-state response of quantum devices to classical control signals and, consequently, their optimal control.

II. LC DEVICE AND EXPERIMENTAL SETUP

The experimental setup implementing the polarization state transformation using liquid crystals is depicted in Fig. 1(b). Light from an 810 nm continuous-wave laser propagates through a horizontal polarizer preparing the input polarization state H. A device based on twisted nematic liquid crystals (LC device), described below, induces a polarization transformation on the input state controlled by an applied electric field. The prepared polarization state is characterized using full quantum tomography. The polarization state is projected into six basis states H, V, D, A, R, and L using a reference polarimeter. It consists of half-wave and quarter-wave plates followed by a polarizer and a silicon photodiode detector. The quantum state is reconstructed from relative measurement frequencies using the maximum likelihood method [41,42].

The LC device consists of three independent LC cells extracted from commercially available twisted nematic displays *Lumex LCD-S101D14TR* by removing auxiliary layers [18]. LCs are enclosed between glass plates with deposited electrodes producing an electric field and alignment layers forming the LC twist. The polarization transformation induced by each cell is controlled by applying

voltage signals to the electrodes to manipulate the spatial orientation of LCs. We used a square wave with a 1 kHz frequency and a 50% duty cycle, whose amplitude ranges from 0 to 10 V p.p. (volts peak to peak). We refer to this amplitude as a control voltage. Connecting three independent LC cells allows us to prepare an arbitrary polarization state by inducing the polarization transformation on a horizontal input state H, i.e., by applying the proper control voltages V_1, V_2, and V_3 to the LC device.

Using this setup, we created a dataset consisting of 27 000 combinations of three control voltages with a corresponding prepared polarization state. The average purity of these polarization states is $(99.7 - 1 + 0.3)\%$, referring to the average value with 5th and 95th percentiles. We divided the randomly shuffled dataset into three parts—training set, validation set, and test set. All discussed models are trained on the training set containing 16 000 data samples to learn the desired mapping. The validation set of 6500 data samples is utilized for optimizing the models' hyperparameters, for example, the architecture of a neural network or a type of radial basis function. And finally, the test set containing the remaining data samples is used to evaluate the models' generalization ability on data never seen before. We note that the reference measurement induces certain infidelity to the dataset compared to the ground truth states. Therefore, as all models' predictions are evaluated compared to this dataset, the error between the ground truth states and the model predictions is a combination of the reference measurement error [independently estimated to be 0.05(5)%] and the error of the model.

III. DIRECT MODELS AND DATASET-SIZE SCALING

First, we report on predicting the prepared polarization state given the three control voltages, i.e., modeling the direct transformation. Each polarization state is described using a density matrix ρ. To ensure that all predictions fulfill the physical requirements on a density matrix, we utilize the Cholesky decomposition. For a Hermitian positive semidefinite matrix M, the decomposition has the form $M = \tau\tau^\dagger$, where τ is a lower triangular matrix with real and positive diagonal elements. Trained models output elements of the τ matrix, which are reconstructed into a Hermitian and positive definite matrix M. This matrix is then normalized into a physically sound density matrix with unity trace, $\rho = M/\mathrm{Tr}[M]$. To measure the closeness between two polarization states $\rho_{1,2}$, we use fidelity calculated as $F = (\mathrm{Tr}[\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}])^2$, whose values range from 0 to 1, and the infidelity obtained as $1 - F$.

The developed deep learning model is based on a fully connected deep neural network transforming three control voltages V_1, V_2, V_3 into four real-valued parameters of the τ matrix (two real diagonal and one complex off diagonal). We performed a hyperparameter optimization of the deep learning model, utilizing the MADS algorithm implemented in the black-box optimization software Nomad [43,44]. The optimized parameters include the number of hidden layers, number of neurons in the hidden layer, batch size, initial learning rate, and dropout regularization. The optimum dropout rate is consistently found to be zero. The optimal network consists of more than 400 000 trainable parameters arranged into 19 hidden layers with 156 neurons in each layer. We use a rectified linear unit as an activation function in each hidden layer and a linear activation function in the output layer. We train the model using Adam [45] as a stochastic gradient descend optimizer to minimize the mean squared error (MSE) loss function. Having a significantly higher number of trainable parameters than samples in the training dataset does not pose a problem in the DNN context. Large neural networks are biased towards simpler solutions [46,47], which allows the use of complex network architectures without overfitting.

Referring to the average value with the 5th and 95th percentiles, we achieve an infidelity of $(4 - 4 - 1) \times 10^{-4}$ of target polarization preparation. Here the average value lies outside the 5th–95th percentiles interval due to outliers having larger infidelity values. The DNN model is 2 orders of magnitude more accurate than linear interpolation with an infidelity of $(1 - 1 + 1) \times 10^{-2}$. We also compare the DNN model to RBF interpolation, an approximation method based on a weighted sum of radial basis functions [48]. The output variable y_i is given by $\sum_j a_{ij}\varphi(\|x - x_j\|)$, where x is a vector of input variables and x_j is a fixed input data point termed a center. The value of a radial basis function φ depends only on the distance from the corresponding center. A set of differently centered functions of the same type forms the basis. We tested linear, cubic, quintic, multiquadric, inverse multiquadric, and Gaussian functions. The weights a_{ij} are adjusted for a minimum error on the training set. Using the validation set, we find the optimum basis function to be cubic. The resulting RBF model achieves an infidelity of $(2 - 2 + 2) \times 10^{-3}$, which is an order of magnitude less accurate than the DNN model. The left side of Table I summarizes these results, together with the average computational time per data sample and a single CPU core. The DNN model is significantly faster in its predictions when compared to the other two methods.

Furthermore, we study the dependence of the direct models' infidelities on the number of samples in the training set; see Fig. 2(a). We evaluated each point in the chart following the same procedure. First, we divided the training set into smaller disjoint subsets with the same number of samples. Each subset is used to train all three models—linear interpolation, RBF interpolation, and DNN. We used the whole validation set to optimize the

TABLE I. The comparison of infidelity and computational time for linear interpolation, radial basis function interpolation, and the deep neural network. The infidelity values evaluated on the test dataset refer to the average, the 5th percentile, and the 95th percentile.

	Direct model		Compound model	
	Infidelity	Time per sample (s)	Infidelity	Time per sample (s)
Linear interpolation	$(1-1+1)\times10^{-2}$	4×10^{-3}	$(1-1+1)\times10^{-2}$	3×10^{-3}
Radial basis function	$(2-2+2)\times10^{-3}$	4×10^{-4}	$(5-5+30)\times10^{-2}$	5×10^{-4}
Deep neural network	$(4-4-1)\times10^{-4}$	8×10^{-5}	$(2-2+5)\times10^{-4}$	1×10^{-4}

hyperparameters and choose the best model for each method. The infidelities of the best models are evaluated on the test set and visualized in the chart. Each curve then represents the results of the best possible model given the number of samples in the training set. Not only does the deep learning model achieve the lowest errors and decrease faster with the size of the training set compared to both other methods, but it also saturates at errors lower by orders of magnitude, as depicted in Fig. 2(a). The results can also be read as the deep learning model needs an order of magnitude lower number of experimental measurements to characterize the device with the required accuracy.

So far, we targeted the best performing DNN model using the process of hyperparameters optimization without limiting the size of the model. The optimal model lies within an overparametrized regime of the double-descent curve [46,47]. The ratio of the number of trainable parameters to the number of training samples, termed overparametrization, is shown in Fig. 2(b) for the optimal DNN models trained with various numbers of training samples. The overparametrization reaches dozens of hundreds for very small training sets and decreases to approximately 50 for the large training sets. As seen, one can reduce the overparametrization to some extent by using larger datasets. Moreover, we explored suboptimal DNNs by repeatedly lowering the number of trainable parameters of the optimal DNN by a factor of 0.5 while keeping the ratio of neurons per layer to the number of hidden layers approximately constant. The suboptimal DNNs are trained in the same way as the optimum one except for hyperparameter optimization. As shown in Fig. 2(c), we can significantly lower the number of trainable parameters while only slightly decreasing the accuracy. In particular, the suboptimal DNN model with 13 800 trainable parameters reaches an infidelity of $(7-7+6)\times10^{-4}$, which represents only a slight increase compared with the optimal DNN with an infidelity of $(4-4-1)\times10^{-4}$ and more than 400 thousand parameters. We also found the smallest suboptimal DNN performing at the same level of infidelity as the best RBF model utilizing 16 000 parameters. For this, the DNN requires only 3500 trainable parameters. It seems that deep fully connected networks are much more efficient and provide better scalability than interpolation approaches, including the RBF method.

IV. INVERSE AND COMPOUND MODEL

Our main goal is predicting control voltages required to prepare the target polarization state by the LC device, i.e., modeling the inverse transformation. Because of ambiguous mapping between polarization state and control voltages, the direct model cannot be easily inverted. Also, no physically fundamental or otherwise preferred metric exists in the space of the three control voltages (classical control signals in general), but the resulting inverse model would depend on the chosen metric. To eliminate the ambiguity, we train the inverse deep learning model by connecting it to the already optimized direct model, creating an autoencoderlike compound network, as visualized in Fig. 3. Only the inverse part is trainable during the learning process, while the parameters of the direct model stay fixed. Not only does this solve the ambiguous mapping, but allows us to evaluate the performance using fidelity between the desired input state and the predicted output state. We note that the evaluated (in)fidelity describes the compound model as a whole, not just the inverse section. However, given the negligible error of the direct model, the compound model's fidelity can be viewed as an approximated accuracy of the inverse model.

The inverse section of the compound model is a fully connected DNN, which takes four elements of the polarization state density matrix for its input and predicts the three control voltages. These voltages are then transformed by the fixed direct model back into a polarization state. Utilizing the MADS algorithm, we optimized hyperparameters of the inverse part of the compound model. With 91 neurons in each of the 14 hidden layers, the optimum inverse model contains more than 100 000 trainable parameters. Each hidden layer uses a rectified linear unit as an activation function, whereas the output layer utilizes the sigmoid activation function to predict control voltages rescaled to range from 0 to 1. Using the Adam optimizer and MSE loss function, the optimized compound model achieves an infidelity of $(2-2+5)\times10^{-4}$ between the input and predicted output states, evaluated on the test set.

In the same way as for DNNs, the inverse model of RBF interpolation (and linear interpolation) is a mapping from elements of the density matrix to the three control voltages. This inverse model is connected to the pretrained and optimized direct RBF model (direct linear interpolation

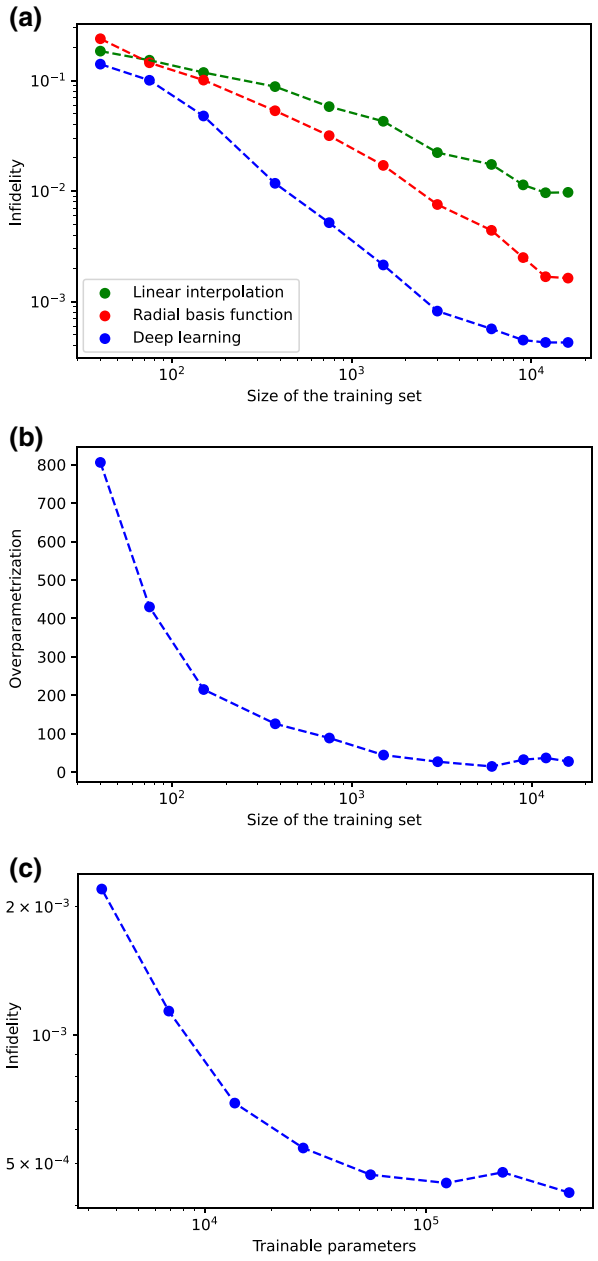


FIG. 2. (a) The dependence of the optimal models' test set infidelity on the number of samples in the training set. The DNN models (blue) not only achieve the lowest errors but improve faster with the size of the training dataset than linear interpolation (green) and RBF interpolation (red). Furthermore, all three methods seem to saturate at certain infidelities for larger datasets, with the deep learning model saturating at the lowest error. (b) The ratio of the number of trainable parameters to the training set size, i.e., the overparametrization of the optimal DNNs, as a function of the number of training samples. The overparametrization decreases for large training sets. (c) The test set infidelity of suboptimal models with the reduced number of trainable parameters. The left-most model consisting of approximately 3500 parameters reaches the infidelity level comparable with the optimum RBF interpolation using 16 000 parameters.

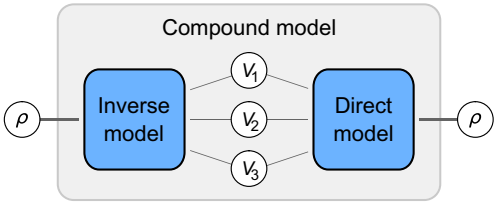


FIG. 3. Representation of the compound model created by connecting the inverse model to the already-trained direct model. The model's inverse part transforms parameters of a polarization density matrix to three control voltages. The direct part then converts these voltages back into a polarization state. The direct part is fixed during the learning process, and only the inverse part is trainable.

model). The resulting RBF compound model is trained by optimizing weights of radial basis functions in the inverse model using the training dataset. The training is performed for various radial functions to find the best model using the validation dataset. Compared to infidelities of linear interpolation, $(1-1+1)\times 10^{-2}$, and RBF interpolation, $(5-5+30)\times 10^{-2}$, the deep learning model represents a significantly more accurate method to control LC devices and predict their operation. As shown on the right side of Table I, the DNN model is also the fastest in its predictions.

V. SINGLE-PHOTON POLARIZATION STATE PREPARATION

The negligible infidelities obtained by application of the DNN models to the test dataset show unprecedentedly accurate continuous experimental control of laser beam polarization. To extend the applicability of our method even further, we implement the developed model in a single-photon polarization state preparation. Using the inverse part of our model, we predicted the control voltages for more than 1000 polarization states, with specific positions on the Bloch sphere, forming an image of the Palacký University logo; see Fig. 5(a). We prepared and fully characterized these polarization states carried by single photons.

The experimental setup is depicted in Fig. 4 in configuration (a). A 405-nm continuous-wave horizontally polarized laser beam propagates through a dichroic mirror (DM) and a dual-wavelength polarizing beam splitter (PBS) into a periodically poled potassium titanyl phosphate (PPKTP) crystal. In this crystal, two correlated 810-nm photons (signal and idler) are generated in the process of collinear type-II spontaneous parametric down-conversion (SPDC). The signal photon (reflected by DM) is spatially and polarization filtered by a single-mode fiber (SMF) and a horizontal polarizer (P), respectively. Next, the LC device applies the target polarization transformation. A resulting polarization state of the signal photon is then analyzed using the full quantum-state tomography [41,42]. The state is

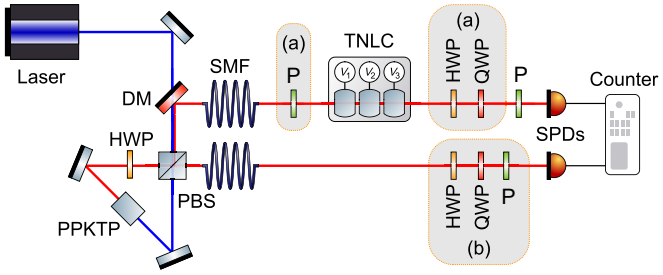


FIG. 4. (a) Scheme of heralded single-photon polarization state preparation using the idler photon (lower red path) for heralding. The components denoted as (b) are not present in this configuration. The PPKTP crystal inside a Sagnac interferometer is pumped unidirectionally (blue line). The heralded signal photon (upper red line) in the horizontal polarization state (after polarizer P) is transformed by the LC device and analyzed by quantum tomography using wave plates (HWP and QWP), a polarizer (P), and a single-photon detector (SPD). The coincidence detection events of the signal and idler detectors are acquired by a counter. (b) Setup for remote quantum state preparation using an entangled pair of photons. The components denoted as (a) are not present in this configuration. The generation of entangled photon pairs requires bidirectional pumping of the PPKTP crystal. The signal photon is projected to a target polarization state by the LC device and the polarizer. The polarization state of the idler photon is analyzed by quantum tomography; see the text for details.

sequentially projected into six eigenstates of Pauli operators, i.e., H, V, D, A, R, and L polarizations, using half- and quarter-wave plates (HWP and QWP) and a polarizer (P), and detected by a single-photon detector (SPD). The idler photon is navigated directly to the second SPD. We measured the six projections of the signal photon in coincidence with the idler photon with a 4.5 ns coincidence window and 6 s acquisition time for each projection. The coincidence detection rate is approximately 30×10^3 per second. The polarization states are reconstructed using the maximum likelihood method; see Fig. 5(b). The resulting logo formed by the prepared polarization states agrees very well with that consisting of the modeled states. The average fidelity between these sets of states reaches $F = 0.998 - 0.006 + 0.002$.

Furthermore, we demonstrate a remote state preparation [4,5,49]. The protocol requires a bipartite entangled state; its first qubit (signal photon) is projected to a target state (and detected) while the second qubit (idler photon) collapses to the same quantum state or a state with a local unitary operation applied. Specifically, the density matrix ρ_2 of the remotely prepared polarization state reads $\rho_2 = \mathrm{Tr}_1[\rho_{12}(\Pi_1 \otimes I_2)]/\mathrm{Tr}[\rho_{12}(\Pi_1 \otimes I_2)]$, where ρ_{12} is the density matrix of the entangled state, Π_1 is a projector representing the LC device followed by a polarizer and the detector acting on the first subsystem, I_2 is the identity matrix on the second subsystem, and Tr_1 is a partial trace over the first subsystem. We use the entangled state close to the ideal singlet state $\rho_{12} = |\Psi^-\rangle\langle\Psi^-|$, $|\Psi^-\rangle = (|HV\rangle - |VH\rangle)/\sqrt{2}$, which requires an additional application of the $i\sigma_y$ operation on the second subsystem to project the second photon onto the target state. The corresponding experimental setup is shown in Fig. 4 as configuration (b). The polarization of the 405-nm laser is set to a diagonal state, and the PPKTP crystal in the Sagnac interferometer is pumped from both sides [50,51]. The dual-wavelength half-wave plate in the Sagnac loop swaps horizontal and vertical polarizations. The resulting two SPDC processes interfere and produce the singlet state, which is coupled to the rest of the setup by single-mode optical fibers. Unfortunately, the signal fiber is 7 m long due to space restriction in the lab, which imposes a unitary transformation randomly evolving within the measurement time. This drift would obscure the precision of the remote preparation, and an active fiber stabilization would complicate the

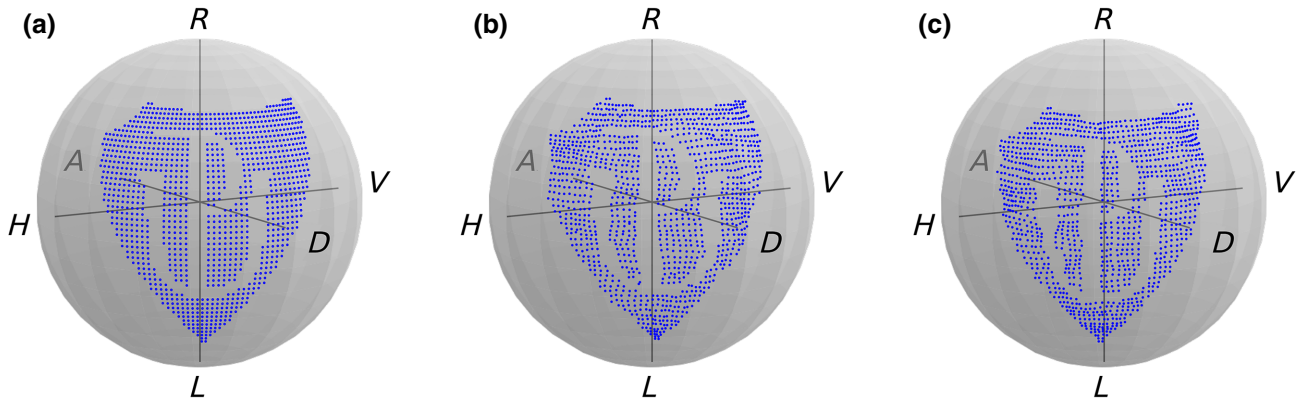


FIG. 5. Visualization of the polarization states forming the Palacký University logo on the Bloch sphere. (a) The target states for which we predict the optimal control voltages using the inverse part of the compound model. (b) The logo consisting of measured polarization states from heralded single-photon preparation, applying the predicted control voltages. (c) Polarization states obtained in the process of remote single-photon state preparation.

setup considerably. Instead, we demonstrated the remote preparation numerically, including all relevant experimental imperfections. Namely, we performed the full two-qubit state reconstruction and estimated the experimental density matrix ρ_{12}. Its purity is $P = 0.978(1)$, concurrence $C = 0.978(1)$, and fidelity with the ideal singlet state $F = 0.987(1)$. The measurement operators Π_1 are obtained as output predictions of the compound model for the input polarization states forming the Palacký University logo. The resulting states of the remote single-photon preparation are visualized in Fig. 5(c). The average fidelity of these states reaches $F = 0.988 - 0.001 + 0.002$. The remote-preparation fidelity is 0.01 smaller than the local-preparation fidelity, which is caused primarily by the imperfect purity of the entangled state.

VI. CONCLUSION

We have reported on modeling the transfer function of the complex multicell twisted nematic liquid crystal device using deep neural networks. The model is trained using an experimentally acquired dataset containing control voltages and tomographically measured polarization states. The trained model predicts the output polarization from the control voltages at the unprecedented fidelity level. The accuracy of the model is compared to commonly used approaches, namely linear interpolation and radial basis function interpolation. The deep learning model is more accurate and faster than both the reference methods by orders of magnitude. Also, the deep learning model is resource efficient; it requires significantly fewer samples than other tested approaches for the given accuracy.

Our main result lies in solving the ambiguity of the control voltages in the inverse transformation. Here the optimum control voltages are predicted for a given polarization. Various combinations of control voltages can result in almost similar polarization states. Furthermore, there is no preferred metric in the space of classical control signals (control voltages in our case). We solved these issues by creating the compound model consisting of a trainable inverse part and a fixed direct part trained in the previous step. We further verified our results by employing the deep learning models in local single-photon polarization state preparation and remote quantum state preparation. Our results open the path to ultraprecise polarimetry using liquid crystals with classical light as well as with single-photon signals in quantum information processing.

Even though our work focuses on polarization encoding, we expect similar behavior and scaling in systems transforming different degrees of freedom of light such as spatial modes or which-way information in interferometric networks. The developed approach allows for near-perfect bidirectional classical control of the polarization-encoded quantum system and is easily transferable to other photonics quantum systems.

The code and data that support the findings of this study are publicly available on GitHub [52].

ACKNOWLEDGMENTS

We acknowledge funding from the Czech Science Foundation, under Grant No. 21-18545S; the European Union's Horizon 2020 (2014–2020) research and innovation framework programme, QuantERA grant HYPER-U-P-S; The Ministry of Education, Youth and Sports of the Czech Republic under Grant No. 8C18002; and Palacký University under Grants No. IGA-PrF-2021-006 and No. IGA-PrF-2022-005. We acknowledge use of cluster computing resources provided by the Department of Optics, Palacký University Olomouc. We thank J. Provazník for maintaining the cluster and providing support. We also acknowledge the participation of I. Straka in developing the entanglement source.

[1] Z. Zhuang, S.-W. Suh, and J. S. Patel, Polarization controller using nematic liquid crystals, Opt. Lett. **24**, 694 (1999).

[2] R. B. A. Adamson, L. K. Shalm, and A. M. Steinberg, Preparation of pure and mixed polarization qubits and the direct measurement of figures of merit, Phys. Rev. A **75**, 012104 (2007).

[3] R. B. A. Adamson and A. M. Steinberg, Improving Quantum State Estimation with Mutually Unbiased Bases, Phys. Rev. Lett. **105**, 030406 (2010).

[4] N. A. Peters, J. T. Barreiro, M. E. Goggin, T.-C. Wei, and P. G. Kwiat, Remote State Preparation: Arbitrary Remote Control of Photon Polarization, Phys. Rev. Lett. **94**, 150502 (2005).

[5] H. Defienne, B. Ndagano, A. Lyons, and D. Faccio, Polarization entanglement-enabled quantum holography, Nat. Phys. **17**, 591 (2021).

[6] A. Lohrmann, C. Perumgatt, and A. Ling, Manipulation and measurement of quantum states with liquid crystal devices, Opt. Express **27**, 13765 (2019).

[7] A. Villar, A. Lohrmann, X. Bai, T. Vergoossen, R. Bedington, C. Perumangatt, H. Y. Lim, T. Islam, A. Reezwana, Z. Tang, R. Chandrasekara, S. Sachidananda, K. Durak, C. F. Wildfeuer, D. Griffin, D. K. L. Oi, and A. Ling, Entanglement demonstration on board a nano-satellite, Optica **7**, 734 (2020).

[8] K. A. G. Fisher, R. Prevedel, R. Kaltenbaek, and K. J. Resch, Optimal linear optical implementation of a single-qubit damping channel, New J. Phys. **14**, 033016 (2012).

[9] A. Orieux, L. Sansoni, M. Persechino, P. Mataloni, M. Rossi, and C. Macchiavello, Experimental Detection of Quantum Channels, Phys. Rev. Lett. **111**, 220501 (2013).

[10] D. S. Naik, C. G. Peterson, A. G. White, A. J. Berglund, and P. G. Kwiat, Entangled State Quantum Cryptography: Eavesdropping on the Ekert Protocol, Phys. Rev. Lett. **84**, 4733 (2000).

[11] J. T. Barreiro, T.-C. Wei, and P. G. Kwiat, Beating the channel capacity limit for linear photonic superdense coding, Nat. Phys. **4**, 282 (2008).
[12] H. Abbaszadeh, M. Fruchart, W. van Saarloos, and V. Vitelli, Liquid-crystal-based topological photonics, Proc. Natl. Acad. Sci. **118**, e2020525118 (2021).
[13] A. Yariv and A. P. Yeh, *Optical Waves in Crystals: Propagation and Control of Laser Radiation* (Wiley, New York, 1984).
[14] J. A. Coy, M. Zaldarriaga, D. F. Grosz, and O. E. Martinez, Characterization of a liquid crystal television as a programmable spatial light modulator, Opt. Eng. **35**, 15 (1996).
[15] A. Marquez, J. Campos, M. J. Yzuel, I. S. Moreno, J. A. Davis, C. C. Iemmi, A. Moreno, and A. Robert, Characterization of edge effects in twisted nematic liquid crystal displays, Opt. Eng. **39**, 3301 (2000).
[16] A. Marquez, C. C. Iemmi, I. S. Moreno, J. A. Davis, J. Campos, and M. J. Yzuel, Quantitative prediction of the modulation behavior of twisted nematic liquid crystal displays based on a simple physical model, Opt. Eng. **40**, 2558 (2001).
[17] M. Yamauchi, Jones-matrix models for twisted-nematic liquid-crystal devices, Appl. Opt. **44**, 4484 (2005).
[18] M. Bielak, R. Stárek, V. Krčmarský, M. Mičuda, and M. Ježek, Accurate polarization preparation and measurement using liquid crystal displays, Opt. Express **29**, 33037 (2021).
[19] C. Monroe and J. Kim, Scaling the ion trap quantum processor, Science **339**, 1164 (2013).
[20] M. H. Devoret and R. J. Schoelkopf, Superconducting circuits for quantum information: An outlook, Science **339**, 1169 (2013).
[21] J. Wang, F. Sciarrino, A. Laing, and M. G. Thompson, Integrated photonic quantum technologies, Nat. Photonics **14**, 273 (2019).
[22] T. Fösel, P. Tighineanu, T. Weiss, and F. Marquardt, Reinforcement learning with neural networks for quantum feedback, Phys. Rev. X **8**, 031084 (2018).
[23] V. Cimini, I. Gianani, N. Spagnolo, F. Leccese, F. Sciarrino, and M. Barbieri, Calibration of Quantum Sensors by Neural Networks, Phys. Rev. Lett. **123**, 230502 (2019).
[24] V. Cimini, E. Polino, M. Valeri, I. Gianani, N. Spagnolo, G. Corrielli, A. Crespi, R. Osellame, M. Barbieri, and F. Sciarrino, Calibration of Multiparameter Sensors via Machine Learning at the Single-Photon Level, Phys. Rev. Appl. **15**, 044003 (2021).
[25] J. Wallnöfer, A. A. Melnikov, W. Dür, and H. J. Briegel, Machine Learning for Long-Distance Quantum Communication, PRX Quantum **1**, 010301 (2020).
[26] P. Sgroi, G. M. Palma, and M. Paternostro, Reinforcement Learning Approach to Nonequilibrium Quantum Thermodynamics, Phys. Rev. Lett. **126**, 020601 (2021).
[27] J. Brown, P. Sgroi, L. Giannelli, G. S. Paraoanu, E. Paladino, G. Falci, M. Paternostro, and A. Ferraro, Reinforcement learning-enhanced protocols for coherent population-transfer in three-level quantum systems, New J. Phys. **23**, 093035 (2021).
[28] A. Suprano, D. Zia, E. Polino, T. Giordani, L. Innocenti, A. Ferraro, M. Paternostro, N. Spagnolo, and F. Sciarrino, Dynamical learning of a photonics quantum-state engineering process, Adv. Photonics **3**, 066002 (2021).
[29] F. Flamini, L. Magrini, A. S. Rab, N. Spagnolo, V. D'Ambrosio, P. Mataloni, F. Sciarrino, T. Zandrini, A. Crespi, R. Ramponi, and R. Osellame, Thermally reconfigurable quantum photonic circuits at telecom wavelength by femtosecond laser micromachining, Light Sci. Appl. **4**, e354 (2015).
[30] J. Carolan, C. Harrold, C. Sparrow, E. Martin-Lopez, N. J. Russell, J. W. Silverstone, P. J. Shadbolt, N. Matsuda, M. Oguma, M. Itoh, G. D. Marshall, M. G. Thompson, J. C. F. Matthews, T. Hashimoto, J. L. O'Brien, and A. Laing, Universal linear optics, Science **349**, 711 (2015).
[31] M. Jacques, A. Samani, E. El-Fiky, D. Patel, Z. Xing, and D. V. Plant, Optimization of thermo-optic phase-shifter design and mitigation of thermal crosstalk on the SOI platform, Opt. Express **27**, 10456 (2019).
[32] M. Werninghaus, D. J. Egger, F. Roy, S. Machnes, F. K. Wilhelm, and S. Filipp, Leakage reduction in fast superconducting qubit gates via optimal control, npj Quantum Inf. **7**, 14 (2021).
[33] N. Wittler, F. Roy, K. Pack, M. Werninghaus, A. S. Roy, D. J. Egger, S. Filipp, F. K. Wilhelm, and S. Machnes, Integrated Tool Set for Control, Calibration, and Characterization of Quantum Devices Applied to Superconducting Qubits, Phys. Rev. Appl. **15**, 034080 (2021).
[34] R. Trotta, E. Zallo, C. Ortix, P. Atkinson, J. D. Plumhof, J. van den Brink, A. Rastelli, and O. G. Schmidt, Universal Recovery of the Energy-Level Degeneracy of Bright Excitons in InGaAs Quantum Dots without a Structure Symmetry, Phys. Rev. Lett. **109**, 147401 (2012).
[35] J. D. Plumhof, R. Trotta, A. Rastelli, and O. G. Schmidt, Experimental methods of post-growth tuning of the excitonic fine structure splitting in semiconductor quantum dots, Nanoscale Res. Lett. **7**, 1 (2012).
[36] T. A. Baart, P. T. Eendebak, C. Reichl, W. Wegscheider, and L. M. K. Vandersypen, Computer-automated tuning of semiconductor double quantum dots into the single-electron regime, Appl. Phys. Lett. **108**, 213104 (2016).
[37] J. D. Teske, S. S. Humpohl, R. Otten, P. Bethke, P. Cerfontaine, J. Dedden, A. Ludwig, A. D. Wieck, and H. Bluhm, A machine learning approach for automated fine-tuning of semiconductor spin qubits, Appl. Phys. Lett. **114**, 133102 (2019).
[38] J. P. Zwolak, T. McJunkin, S. S. Kalantre, J. P. Dodson, E. R. MacQuarrie, D. E. Savage, M. G. Lagally, S. N. Coppersmith, M. A. Eriksson, and J. M. Taylor, Autotuning of Double-Dot Devices In Situ with Machine Learning, Phys. Rev. Appl. **13**, 034075 (2020).
[39] H. Moon, D. T. Lennon, J. Kirkpatrick, N. M. van Esbroeck, L. C. Camenzind, L. Yu, F. Vigneau, D. M. Zumbühl, G. A. D. Briggs, M. A. Osborne, D. Sejdinovic, E. A. Laird, and N. Ares, Machine learning enables completely automatic tuning of a quantum device faster than human experts, Nat. Commun. **11**, 4161 (2020).
[40] R. Durrer, B. Kratochwil, J. V. Koski, A. J. Landig, C. Reichl, W. Wegscheider, T. Ihn, and E. Greplova, Automated Tuning of Double Quantum Dots into Specific

Charge States Using Neural Networks, Phys. Rev. Appl. **13**, 054019 (2020).
[41] M. Ježek, J. Fiurášek, and Z. Hradil, Quantum inference of states and processes, Phys. Rev. A **68**, 012305 (2003).
[42] Z. Hradil, J. Řeháček, J. Fiurášek, and M. Ježek, in *Quantum State Estimation* (Springer, 2004), p. 59.
[43] S. L. Digabel, Algorithm 909, ACM Trans. Math. Softw. **37**, 1 (2011).
[44] C. Audet, S. L. Digabel, V. R. Montplaisir, and C. Tribes, NOMAD version 4: nonlinear optimization with the MADS algorithm, ArXiv:2104.11627 (2021).
[45] D. P. Kingma and J. Ba, in *ICLR Conference Track Proceedings*, ArXiv:1412.6980 (2015).
[46] M. Belkin, D. Hsu, S. Ma, and S. Mandal, Reconciling modern machine-learning practice and the classical bias-variance trade-off, Proc. Natl. Acad. Sci. USA **116**, 15849 (2019).
[47] S. Bubeck and M. Sellke, A universal law of robustness via isoperimetry, ArXiv:2105.12806 (2021).
[48] D. Broomhead and D. Lowe, Multivariable functional interpolation and adaptive networks, Complex Syst. **2**, 321 (1988).
[49] C. H. Bennett, D. P. DiVincenzo, P. W. Shor, J. A. Smolin, B. M. Terhal, and W. K. Wootters, Remote State Preparation, Phys. Rev. Lett. **87**, 077902 (2001).
[50] A. Fedrizzi, T. Herbst, A. Poppe, T. Jennewein, and A. Zeilinger, A wavelength-tunable fiber-coupled source of narrowband entangled photons, Opt. Express **15**, 15377 (2007).
[51] A. Predojević, S. Grabher, and G. Weihs, Pulsed Sagnac source of polarization entangled photon pairs, Opt. Express **20**, 25022 (2012).
[52] D. Vašinka, Github repository, https://github.com/VasinkaD/NeuralLC (2021).

LASER & PHOTONICS REVIEWS
www.lpr-journal.org

All-Fiber Microsensor of Polarization at Single-Photon Level Aided by Deep Learning

*Martin Bielak, Dominik Vašinka, and Miroslav Ježek**

Polarization of light carries vital information in numerous scientific disciplines, including biomedical imaging, optical diagnostics, and environmental sensing. However, accurate polarization measurements in constrained spaces, under low-light conditions, and at high speeds remain a severe challenge. A compact, all-fiber polarization sensor capable of single-shot, real-time operation with single-photon sensitivity and long-term stability is presented. The sensor leverages intermodal interference within a short segment of a few-mode fiber, coupled to an array of single-photon detectors. The recorded detections are processed by a neural network, enabling precise reconstruction of complete polarization information for fully and partially polarized states. This robust architecture allows for thousands of polarization state measurements per second while achieving exceptional accuracy with Stokes error below 0.01, and even lower at higher photon fluxes. This technology is demonstrated through diverse experimental scenarios, such as resolving structural details in biological tissues with a spatial resolution of 6 µm, characterizing rapid polarization transitions, and monitoring micro-scale birefringence in living or moving specimens.

1. Introduction

From imaging[1,2] and sensing[3] to ellipsometry[4,5] and optical communications,[6,7] light polarization plays a fundamental role across numerous scientific disciplines. In particular, the polarization of light emitted by an optically anisotropic specimen carries essential information about its material structure and optical properties. This capability has led to widespread applications in medicine, where polarization provides insight into the structural and morphological characteristics of various biological tissues and enables label-free detection of pathological changes.[8] Polarization-resolved endoscopy and imaging techniques have proven valuable for detecting, classifying, and staging cancerous lesions in organs such as the stomach, colon, prostate, and larynx.[9,10] However, while polarization measurements can be relatively straightforward in a controlled laboratory environment, achieving highly accurate and reliable characterization under challenging conditions and in complex dynamic scenarios remains a significant obstacle.

Low light intensities, representing the first challenging condition under consideration, come with a reduced signal-to-noise ratio, which complicates extracting meaningful polarization information from data burdened with environmental and shot noise. Similarly, conducting high-speed measurements, as the second condition, is accompanied by low signal levels within the acquisition time. While increasing the incident light intensity can help, it is limited by our ability to deliver the optical power to the particular sensing area and, most ultimately, by the photodamage threshold of the sample. Incorporating single-photon detectors can substantially improve the accuracy of protocols relying on the number of detected photons for polarization state reconstruction. However, another drawback arises as these detectors are relatively large and cannot be efficiently integrated. This issue becomes particularly problematic when faced with the third condition - constrained space. Such spaces not only impose restrictions on the overall size of the measurement setup but also limit the placement and alignment of optical components. Therefore, these challenging conditions pose a severe setback, especially for measurements during invasive medical procedures, such as microendoscopy,[11,12] environmental monitoring,[13,14] or internal/in-situ inspections of materials.[15,16] We address these problems by developing a polarization sensor based on light propagation in disordered media.

Coherent light propagating within a disordered medium is subjected to multimode interference, generating a granular speckle pattern. Despite the apparent degrading effect, the pattern retains encoded information about the incident light.[17–21] Analyzing the light propagation in disordered media and employing suitable post-processing techniques, one can study the disordering effects of the media,[22,23] estimate and classify input images,[12,24–28] learn to modulate the input for projecting a targeted image on the output,[29] reconstruct shapes of propagated ultrashort laser pulses,[30,31] compute and communicate,[32–36]

M. Bielak, D. Vašinka, M. Ježek
Department of Optics
Faculty of Science
Palacký University
17. listopadu 1192/12, 77900 Olomouc, Czechia
E-mail: jezek@optics.upol.cz

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/lpor.202501775

DOI: 10.1002/lpor.202501775

and even make an accurate spectrometer or wavemeter.[37,38] Recently, the speckle pattern has proved to preserve a great deal of information about the polarization state of the incident light.[39,40] Although additional studies successfully advancing the application of disordered-media propagation have been conducted,[41–44] developing a compact polarimeter capable of rapid measurements with single-photon sensitivity continues to be an open problem.

Here, we introduce a highly accurate, single-shot, all-fiber polarization sensor based on intermodal interference in a very short segment of a few-mode fiber. Instead of processing the complete spatial speckle pattern, our approach relies on its sparse sampling in a few isolated points using a fiber array. Combined with fiber-coupled single-photon detectors, this design allows for high sensitivity down to a single-photon level. The all-fiber sensor contains no moving components or complex metastructures and is sufficiently compact for procedures in constrained spaces while operating at a spatial resolution of 6 μm. Additionally, it supports a high-throughput operation capable of resolving several thousand polarization states per second. The setup is aided by a deep machine-learning model reconstructing complete polarization information in real-time. Once calibrated, the sensor maintains unparalleled accuracy, reaching Stokes errors well below 0.01, and long-term stability for over a month, even when measuring partially polarized light. We demonstrate its polarimetric capabilities through various experiments, including scans of dense connective tissue, birefringence measurements with a USAF test target, and dynamic polarization sensing in a moving diatom and during a fast twisted nematic liquid crystal transition. This all-fiber sensor sets new standards for compact, stable, and high-speed polarization measurements, offering a powerful tool for cutting-edge applications in biomedicine, material research, and beyond.

2. Results

2.1. Experimental and Data Processing Setup

The speckle-based polarization sensor illustrated in **Figure 1** employs a common few-mode fiber to characterize the input polarization state. The sensor collects light with a front face of a 5 cm long step-index SMF28 fiber with a numerical aperture of 0.14 and a core diameter of 8.2 μm. While the motivation behind choosing this specific fiber type is to minimize the sensing area and keep the speckle structures large, we emphasize that this preference is optional. Virtually any few-mode or multimode fiber, including significantly shorter variants (see Section 2.2), can be used as a substitute. The main requirement for the fiber is to support several interfering modes necessary for creating a polarization-dependent speckle pattern at the output. Following a 5-mm free-space propagation, we collect discrete samples of this interference pattern using a fiber array comprising seven gradient multimode fibers, each with a 62.5 μm core diameter, forming a 375 μm diameter honeycomb pattern. Each fiber then propagates the collected intensity signal to an independent single-photon avalanche diode, recording the number of detection events. After normalizing the measured counts, this setup associates the sampled polarization state with the corresponding

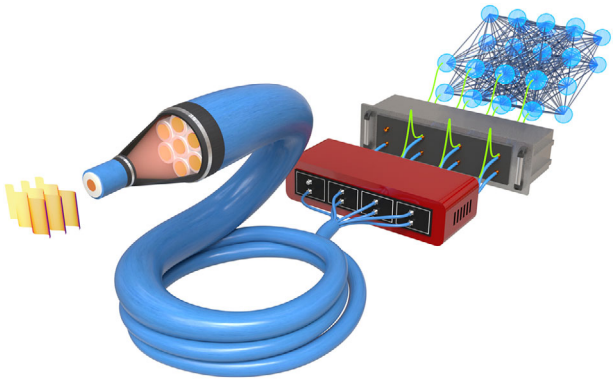


Figure 1. A visual representation of the all-fiber polarization sensor. Polarized light emitted from a specimen is collected by a short piece of a few-mode fiber. The intermodal interference generates a granular speckle pattern at the output. Several isolated samples of this pattern impinge a fiber array and propagate to a corresponding number of single-photon detectors. The recorded detections are electronically processed using deep learning methods to characterize the incident polarization state.

relative frequencies of each detector, which we term a count distribution. For more experimental details, see Section 4.1 and **Figure 7**.

Using the sensor, we analyzed 30 000 polarization states, uniformly covering the entire surface of the Poincaré sphere.[45] Our state preparation involved a continuous laser beam with a central wavelength of 0.8 μm attenuated to the single-photon level and propagated through a linear polarizer. Subsequently, a twisted nematic liquid crystal device,[46] controlled by voltage signals, performed a targeted unitary transformation of the polarization state on this weak coherent light. The device operation underwent calibration through bidirectional modeling,[47] ensuring fast and precise preparation of arbitrary polarization states. Its accuracy was verified using a reference polarimeter based on wave plates. The polarization state preparation using both methods reaches an average fidelity exceeding 0.999, however, the liquid crystal device operates by orders of magnitude faster. Subsequently, our sensor characterized the prepared polarization state, providing the corresponding detection counts. On average, we acquired approximately 150 000 photon detection events per detector within a 50 ms acquisition time window. Moreover, we expand our dataset to include partially polarized light. We establish a count distribution for a mixture of two orthogonal states by weight-summing their respective distributions. This approach is virtually equivalent to directly measuring the corresponding mixed state. Employing this method, we assembled an extensive dataset of polarization states distributed uniformly throughout the entire volume of the Poincaré sphere.

We used the dataset to train a fully connected deep neural network for reconstructing the polarization state given the associated count distribution, see Section 4.3 for details. The optimized and fully trained network was evaluated on a test set of previously unobserved polarization states, also acquired from the experimental setup. We achieved an unprecedented average infidelity of 8×10^{-4} with a $[5 \times 10^{-5}, 2 \times 10^{-3}]$ confidence interval using the 10th and 90th percentile. These results closely approach the fidelity limit imposed by the polarization preparation itself,[46]

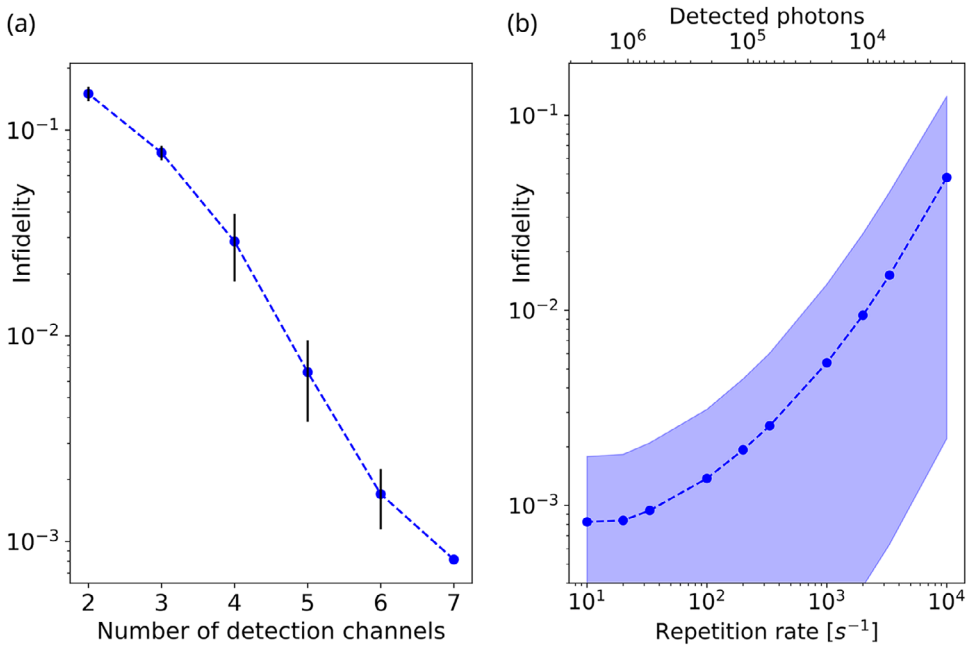


Figure 2. a) Polarization error quantified by infidelity for varying numbers of active detectors. Variability bars represent uncertainties arising from different subset combinations. The achieved infidelities underscore the sufficiency of sparse sampling of the speckle pattern. b) The relation between polarization infidelity, the collective number of detected photons across all detectors, and the measurement repetition rate. The colored area indicates the confidence interval of the experimental test set.

indicating that our method reaches the fundamental accuracy bounds of the experiment. The achieved infidelity translates to a reconstruction error of approximately 0.01 per Stokes parameter. In contrast, previous methods based on random scattering typically reach errors around 0.05 while requiring bulky setups unsuitable for in-situ applications, achieving lower operational speed, lacking the single-photon sensitivity, and not presenting the ability to process partially polarized light.[39–42] To further validate our approach, we performed an additional benchmark using a 1951 USAF birefringent test. As detailed in Section 4.4, we compared our sensor with a polarization measurement based on rotating wave plates. The two methods yielded an average fidelity of ≈ 0.987, underscoring the accuracy of our all-fiber sensor. Additionally, the USAF test also allowed the sensor spatial resolution estimation, which was determined to be $\approx 6(5)$ µm. Together, these results comprehensively demonstrate the exceptional accuracy and robustness of our approach.

2.2. Performance Evaluation

We highlight that our approach eliminates the need for capturing an image of the complete speckle pattern. Instead, we collect a limited number of speckle spots, which are then propagated through the fiber array channels. This substantial reduction in data collection allows for sensitive measurements at extremely low intensities using single-photon detectors. To assess the sensitivity of our approach to the processed portion of the speckle pattern, we investigate the achievable infidelity with regard to the number of employed detection channels. Following the same procedure, we trained additional networks utilizing only a subset of the detection channels as inputs. The infidelities achieved by these networks on corresponding test sets are depicted in **Figure** 2a. The values represent the averaged errors across multiple trained networks, each employing a distinct combination of detection channel subsets. The results illustrated in this graph unequivocally demonstrate that extracting only a few samples from the speckle pattern provides sufficient information about the polarization state.

We also explored the dependence of the polarization sensing accuracy on the measurement repetition rate, i.e., the inverse value of the acquisition time, representing the number of polarization states characterized per second. Figure 2b displays the polarization infidelities of a single network, using all seven channels, evaluated for numerous repetition rates. This analysis allows us to determine the minimum required light intensity (i.e., the number of detected photons) needed to achieve a targeted infidelity value. Similarly, it allows estimation of the attainable fidelity for a given repetition rate. For example, the system achieves an average fidelity of 0.999 at a rate of 33 Hz—approximately the threshold at which the human eye perceives continuous motion. This result highlights the capability of our sensor for accurate, real-time polarization measurements. Furthermore, our setup demonstrates the capacity to measure over 2000 polarization states per second using only 10^4 photons while maintaining an average fidelity above 0.99. It is crucial to emphasize that these performances depend on the number of detected photons rather than the repetition rate alone. Consequently, even significantly higher repetition rates can attain the same fidelities with appropriately increased light intensity, limited virtually only by saturation of employed single-photon detectors. The presented photon efficiency makes the

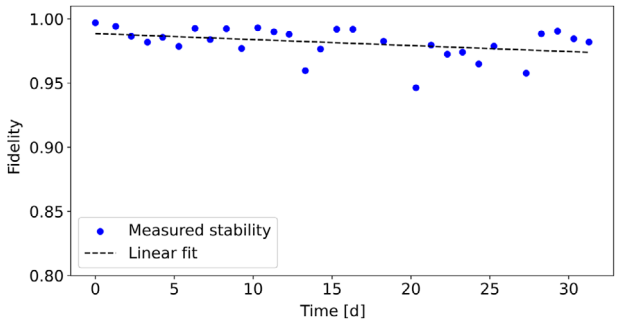


Figure 3. Long-term stability evaluation of the all-fiber polarization sensor employing a 12 mm long few-mode fiber encased in a ceramic ferrule. The system performance was evaluated periodically with a 24-hour repetition interval. The average fidelities (blue dots) are accompanied by a two-parameter linear fit (black line) with a 4.7×10^{-4} d^{-1} negative slope, capturing the slow, gradual decline.

sensor particularly suitable for high-throughput, rapid-sensing applications.

Furthermore, let us discuss the long-term stability of our approach and its robustness against environmental effects. Even when using a 5 cm long, few-mode fiber in an open environment without elaborated mechanical stabilization, the polarization sensor maintained sufficient stability for a week without recalibration or network retraining. We explored the possibility of further enhancing the stability using a 12 mm long fiber encased within a ceramic ferrule. This encasing rigidly fixes the fiber position relative to the fiber array, minimizing mechanical bending and reducing susceptibility to temperature gradients. With this modification, the system maintained its performance for over a month without recalibration. **Figure 3** illustrates the results of our long-term stability test, which was evaluated periodically at 24-hour intervals. The average fidelity was observed to degrade by $\approx 4.7 \times 10^{-4}$ d^{-1}, i.e., 2×10^{-5} h^{-1}. This exceptional stability is attributed to three key factors: the short fiber length, which minimizes propagation-induced instabilities; the ferrule encasing, which mitigates environmental perturbations; and the use of a few-mode fiber, which produces a coarse-grained speckle pattern that increases the collection efficiency and reduces sensitivity to vibrations and drift. This design starkly contrasts with imaging-based applications, where long and massively multimode fibers have to be used with a deteriorating impact on the stability.[12,24–28] Since sufficient intermodal coupling to generate a speckle pattern in a multimode fiber occurs over a propagation distance of several hundred wavelengths,[48,49] we could reduce the employed fiber length even to sub-millimeter lengths. We expect that reducing the fiber to this scale could yield an additional order-of-magnitude improvement in stability. Finally, the sensor can be recalibrated to prolong its stability indefinitely. The recalibration process is straightforward, fully automatable, and can be completed in under an hour, making regular, even daily, recalibration a practical option. In summary, our polarization sensor offers a highly stable, accurate, and efficient solution for real-time, high-speed polarimetry with minimal maintenance requirements and strong resistance to environmental disturbances.

2.3. Applications

We first demonstrate the sensing ability of our approach by conducting a polarization-sensitive scan of dense connective tissue (AmScope PS25W). While the primary application of our sensor significantly differs from polarization imaging, this optically anisotropic specimen is famously known in the polarization microscopy field. Visualizing its typical structures further underscores the accuracy of the presented polarimetric sensor. Using the experimental arrangement outlined in Section 4.1, we replaced the calibration polarization preparation stage with the tissue specimen. The few-mode fiber was positioned proximal to the tissue, capturing the induced polarization transformation. We used motorized translation stages (Newport MFA-CC) with a bidirectional repeatability of ±0.15 µm and a maximum speed of 2500 µm s^{-1} to scan over the tissue area. The reconstructed polarization states were expressed as three Bloch parameters, i.e., renormalized Stokes parameters,[45] and visualized as a false-colored RGB image, see Section 4.4 for details. In **Figure 4**, we present a side-by-side comparison of this polarization scan (middle) with an intensity image (left) captured from an identical tissue region. This intensity image was obtained using a monochrome camera (ImagingSource DMK 23U274) with a pixel size of 4.4×4.4 µm. We illuminated the specimen with 0.8 µm light and projected the resulting image onto the camera through a 25.4 mm focal-length lens, providing an estimated fivefold magnification. Moreover, the third panel (right) depicts a polarization image of the same specimen obtained using a stand-alone polarization microscope. Its experimental setup comprises a pair of crossed polarizers and the monochrome camera. As illustrated in Figure 4, both polarization-based methods reveal underlying polarization structures that remain concealed in the intensity profile. This typical inner structure of the dense connective tissue verifies that our sensor performs polarization characterization, providing complete information about the polarization state.

As the dense connective tissue is a static specimen, the complete polarization analysis provided by the all-fiber sensor can also be performed using an advanced imaging polarimeter comprising rotating wave plates. However, such a technique is insufficient for characterizing a rapidly changing polarization state induced by moving or evolving specimens. These scenarios require significantly faster approaches that have to be performed in a single-shot regime. To demonstrate the dynamic capabilities of our all-fiber sensor, we conducted polarization measurements of two systems: a floating diatom and twisted nematic liquid crystals undergoing a fast voltage-controlled transition. Starting with diatoms, these microscopic unicellular organisms comprise a complex inner structure that can exhibit anisotropic properties. We characterized the birefringence of an Actinoptychus heliopelta diatom positioned within the same experimental arrangement as the previous application. The spatial variation in the Bloch parameters across the diatom is depicted in **Figure 5** with red lines. Compared to the blue lines of background polarization, the diatom properties are visible in the modulation of Bloch parameters. This analysis can count the number of living specimens, estimate their size and morphology, and classify them

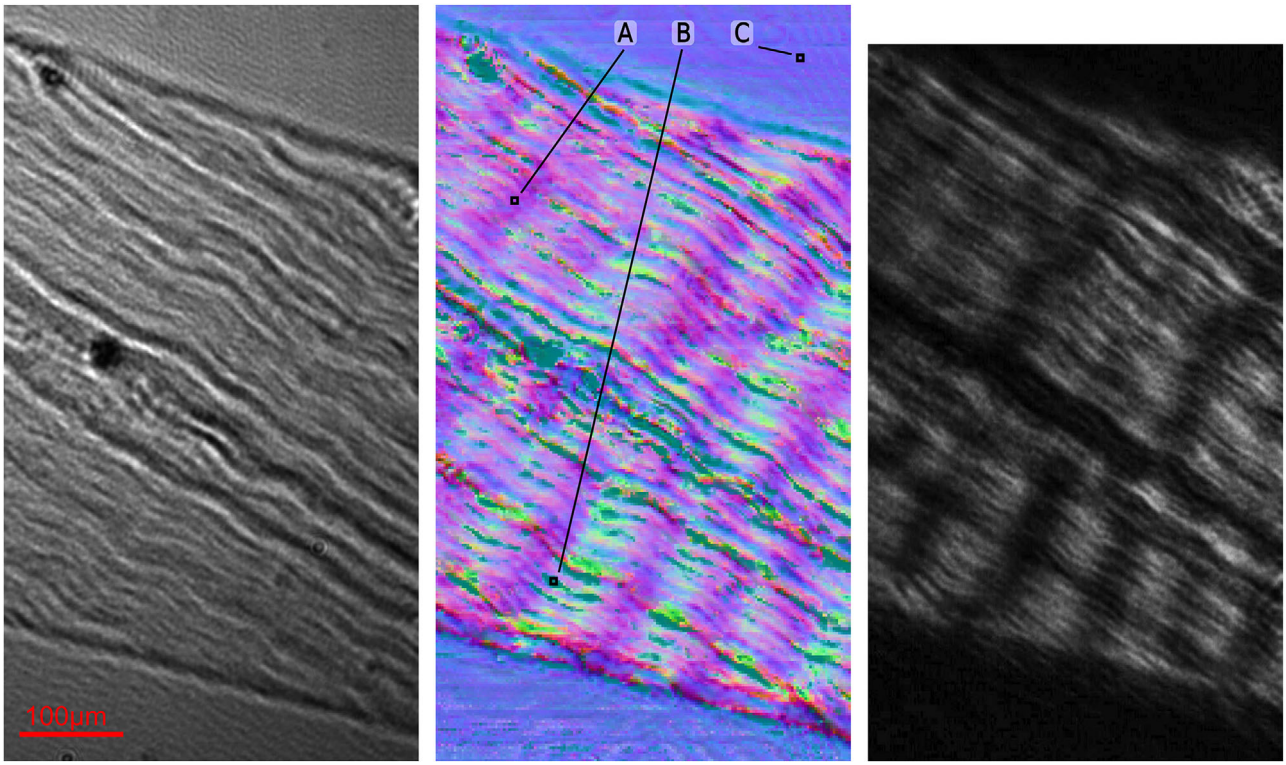


Figure 4. Visualization of dense connective tissue: (left) an intensity image, (middle) a scan using the all-fiber polarization sensor, (right) an image using a stand-alone polarization microscope. The resulting Bloch parameters of the all-fiber scan are represented as an RGB false-colored image. The three highlighted pixels characterize the purple and green segments in the polarization structure alongside a reference background polarization. Their respective Bloch parameters are A = (0.34, −0.43, 0.82), B = (−0.90, 0.05, 0.06), and C = (−0.04, −0.18, 0.98).

based on their birefringent properties. Beyond biological samples, the sensor can also find applications for fast in-situ material inspection, local strain analysis, and micro crystal growth monitoring.

Finally, we demonstrate the high-speed polarimetric capabilities of our sensor using a rapidly evolving specimen: twisted nematic liquid crystals. They allow for fast polarization state transformation, controlled by low-voltage signals that manipulate their spatial alignment. During reorientation, the induced polarization transformation undergoes a rapid transition, typically within a time window ranging from 50 to 600 ms.[46] We characterized the horizontal-to-vertical transition by continuously measuring the prepared polarization state using a 5 ms acquisition window. The reconstructed Bloch parameters corresponding to this process are visualized in **Figure 6**. The first Block parameter exhibits a smooth evolution from $B_1 = 1$ (horizontal polarization) to $B_1 = -1$ (vertical polarization). The remaining two parameters reveal that the transition corresponds to a continuous rotation on a Bloch sphere, passing close to the diagonal state ($B_2 = 1$). The complete transition occurs within approximately 50 ms, validating the ability of our sensor to accurately capture dynamic polarization changes in real time, which is inaccessible to common rotating-wave-plate polarimeters. This demonstration, alongside the previously discussed applications, underscores the versatility and performance of our all-fiber polarization sensor for fast, low-light, high-precision polarimetry.

3. Discussion

We have developed a single-shot polarization sensor based on light propagation through a short rigid piece of a few-mode fiber, resulting in polarization-dependent scattering. This process provides a coarse speckle pattern with low signal losses and unprecedented long-term stability, which is virtually unaffected by fiber bending. The defining aspect of our approach is the sparse sampling of this pattern using a fiber array instead of capturing the whole intensity profile. This characteristic allows the use of any photonic detectors, including high-speed or single-photon detectors, enabling fast response and operation at low light intensities down to the picowatt level. The implementation of deep learning algorithms in the system ensures an accurate reconstruction, providing complete information about the polarization state, including partial polarization, with an infidelity of 8×10^{-4}. This superior performance remains stable for over a month without requiring adjustments or recalibration. Complemented by the rapid operation speed of more than two thousand high-fidelity reconstructions per second, it allows for both real-time and high-repetition-rate polarization sensing with remarkable precision. The all-fiber sensor, featuring a compact design with no moving components, is particularly well-suited for reducing the invasiveness of biomedical procedures in constrained spaces and under low-light conditions. Furthermore, one can enhance the spatial resolution of our sensor by tapering the tip of the few-mode fiber to reach the limit on near-field sensing while pre-

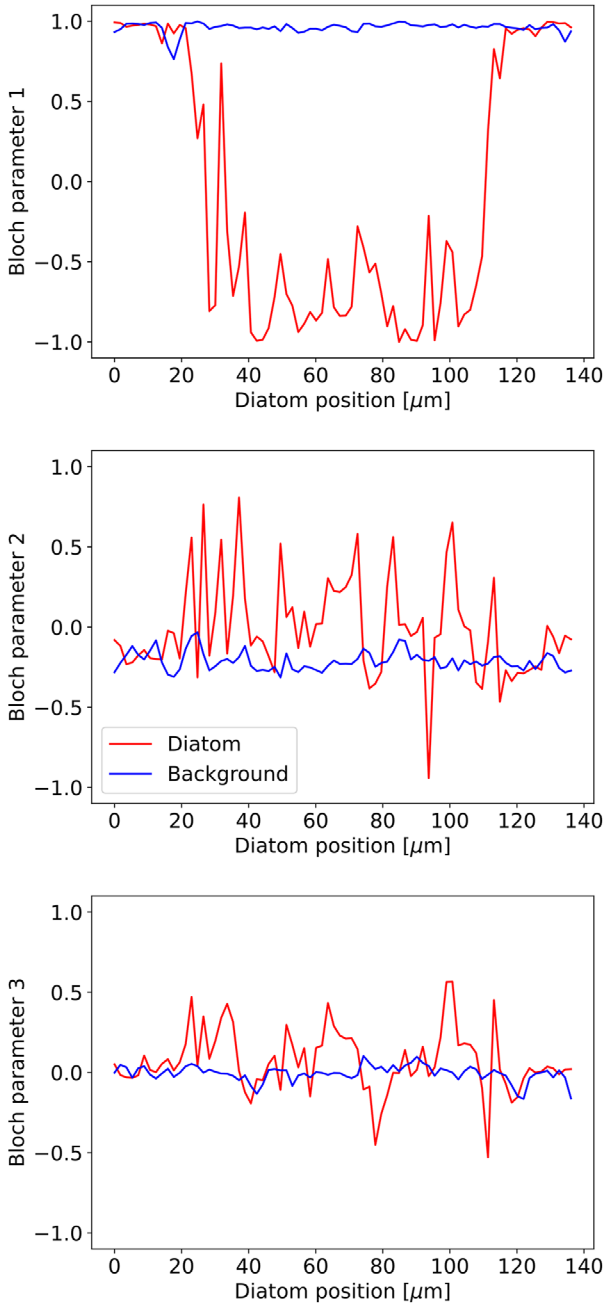


Figure 5. Spatial distribution of a diatom birefringent structure measured by the polarimetric all-fiber sensor and characterized using three Bloch parameters. The red color corresponds to changes induced by an anisotropic diatom passing in front of the fiber tip. For comparison, the blue line represents the same environment without the diatom. The modulation visible in all three Bloch parameters indicates the properties of a polarization-affecting element moving in front of the sensor, allowing further study of the specimen.

serving the single-photon polarization sensitivity. However, there will be a trade-off between the increased spatial resolution and decreased collection efficiency. Also, the stability performance might be negatively affected.

In the view of optically-assisted machine learning, the random propagation in multimode fiber followed by a computational neural network represents an instance of an extreme learning machine or reservoir computing.[35,50] By demonstrating the single-photon sensing capability of these architectures, we have provided the foundation for decreasing energy consumption in optical sensing, similar to what has been studied recently for computing.[51] Furthermore, to fully quantify the flow of information about the measured quantity (polarization state of light in our case), one may contemplate utilizing the Fisher information concept that has been only recently adapted for wave scattering.[52] The crucially missing part, and possible future focus of the research community, might be the analysis of the flow of Fisher information in neural networks.[53,54] This would complement the analysis of the physical part of the sensor and allow for the full theoretical information analysis of the deep-learning aided metrology based on random wave scattering.

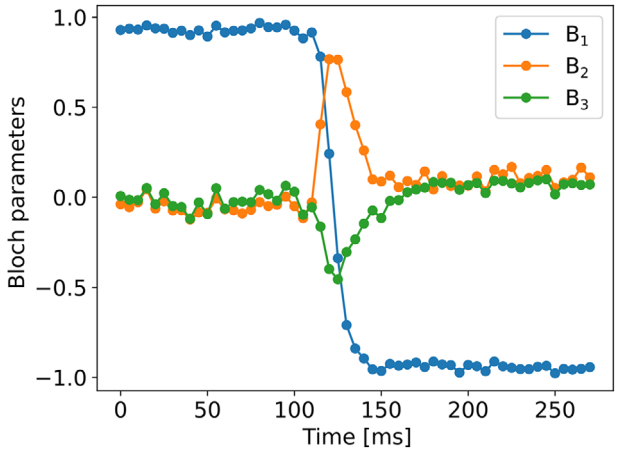


Figure 6. Time-resolved measurement of a polarization transition induced by voltage-controlled twisted nematic liquid crystals. The all-fiber sensor continuously captures the complete polarization information with a 5 ms acquisition window at the approximate power level of 10^5 detected photons. The transition from horizontal to vertical polarization is represented through the evolution of the three Bloch parameters, each shown in a different color.

In summary, we presented a novel polarization sensor that is explicitly demonstrated for arbitrary polarization states, including partially polarized light, and simultaneously offers: single-photon sensitivity (picowatt powers transmitting a sample); rapid operation with thousands of reconstructions per second; high accuracy reaching Stokes errors below 0.01, and even lower at higher photon fluxes; long-term stability for over a month, unprecedented among imagers and sensors using multimode fibers; high spatial resolution of 6 μm; and a compact, all-fiber design without moving components. In addition, we analyzed the resource budget of the sensor and conducted a thorough confidence analysis, which is often neglected for deep learning applications. Our polarization sensor significantly advances the field, answering the demand for a combination of compactness, accuracy, low light sensitivity, and high operational speed, which opens up new possibilities for highly accurate polarization sensing across various applications. Furthermore, our contribution lays down novel research pathways to single-photon-level sensing using random optical networks and media. It also poses new questions on information and confidence evaluation of systems

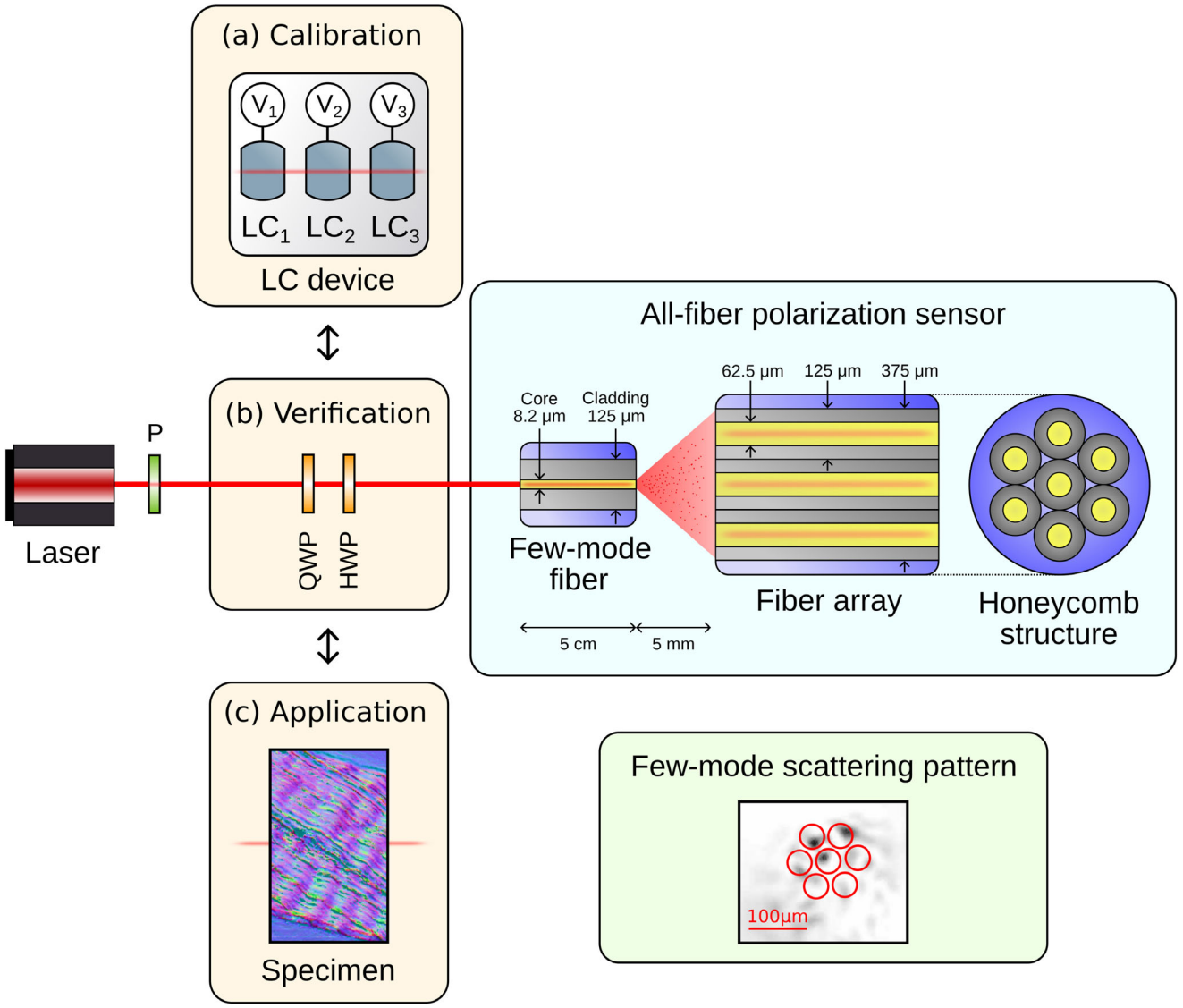


Figure 7. A detailed visualization of the experimental setup for the all-fiber sensor. The twisted nematic liquid crystal device, labeled as (a), is present only during dataset acquisition for deep neural network training. In the verification stage, depicted in (b), a reference polarimeter based on wave plates is employed to validate the accurate operation of the liquid crystal device. Finally, this segment is replaced with a polarization-modulating specimen (c) during the USAF test, dense connective tissue, diatom measurements, and fast liquid crystal transition. The green panel depicts a polarization-dependent speckle pattern with inverted intensities generated by the few-mode fiber. The red circles represent the sparse areas sampled by the subsequent fiber array.

consisting of random media together with deep neural networks. The ability to extract meaningful data from minimal input has far-reaching implications in fields where resource efficiency and sensitivity are paramount.

4. Materials and Methods

Experimental Setup: In addition to the information provided in the main text, a detailed scheme of the all-fiber polarization sensor is depicted in **Figure 7**. This scheme comprised three configurations. In the first configuration, the twisted nematic liquid crystal device[46] induced polarization transformation on the attenuated coherent signal generated by a fiber-coupled continuous 810 nm semiconductor laser diode (QPhotonics QFLD). The laser diode exhibited typical fluctuations, drifts, and mode hops within a 1 nm wavelength range. With a more stable source, such as a distributed-feedback laser diode or solid-state laser, the wavelength stability would increase significantly, which would completely neglect any negative effects on the performance of our sensor. Moreover, the sensor can be recalibrated for an arbitrary wavelength within an hour. The accurate operation of the liquid crystal device in the second configuration was independently verified by replacing it with a reference polarimeter based on rotating wave plates. The third configuration represented the sensor application and comprised the birefringent specimen under investigation. In all cases, the modified signal was collected by the

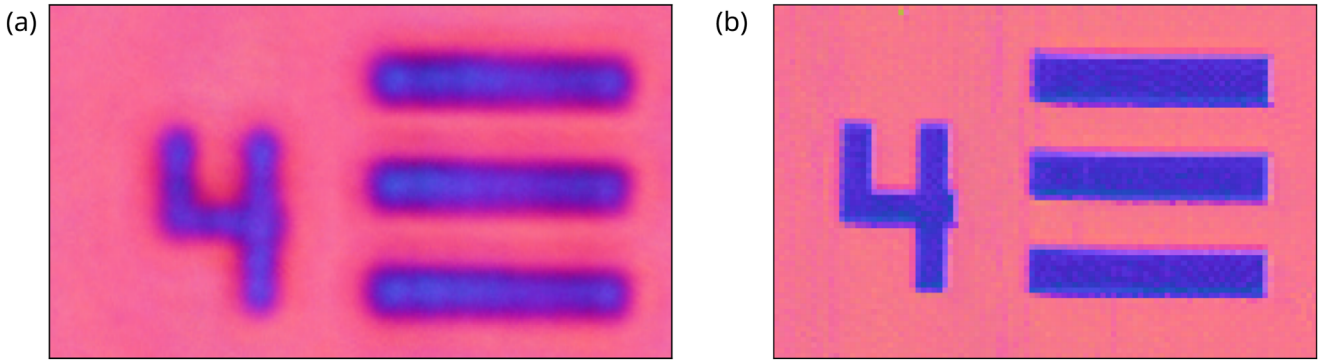


Figure 8. A comparison of polarization measurement on a region of the 1951 USAF birefringent resolution test. a) A polarization image obtained using standard imaging based on rotating wave plates. b) An all-fiber sensor scan of the same region. Both results are visualized as false-colored RGB images using Bloch parameters.

5 cm long SMF28 few-mode fiber. The number of modes supported by this fiber was characterized in Section 4.2.

The generated speckle pattern propagated through a 5 mm free space before reaching the array of seven gradient multimode fibers GIF625 (SQS Vláknová optika, 62.5/125/250 µm MM OM1 Fiber). This propagation allowed the speckle grain size to match the multimode fiber core. Arranged in a honeycomb structure with a 375 µm diameter, the multimode fibers had a core diameter of 62.5 µm and a cladding diameter of 125 µm. Each fiber sparsely sampled the intensity from the speckle pattern and forwarded the signal to an independent single-photon avalanche diode (Excelitas SPCM). The entire experimental setup was located in an unsealed open space of a temperature-maintained laboratory, with the temperature varying within ≈ 1.5 °C and the relative humidity ranging from 25% to 35% during the measurements. Despite these fluctuations, unprecedented stability and accuracy were demonstrated over many weeks (see Section 2.2). This level of robustness was fully sufficient for any laboratory measurements. Moreover, in typical biomedical and material research, the environmental conditions, such as sample temperature, are deliberately controlled and stabilized. Under such conditions, the long-term stability of the sensor will be even greater than demonstrated.

Supported Fiber Modes Analysis: Determining the exact number of modes M supported by the SMF28 few-mode fiber at the given wavelength proves intricate. Most mathematical formulas are precise only in the limit case of a high mode number. Nevertheless, M was estimated by evaluating the normalized frequency parameter V,

$$V = \frac{2\pi}{\lambda} r_{\text{core}} \text{NA}, \tag{1}$$

where $\lambda = 810$ nm is the wavelength of propagated light, $r_{\text{core}} = 4.1$ µm is the fiber core radius, and NA = 0.14 is the numerical aperture. The calculated value of $V = 4.45 > 2.405$ indicates that the fiber supports more than one mode per polarization direction. Subsequently, the number of modes M is approximately given by the relation

$$M \approx \frac{V^2}{2} = 9.9 \tag{2}$$

valid under the assumption that V is large. Since this condition was not fully satisfied, publicly available software developed was also utilized for mode computation in multimode fibers[55] and the number of modes $M = 8$ was obtained.

Numerical Setup: The physical setup described above was assisted by a numerical processing comprising a deep neural network model. This trained network, consisting of 250 neurons per 4 hidden layers, performed a nonlinear transformation on the count distribution and returns four real-valued outputs. These represented elements of a two-by-two triangular matrix τ with real-valued diagonal entries and a single complex off-diagonal entry. Utilizing the Cholesky decomposition, a Hermitian positive-definitive matrix $\tau\tau^\dagger$ was reconstructed. Upon normalizing its trace to unity, $\rho = \frac{\tau\tau^\dagger}{\text{Tr}[\tau\tau^\dagger]}$ represented a physically sound coherence matrix of the polarization state,[45] mathematically equivalent to a density matrix of a two-level quantum system. Therefore, the complete model can be interpreted as providing the polarization coherence matrix of the collected light given the measured count distribution. Alternatively, the polarization state can be described by the Bloch parameters (B_1, B_2, B_3), $\rho = \frac{1}{2}\left(1 + \sum_{j=1}^{3} B_j \sigma_j\right)$, where σ_j are Pauli matrices.

The network was trained using the Adam optimizer and a mean squared error loss function. Various hyperparameters, including network architecture, were fine-tuned to optimize the network performance. For this purpose, a validation set, separated from the training data, was employed to find the optimal hyperparameter combination with respect to a fidelity metric

$$F = \text{Tr}\left[\sqrt{\sqrt{\rho}\cdot\sigma\cdot\sqrt{\rho}}\right]^2 \in [0,1] \tag{3}$$

where ρ is the polarization coherence matrix provided by the network and σ is the ground truth matrix of the input polarization state. This metric, which quantifies the closeness between the states, can be expressed in terms of error as infidelity $1 - F$. Finally, the separated experimentally acquired test set was used for the performance evaluation of the optimal network.

1951 USAF Birefringent Test: A scan of a 1951 USAF birefringent resolution test was conducted with the developed polarization sensor and compared with a polarization measurement based on rotating wave plates. **Figure 8** shows a polariza-

tion image obtained with the standard method (a) alongside the scan using our all-fiber sensor (b). Both panels visualize false-colored RGB images using Bloch parameters (B_1, B_2, B_3) with $R = \frac{B_1+1}{2}$, $G = \frac{B_2+1}{2}$, and $B = \frac{B_3+1}{2}$ corresponding to red, green, and blue channel intensities. By averaging over the background and polarization-modulating regions, a fidelity value of ≈0.987 was estimated between the two polarization measurement approaches. This comparison unequivocally demonstrated the ability of the sensor to perform highly accurate polarization measurements.

Additionally, the image shown in panel (b) allows for the evaluation of the spatial resolution of the sensor. To this end, horizontally and vertically oriented edges were fitted at several positions. Characterized as the 20% to 80% width, the resolution was estimated to be ≈6(5) µm. Besides the distance from a specimen to the fiber tip, the resolution is dictated by the few-mode fiber diameter and can be modified by employing different fiber types. Moreover, for applications requiring higher resolution, utilizing fiber tip tapering could improve it up to the near-field scanning optical microscopy at the cost of the collection efficiency and possibly stability.

Acknowledgements

This work was supported by the Czech Science Foundation (project 21-18545S), Ministry of Education, Youth, and Sports of the Czech Republic (project OP JAC CZ.02.01.01/00/23_021/0008790). M.B. and D.V. acknowledge the support by Palacký University (projects IGA-PrF-2024-008 and IGA-PrF-2025-010). The authors acknowledge the use of cluster computing resources provided by the Department of Optics, Palacký University Olomouc. The authors thank Jan Provazník for maintaining the cluster and providing support. The authors also thank Josef Hloušek for developing a 3D model shown in Figure 1 and Jaromír Běhal for a fruitful discussion on birefringent biological samples.

Open access publishing facilitated by Univerzita Palackeho v Olomouci, as part of the Wiley - CzechELib agreement.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

M.B. and D.V. contributed equally to this work. M.B. assembled the sensor and conducted the experiments. D.V. wrote the deep learning and data processing codes, and drafted manuscript. M.J. initiated and supervised the project. All authors contributed in data interpretation and revising manuscript.

Data Availability Statement

The code and data that support the findings of this study are publicly available on GitHub: https://github.com/VasinkaD/Polarization-Deep-Sense.

Keywords

all-fiber sensor, deep learning, in-situ and microendoscopic measurement, light polarization, single-photon and low-light sensing

Received: July 8, 2025
Revised: September 19, 2025
Published online:

[1] D. M. Jameson, J. A. Ross, *Chem. Rev.* **2010**, *110*, 2685.
[2] C. V. Rimoli, C. A. Valades-Cruz, V. Curcio, M. Mavrakis, S. Brasselet, *Nat. Commun.* **2022**, *13*, 301.
[3] Z. Zhan, M. Cantono, V. Kamalov, A. Mecozzi, R. Müller, S. Yin, J. C. Castellanos, *Science* **2021**, *371*, 931.
[4] J. Liu, D. Zhang, D. Yu, M. Ren, J. Xu, *Light Sci. Appl.* **2021**, *10*, 55.
[5] S. Wen, X. Xue, S. Wang, Y. Ni, L. Sun, Y. Yang, *Light Sci. Appl.* **2024**, *13*, 88.
[6] J. Wang, *Photonics Res.* **2016**, *4*, B14.
[7] Z.-Y. Chen, L.-S. Yan, Y. Pan, L. Jiang, A.-L. Yi, W. Pan, B. Luo, *Light Sci. Appl.* **2016**, *6*, e16207.
[8] J. Qi, D. S. Elson, *J. Biophotonics* **2017**, *10*, 950.
[9] P. Doradla, K. Alavi, C. Joseph, R. Giles, *J. Biomed. Opt.* **2013**, *18*, 090504.
[10] J. Qi, T. Tatla, E. Nissanka-Jayasuriya, A. Y. Yuan, D. Stoyanov, D. S. Elson, *Nat. Biomed. Eng.* **2023**, *7*, 971.
[11] S. Turtaev, I. T. Leite, T. Altwegg-Boussac, J. M. P. Pakan, N. L. Rochefort, T. Čižmár, *Light Sci. Appl.* **2018**, *7*, 1.
[12] Z. Wen, Z. Dong, Q. Deng, C. Pang, C. F. Kaminski, X. Xu, H. Yan, L. Wang, S. Liu, J. Tang, W. Chen, X. Liu, Q. Yang, *Nat. Photonics* **2023**, *17*, 679.
[13] M. Valentino, J. Běhal, V. Bianco, S. Itri, R. Mossotti, G. D. Fontana, T. Battistini, E. Stella, L. Miccio, P. Ferraro, *Sci. Total Environ.* **2022**, *815*, 152708.
[14] J. Běhal, M. Valentino, L. Miccio, V. Bianco, S. Itri, R. Mossotti, G. Dalla Fontana, E. Stella, P. Ferraro, *ACS Photonics* **2022**, *9*, 694.
[15] Y. Deng, M. Wang, Y. Zhuang, S. Liu, W. Huang, Q. Zhao, *Light Sci. Appl.* **2021**, *10*, 76.
[16] S. Coppola, V. Vespini, J. Behal, V. Bianco, L. Miccio, S. Grilli, L. De Sio, P. Ferraro, *ACS Appl. Mater. Interfaces* **2024**, *16*, 19453.
[17] S. Rotter, S. Gigan, *Rev. Mod. Phys.* **2017**, *89*, 015005.
[18] J. Bertolotti, O. Katz, *Nat. Phys.* **2022**, *18*, 1008.
[19] H. Cao, A. P. Mosk, S. Rotter, *Nat. Phys.* **2022**, *18*, 994.
[20] K. Vynck, R. Pierrat, R. Carminati, L. S. Froufe-Pérez, F. Scheffold, R. Sapienza, S. Vignolini, J. J. Sáenz, *Rev. Mod. Phys.* **2023**, *95*, 045003.
[21] H. Cao, T. Čižmár, S. Turtaev, T. Tyc, S. Rotter, *Adv. Opt. Photonics* **2023**, *15*, 524.
[22] E. Valent, Y. Silberberg, *Optica* **2018**, *5*, 204.
[23] M. W. Matthès, Y. Bromberg, J. de Rosny, S. M. Popoff, *Phys. Rev. X* **2021**, *11*, 021060.
[24] T. Čižmár, K. Dholakia, *Nat. Commun.* **2012**, *3*, 1027.
[25] M. Plöschner, T. Tyc, T. Čižmár, *Nat. Photonics* **2015**, *9*, 529.
[26] N. Borhani, E. Kakkava, C. Moser, D. Psaltis, *Optica* **2018**, *5*, 960.
[27] B. Rahmani, D. Loterie, G. Konstantinou, D. Psaltis, C. Moser, *Light Sci. Appl.* **2018**, *7*, 69.
[28] P. Caramazza, O. Moran, R. Murray-Smith, D. Faccio, *Nat. Commun.* **2019**, *10*, 2029.
[29] B. Rahmani, D. Loterie, E. Kakkava, N. Borhani, U. Teğin, D. Psaltis, C. Moser, *Nat. Mach. Intelligence* **2020**, *2*, 403.
[30] R. Ziv, A. Dikopoltsev, T. Zahavy, I. Rubinstein, P. Sidorenko, O. Cohen, M. Segev, *Opt. Express* **2020**, *28*, 7528.
[31] W. Xiong, B. Redding, S. Gertler, Y. Bromberg, H. D. Tagare, H. Cao, *APL Photonics* **2020**, *5*, 096106.
[32] H. Defienne, M. Barbieri, I. A. Walmsley, B. J. Smith, S. Gigan, *Sci. Adv.* **2016**, *2*, e1501054.
[33] S. Leedumrongwatthanakun, L. Innocenti, H. Defienne, T. Juffmann, A. Ferraro, M. Paternostro, S. Gigan, *Nat. Photonics* **2020**, *14*, 139.
[34] N. H. Valencia, S. Goel, W. McCutcheon, H. Defienne, M. Malik, *Nat. Phys.* **2020**, *16*, 1112.
[35] D. Pierangeli, G. Marcucci, C. Conti, *Photonics Res.* **2021**, *9*, 1446.
[36] B. Courme, P. Cameron, D. Faccio, S. Gigan, H. Defienne, *PRX Quantum* **2023**, *4*, 010308.
[37] B. Redding, S. F. Liew, R. Sarma, H. Cao, *Nat. Photonics* **2013**, *7*, 746.

[38] R. K. Gupta, G. D. Bruce, S. J. Powis, K. Dholakia, *Laser Photonics Rev.* **2020**, *14*, 2000120.
[39] M. Juhl, K. Leosson, *ACS Photonics* **2020**, *7*, 203.
[40] M. Facchin, G. D. Bruce, K. Dholakia, *OSA Continuum* **2020**, *3*, 1302.
[41] N. A. Rubin, P. Chevalier, M. Juhl, M. Tamagnone, R. Chipman, F. Capasso, *Opt. Express* **2022**, *30*, 9389.
[42] D. Pierangeli, C. Conti, *Nat. Commun.* **2023**, *14*, 1831.
[43] M. Stibůrek, P. Ondráčková, T. Tučková, S. Turtaev, M. Šiler, T. Pikálek, P. Jákl, A. Gomes, J. Krejčí, P. Kolbábková, H. Uhlířová, T. Čižmár, *Nat. Commun.* **2023**, *14*, 1897.
[44] S. Goel, C. Conti, S. Leedumrongwatthanakun, M. Malik, *Opt. Express* **2023**, *31*, 32824.
[45] S. Huard, *Polarization of Light*, Wiley, New York, **1997**.
[46] M. Bielak, R. Stárek, V. Krčmarský, M. Mičuda, M. Ježek, *Opt. Express* **2021**, *29*, 33037.
[47] D. Vašinka, M. Bielak, M. Neset, M. Ježek, *Phys. Rev. Appl.* **2022**, *17*, 054042.
[48] Q. Wang, G. Farrell, W. Yan, *J. Lightwave Technol.* **2008**, *26*, 512.
[49] Q. Wu, Y. Qu, J. Liu, J. Yuan, S.-P. Wan, T. Wu, X.-D. He, B. Liu, D. Liu, Y. Ma, Y. Semenova, P. Wang, X. Xin, G. Farrell, *IEEE Sens. J.* **2021**, *21*, 12734.
[50] P. Mujal, R. Martínez-Peña, J. Nokkala, J. García-Beni, G. L. Giorgi, M. C. Soriano, R. Zambrini, *Adv. Quantum Technol.* **2021**, *4*, 2100027.
[51] T. Wang, S.-Y. Ma, L. G. Wright, T. Onodera, B. C. Richard, P. L. McMahon, *Nat. Commun.* **2022**, *13*, 123.
[52] J. Hüpfl, F. Russo, L. M. Rachbauer, D. Bouchet, J. Lu, U. Kuhl, S. Rotter, *Nat. Phys.* **2024**, *20*, 1294.
[53] S. Rotter, in *Machine Learning and (Quantum) Physics Workshop*, Obergurgl, Austria, Obergurgl, Austria, **2023**, http://hdl.handle.net/20.500.12708/176580.
[54] S. Rotter, in *Nonlinear Optics and its Applications 2024* (Eds: A. C. Peacock, G. Tissoni, J. M. Dudley, B. Stiller), Vol. *PC13004*, SPIE Photonics Europe, Strasbourg, France **2024**, PC1300401.
[55] S. M. Popoff, Replicate repository **2018**, https://replicate.com/wavefrontshaping/pymmf.

PHYSICS

Deep learning of quantum entanglement from incomplete measurements

Dominik Koutný[1]*, Laia Ginés[2], Magdalena Moczała-Dusanowska[3], Sven Höfling[4], Christian Schneider[5], Ana Predojević[2], Miroslav Ježek[1]*

The quantification of the entanglement present in a physical system is of paramount importance for fundamental research and many cutting-edge applications. Now, achieving this goal requires either a priori knowledge on the system or very demanding experimental procedures such as full state tomography or collective measurements. Here, we demonstrate that, by using neural networks, we can quantify the degree of entanglement without the need to know the full description of the quantum state. Our method allows for direct quantification of the quantum correlations using an incomplete set of local measurements. Despite using undersampled measurements, we achieve a quantification error of up to an order of magnitude lower than the state-of-the-art quantum tomography. Furthermore, we achieve this result using networks trained using exclusively simulated data. Last, we derive a method based on a convolutional network input that can accept data from various measurement scenarios and perform, to some extent, independently of the measurement device.

INTRODUCTION

Physical measurements performed on individual parties of an entangled system reveal strong correlations (*1*), which give rise to nonclassical and nonlocal effects (*2*, *3*). Aforesaid effects are the essential element of fundamental tests of quantum mechanics, including direct experimental verification of quantum nonlocality (*4*–*6*). The critical role of entanglement was demonstrated also on the opposite scale of the complexity spectra in macroscopic phase transitions (*7*–*9*). Besides the fundamental aspects, entanglement is an essential tool for quantum information processing, and it allows for reaching the quantum advantage (*10*, *11*). Modern quantum communication networks rely crucially on entanglement sources (*12*–*15*). Consequently, the characterization of entanglement is paramount for both fundamental research and quantum applications (*16*, *17*).

Here, we adopt methods of deep learning to tackle the longstanding problem of efficient and accurate entanglement quantification. Our approach determines the degree of entanglement of a generic quantum state directly from an arbitrary set of local measurements. Despite the deep learning models being trained on simulated measurements, they excel when applied to real-world measurement data. We quantify photonic entanglement generated by two distinct systems: a nonlinear parametric process and a semiconductor quantum dot.

Reliable entanglement quantification represents an open problem in quantum physics. Direct measurement of entanglement can be achieved by exploiting quantum interference of two (or more) identical copies of a physical system (*18*–*22*). This multicopy approach roots in measuring nonlinear functions of quantum states (*23*, *24*). However, such measurements are experimentally highly demanding, which has spurred the research of single-copy entanglement detection using only local measurements, such as quantum tomography.

Quantum tomography provides the full description of a quantum state including the degree of entanglement (*25*, *26*). However, the total number of measurements required for quantum tomography increases exponentially with the number of qubits or quantum degrees of freedom, which renders the approach inherently not scalable (*27*–*29*). Several methods have been developed to make this scaling more favorable, nevertheless, by imposing an a priori structure or symmetry to the system (*30*–*33*). When a few-parameter model of quantum state is assumed, quantum estimation can be used for optimal inferring of the state entanglement (*34*–*36*). Another approach to emulate quantum correlations (*37*) with fewer resources relies on neural network quantum states (*38*–*42*). However, this method suffers from the sign problem, solving of which requires further assumptions about the state (*43*, *44*). The neural network quantum state approach was used for quantum tomography under nonideal experimental conditions (*45*–*49*). However, how much information is needed for representing a generic quantum state at a given level of accuracy remains an open question (*50*, *51*).

Instead of characterizing the whole system, one might target only mean values of a set of selected observables, which substantially reduces the required number of measurements. This approach, termed shadow tomography (*52*), can also be applied to estimate entanglement entropy of a small subsystem, basically reconstructing its reduced quantum state (*53*, *54*). An alternative method uses random measurements to estimate the second-order Rényi entropy of a subsystem (*55*–*58*). However, quantification of entanglement distributed over the whole system lies beyond the scope of such methods.

Entanglement witnessing seems to be a viable alternative to the tomographic methods, when we only aim at distinguishing between entangled and nonentangled states (or between entanglement classes) without quantifying the degree of entanglement and its

[1]Department of Optics, Faculty of Science, Palacký University, 17. listopadu 12, 77146 Olomouc, Czechia. [2]Department of Physics, Stockholm University, 10691 Stockholm, Sweden. [3]Princeton Institute of Materials, Princeton University, Princeton, NJ 08544, USA. [4]Technische Physik, Physikalisches Institut and Würzburg-Dresden Cluster of Excellence ct.qmat, Universität Würzburg, Am Hubland, D-97074 Würzburg, Germany. [5]Institute of Physics, University of Oldenburg, D-26129 Oldenburg, Germany.
*Corresponding author. Email: koutny@optics.upol.cz (D.K.); jezek@optics.upol.cz (M.J.)

detailed structure. Nevertheless, the witnessing may still require the full knowledge of the underlying quantum state, as is the case of the positive partial transpose criterion (*1*). The witness cannot be directly measured; however, it can be approximated using a completely positive map (*59*), which is equivalent to the full quantum state tomography (*60*, *61*). Other witnessing methods are based on the minimum local decomposition (*62*, *63*), semidefinite programming (SDP) (*64*, *65*), entanglement polytopes (*66*), or correlations in random measurements (*67*, *68*). Entanglement witnessing can also be facilitated by using neural networks classifiers (*69*–*71*). Despite the success of the entanglement witnessing, it provides only witnesses or lower bounds and often requires some a priori information about the state.

In summary, the connection between the entanglement present in a physical system and the measurements of the correlations of its subsystems is highly nontrivial (*72*, *73*). It seems that full entanglement characterization using single-copy local measurements can only be accomplished with the complete quantum state tomography and, consequently, with exponential scaling of the number of required measurements (*60*, *61*, *74*, *75*). The open question remains what one can learn about entanglement from an incomplete observation.

In this work, we use deep neural networks (DNNs) to tackle the problem of entanglement characterization. We develop a method that allows us to quantify the degree of entanglement and quantum correlations in a generic partially mixed state using a set of informationally incomplete measurements. The entanglement quantifiers that we obtain using DNN approach are substantially more accurate compared with the values attainable using the state-of-the-art quantum tomography methods. In addition, we demonstrate a measurement-independent quantification of entanglement by developing a deep convolutional network that accepts an arbitrary set of projective measurements without retraining. The DNN-based approaches that we introduce here can be immediately applied for certification and benchmarking of entanglement sources, which we demonstrate by using photonic sources of entangled photons based on spontaneous parametric downconversion and a semiconductor quantum dot.

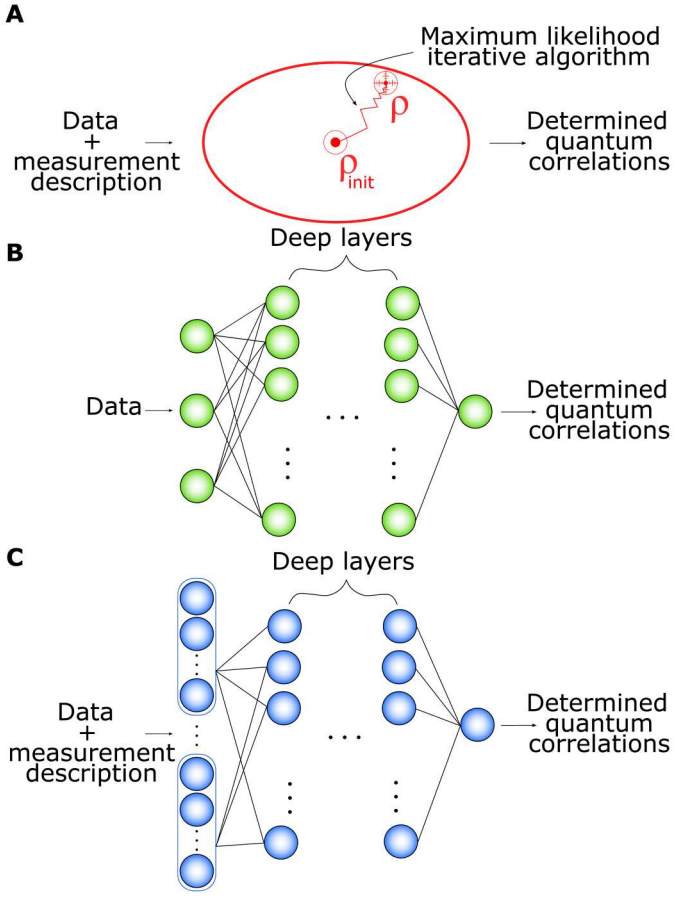


Fig. 1. Schematics of the three methods that we used to infer the quantum correlations. (**A**) The maximum likelihood algorithm (MaxLik) finds the most likely quantum state ρ based on the measured data and an initial guess ρ_{init}. (**B**) Green DNN represents a fully connected neural network that infers directly the concurrence and the mutual information from specific measurements (specific measurement projectors), whereas (**C**) the blue DNN works with an arbitrary measurement projectors. The input for the former is the measured data. The measurement-independent DNN has a first layer convolutional, and it inputs both the data and the measurement description.

RESULTS

Even in a well-understood system, such as a pair of qubits, a reliable quantification of entanglement requires full state tomography (*74*). In other words, to infer the degree of entanglement, we need to determine the quantum state. A common approach to implement photonic qubit tomography is to measure the full basis of three Pauli operators. Such a measurement for a two-qubit state consist of $6^2 = 36$ local projectors (*26*). Omitting randomly some projectors in this measurement scheme decreases the accuracy of the quantum tomography and, consequently, the entanglement evaluation. Here, we show that this problem can be overcome by using DNNs that allow us to gain knowledge on the degree of entanglement without the need to know the quantum state.

To demonstrate the advantage of the DNN approach, we use two quantifiers: the concurrence (*1*) and the mutual information (*76*) for a two-qubit and a three-qubit system, respectively. The concurrence is widely used in experiments for characterization of entangled photon pair sources. Its value is bounded from below by 0 for separable states and from above by 1 for maximally entangled states. On the other hand, the concurrence cannot be easily generalized to higher-dimensional quantum systems and systems of more than two parties. Therefore, the second quantifier that we use is the mutual information, which can be generalized to multipartite systems of qudits, and its value reflects the information shared between the parties of a larger system.

We use three different approaches to determine the concurrence and the mutual information from an incomplete set of data. We show them schematically in Fig. 1. We use the maximum likelihood algorithm (MaxLik) (Fig. 1A, red), measurement-specific DNNs (Fig. 1B, green), and a measurement-independent DNN (Fig. 1C, blue). The maximum likelihood is an algorithm that finds the quantum state (ρ) iteratively, starting from an initial guess (ρ_{init}), which is typically set to maximally mixed state (*77*). Having at hand the quantum state ρ allows us to quantify the entanglement (see Materials and Methods). In contrast, the approaches based on DNN learn the concurrence and mutual information directly from the measured data. While the measurement-specific DNN is

designed for a predefined set of measurement projectors M_m, the measurement-independent DNN relaxes the condition on measuring the a priori known projectors and predicts concurrence and mutual information independently on the measurement settings. This approach has a convolutional first layer, and it inputs the measured data together with the description of the respective projectors. During the training, the DNNs are provided with the theoretical probabilities $\mathrm{Tr}\{\rho M_m\}$ and, in the case of the measurement-independent DNN, also the description of the measurement M_m. For the detailed information about the structure of the DNNs, the dataset, and the training procedure, see Materials and Methods.

We compare the three approaches on the basis of how accurately they can infer the concurrence (mutual information) from an incomplete set of data. Here, the MaxLik serves as a benchmark to the other two methods that are DNN-based. We chose to evaluate the performance of all three approaches by computing the mean absolute error (MAE). The MAE is calculated as $\langle |x_i - y_i| \rangle$ with x_i being the true value and y_i being the predicted value of the concurrence (mutual information). To make our comparison universal, the average is taken over a set of states and several combinations of measurement projectors, i.e., a test set. The total number of combinations of k projectors from the maximum of 36 is $\frac{36!}{k!(36-k)!}$. As this number can be excessively high, we randomly selected a smaller subset of combinations. Therefore, to evaluate the performance of a measurement-specific DNN, we train 12 randomly selected networks for each k-projector measurement and evaluate the average and SD of their MAEs. For the MaxLik and measurement-independent DNN, the averaging is performed over hundreds of randomly selected measurements.

The performance of the three approaches is presented in Fig. 2, where we show how MAE depends on the number of measurement projectors that we used to obtain the result. Figure 2 (A and B) shows the MAE for the concurrence and the mutual information, respectively, while Fig. 2C addresses the MAE of the three-qubit mutual information matrix. The MaxLik approach is presented using the red triangles in all panels. For the informationally complete data, i.e., when all 36 projectors are measured, the MaxLik MAE is on the order of 10^{-5} to 10^{-4}. In this scenario, MaxLik converges to the true quantum state, and the error only reflects the numerical errors caused by the computing precision. As we can see in Fig. 2 (A and B), the MAE of the MaxLik starts increasing if only a few out of the 36 projectors are absent. In contrast to the MaxLik, DNNs perform well even for a severely reduced number of projectors. The measurement-specific DNNs (shown in green circles) predict the concurrence and mutual information with the MAE of approximately 0.01 even when only 24 projectors are used. For the same number of projectors, MaxLik MAE is 0.1. Consequently, measurement-specific DNNs result in a precision that is, on average, 10 times higher. If we further reduce the number of projectors, then the MAE for the measurement-specific DNNs starts to increase; however, it keeps being substantially smaller compared to the MAE of the MaxLik. The uncertainty region of MAE also remains at least two times smaller (up to 10 times while working with more than 18 projectors). The measurement-independent DNN error is shown in Fig. 2 using the blue squares. Compared to the performance of the MaxLik, the measurement-independent DNN quantifies the concurrence and the mutual information with a lower MAE however worse than using the measurement-specific strategy. In practice, one can resource to the measurement-independent DNN for preliminary detecting the entanglement in the system, even changing the measurement on the fly, and improve the entanglement quantification by training a particular measurement-specific DNN later.

To further validate our approach, we compare the values of the concurrence determined by MaxLik, measurement-specific DNNs, and measurement-independent DNN using a state that the network has never seen before, the Werner state $\rho_W(p) = p\rho_{\psi^-} + \frac{1-p}{4}1$, where ρ_{ψ^-} is a projector into maximally entangled Bell state spanning the asymmetric subspace of two qubits. The parameter p runs from 0 (mixed state) to 1 (maximally entangled state). The concurrence for the Werner state is a piecewise linear function of the parameter p, and it takes the exact form $C(\rho_W) = \max[0, (3p-1)/2]$. The results are shown in Fig. 3. In the panels (A to D), we show the concurrence and the corresponding uncertainty regions for 36, 28, 18, and 8 projective measurements, respectively. For 28 and 18 measurement projectors, both the DNN approaches follow the ideal concurrence values, while the MaxLik deviates substantially. The measurement-specific DNNs yield nontrivial results even in the case of only 8 measurement projections.

As mentioned previously, the mutual information can be generalized to the systems of more than two qubits. To show that we can also generalize the DNN-based approach to larger quantum

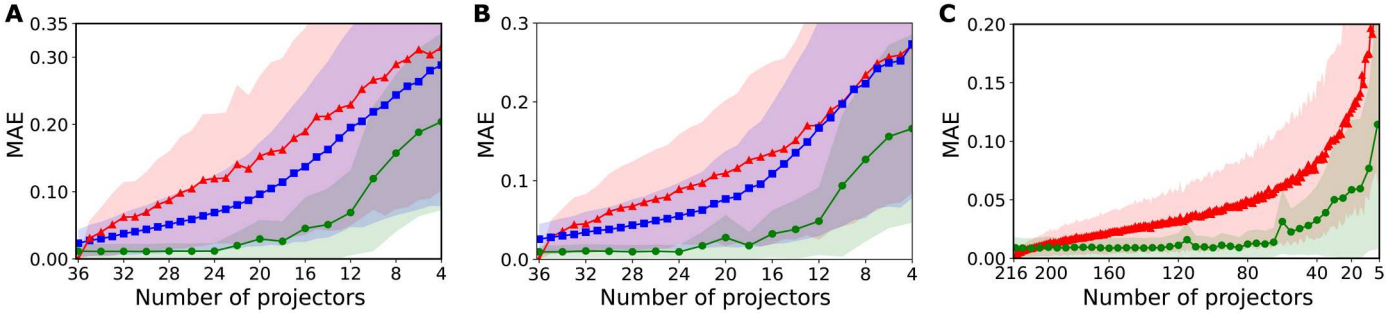


Fig. 2. Entanglement quantification error for the two- and three-qubit systems. The mean absolute error (MAE) versus the number of measurement projectors for (**A**) two-qubit concurrence, (**B**) two-qubit mutual information, and (**C**) three-qubit mutual information matrix. Red triangles depict MAE for the MaxLik, blue squares stand for the values of MAE computed from measurement-independent DNN, and, lastly, green circles represent the values of MAE computed from measurement-specific DNNs. The uncertainty regions are depicted in the corresponding colors and may overlap. The DNNs outperform the MaxLik approach in terms of entanglement quantification accuracy and its consistency, given by smaller errors and uncertainty intervals, even for substantially incomplete measurements.

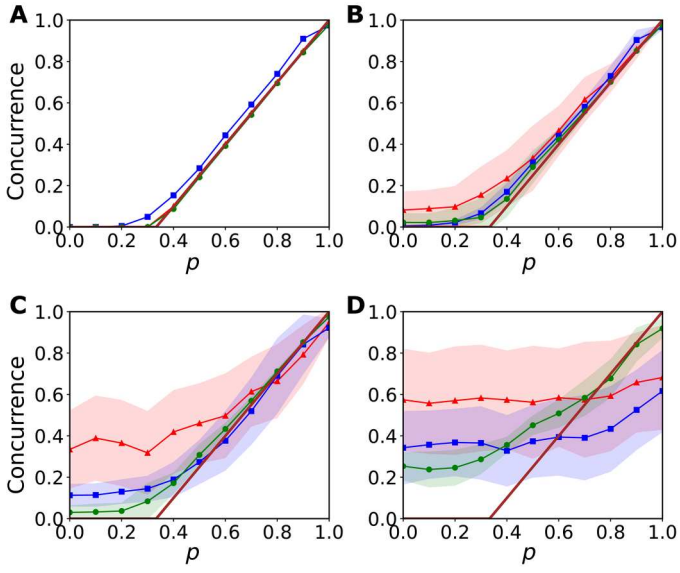


Fig. 3. Entanglement quantification error for the Werner state. The dependence of values of concurrence for the two-qubit Werner state $\rho_W(p)$ characterized by the parameter $p \in [0,1]$. Values of the concurrence determined from (**A**) 36, (**B**) 28, (**C**) 18, and (**D**) 8 measurement settings. In each panel, the red triangles depict the average values of the concurrence determined by the MaxLik with corresponding uncertainty region, the blue squares stand for the measurement-independent DNN predictions, and the green circles represent predictions given by measurement-specific DNNs. The brown line shows the theoretical values of the concurrence for the Werner state. Both measurement-independent DNN and measurement-specific DNNs outperform the MaxLik in entanglement quantification of the Werner state.

systems, we apply our method to a three-qubit system. In such a system, the mutual information matrix has three independent entries $\mathcal{I} \equiv \{\mathcal{I}_{AB}, \mathcal{I}_{AC}, \mathcal{I}_{BC}\}$, with subscripts referring to the three different ways of partition. To determine all three numbers $\mathcal{I}$ simultaneously, we have to perform a full tomographic measurement on each qubit, which leave us with $6^3 = 216$ projections. Following the procedure introduced for the two-qubit case, we built measurement-specific DNNs, each mapping measurement data to the three-component vector $\mathcal{I}$. Deep layers have the same structure as for quantification of mutual information in the two-qubit case. Final results are shown in Fig. 2C. The MAE of $\mathcal{I}$ is averaged over its three independent elements and over randomly generated quantum states. DNN predictions are, on average, akin to the MaxLik ones in the regime close to the complete data. However, with only about a third of all projections, measurement-specific DNNs predict the full mutual information matrix on average with a five times smaller error than the MaxLik.

Table 1. The summary of the mutual information quantification from incomplete measurements consisting of one-fourth of all possible Pauli projectors in each case. The MaxLik and the measurement-specific DNNs are compared up to five-qubit quantum systems. The ratio of the mean absolute errors (MAEs) of the methods shows an increasing improvement in the performance of the DNN approach for entanglement quantification in higher-dimensional systems.

Number of qubits	MAE		Ratio of MaxLik and DNN MAEs
	MaxLik	DNN	
2	0.20 ± 0.16	0.09 ± 0.09	2.2
3	0.068 ± 0.055	0.023 ± 0.020	3.0
4	0.019 ± 0.014	0.005 ± 0.001	3.8
5	0.039 ± 0.032	0.009 ± 0.001	4.3

Our approach needs modest computational resources. In particular, the two- and three-qubit measurement-specific networks (for one-fourth of all Pauli projectors compared to the complete measurement) have approximately 37,000 and 42,000 parameters, respectively. The optimal performance of networks for three qubits does not require substantially more parameters than for two-qubit networks. We further verified this optimistic scaling by training four- and five-qubit measurement-specific networks (for one-fourth of all possible Pauli projectors in each case). These networks require 69,000 and 231,000 parameters, respectively, and outperform the MaxLik even more than two-qubit and three-qubit networks; see Table 1. Namely, the measurement specific networks reach 2.2, 3.0, 3.8, and 4.3 times lower MAE of mutual information matrix than the MaxLik for two, three, four, and five qubits, respectively. On the basis of this finding, we expect that, by keeping the ratio of the MaxLik accuracy and the DNN accuracy constant, the required fraction of the projectors with respect to the full tomography will decrease.

Last, we demonstrate the performance of DNN-based entanglement quantification using experimental data acquired under nonideal conditions and with limited statistical sampling. We study two distinct entanglement sources. The first one is based on continuously pumped spontaneous parametric downconversion. The photon pair generation process is inherently random, and the resulting entangled state depends on the choice of the temporal coincidence window and other experimental conditions such as background noise. Adjusting of the experimental parameters affects the degree of entanglement in the produced state. We quantify concurrence using the DNNs and the MaxLik approach for various experimental settings ranging from the maximally entangled singlet Bell state to a noisy state with a negligible concurrence. Figure 4 (A and B) shows the results for an almost pure entangled state and a partially mixed state with the concurrence of 0.985 ± 0.001 and 0.201 ± 0.002, respectively. In both cases, DNN approaches outperform the MaxLik approach. The measurement-specific DNNs remain very accurate (MAE <0.04) all the way down to 14 projections. Even the measurement-independent DNN outperforms the MaxLik in the generic case of partially mixed state for any number of measurement projectors. The maximally entangled state represents the only case where the MaxLik performs slightly better than the measurement-independent DNN (but worse than measurement-specific DNNs). This behavior results from high purity and sparsity of the state and, consequently, from the sparsity of the measurement data. When randomly selecting a subset of projectors, there is a high possibility of having a majority or even all the measurements with a negligible number of detection counts. It seems that the predictive strength of the measurement-independent DNN is limited for such a scenario. However, the MaxLik approach is biased toward pure states in the case of heavily undersampled data (*78*, *79*), and the positivity constraint tends to a sparse (low-rank)

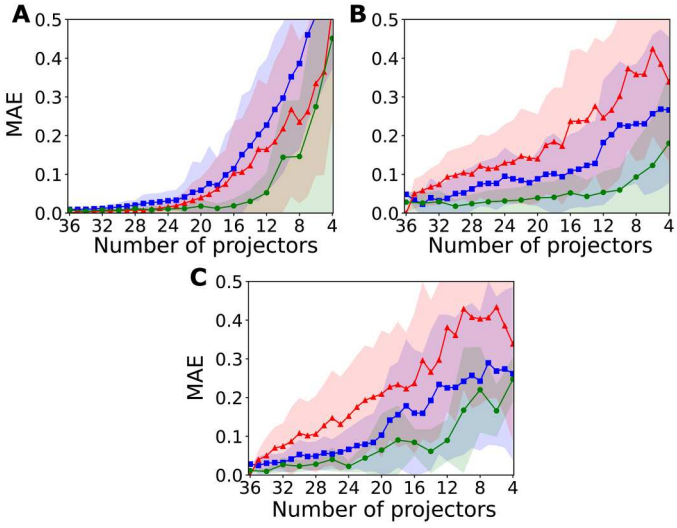


Fig. 4. Performance of MaxLik and DNN-based approaches for an experimental datasets. We show the dependence of the MAE on the number of projectors. (**A** and **B**) Spontaneous parametric downconversion sources and (**C**) semiconductor quantum dot source. The concurrence of experimentally prepared quantum states was determined from the full MaxLik tomography to (A) 0.985 ± 0.001, (B) 0.201 ± 0.002, and (C) 0.18 ± 0.01. The MAE for the measurement-specific DNNs is depicted in green circles, for measurement-independent DNN in blue squares, and for the MaxLik approach in red triangles.

states (*80*). This bias artificially increases the resulting concurrence and reduces its error.

The second experimental system consists of a semiconductor quantum dot resonantly pumped by picosecond pulses. The biexciton-exciton cascade emission produces pairs of photons in a partially polarization entangled state. The degree of entanglement is reduced by the presence of the fine-structure splitting, reaching the concurrence of 0.18 ± 0.01. Figure 4C shows the MAE for such a mixed quantum state. As for the source based on spontaneous downconversion, both DNN approaches beat, on average, the MaxLik method in accuracy. Let us point out that the DNN-based approaches were trained to predict quantum correlations from the theoretical probabilities computed from the ideal quantum states and measurement. Figure 3 thus demonstrates the robustness of our approaches to noisy experimental data.

DISCUSSION

We demonstrated that, by exploiting novel methods of neural networks and deep learning, we can outperform the traditional and commonly used techniques for quantification of quantum correlations such as state tomography. For the systems of two qubits, we built two different neural network–based approaches, namely, measurement-specific and measurement-independent DNNs. Both approaches predict concurrence and mutual information from data with a higher accuracy than the commonly used quantum state tomography. The best performing approach is the measurement-specific DNNs, which are trained to predict the concurrence or mutual information from a fixed set of projectors. Furthermore, we generalized to the system of three qubits, where we show that the measurement-specific DNNs represent a more accurate method to quantify the mutual information matrix than the maximum likelihood one. We demonstrated the feasibility of the measurement-specific DNNs training up to five qubits. Our approaches not only benefit from a high accuracy when working with fewer measurement projectors but also are substantially faster compared to the standard tomography–based methods. Furthermore, we demonstrate the robustness of our approach using two experimental systems: a nonlinear parametric process and a semiconductor quantum dot. The DNN approaches can be further studied and modified to adaptively find a minimal set of projectors that infer the entanglement accurately.

MATERIALS AND METHODS

Quantifying quantum correlations

To quantify the quantum correlation, we use the concurrence and the mutual information, for the two- and three-qubit cases, respectively. The concurrence is a two-qubit monotone entanglement measure (*1*) widely used for the characterization of bipartite entanglement commonly present in sources of entangled photon pairs. Knowing the quantum state the concurrence is defined as

$$\mathscr{C}(\rho) \equiv \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\} \quad (1)$$

with λ_1, ..., λ_4 being the eigenvalues (sorted in decreasing order) of the Hermitian matrix $T = \sqrt{\sqrt{\rho}\tilde{\rho}\sqrt{\rho}}$, here $\tilde{\rho} = \sigma_y \otimes \sigma_y \rho^* \sigma_y \otimes \sigma_y$, where ρsup>/sup> standing for complex conjugate and σ_y is one of the Pauli matrices represented in a computational basis as $\sigma_y = i(|1\rangle\langle 0| - |0\rangle\langle 1|)$. For an arbitrary mixed state, the value of concurrence is saturated from below by 0 (*1*) for the separable states $\rho_{AB} = \sum_i \gamma_i \rho_A^i \otimes \rho_B^i$ and from above by 1 for the maximally entangled states of two-qubits.

Mutual information is a quantum correlation measure commonly used in quantum cryptography or for quantifying complexity in many-body systems. For an *n*-qubit quantum system, mutual information matrix reads

$$\mathscr{I}_{ij} = \frac{1}{2}[\mathscr{S}(\rho_i) + \mathscr{S}(\rho_j) - \mathscr{S}(\rho_{ij})] \quad (2)$$

and is constructed from the one and two point von Neumann entropies (*76*), $\mathscr{S}(\rho_i) = -\mathrm{Tr}\{\rho_i \log_d \rho_i\}$ and $\mathscr{S}(\rho_{ij}) = -\mathrm{Tr}\{\rho_{ij} \log_d \rho_{ij}\}$, with ρ_i and ρ_{ij} standing for reduced density matrices, $\rho_i = \mathrm{Tr}_{k \neq i}\{\rho\}$ and $\rho_{ij} = \mathrm{Tr}_{k \neq ij}\{\rho\}$ respectively.

Quantum state tomography

Quantum state tomography is a method to solve the inverse problem of reconstruction of an unknown quantum state. It uses the set of measurement operators (projectors) and relative frequencies $\{f_i\}$ acquired in a measurement. To obtain an informationally complete measurement, we need the relative frequencies for at least $D^2 - 1$ independent projectors $\{M_i\}_{i=0}^N$. Quantum state is reconstructed by maximizing the log-likelihood functional $\mathscr{L}(\rho) \propto \sum_{j=1}^N f_j \log p_j(\rho)$, which can be written (*77*, *81*, *82*) as the iterative map $\rho^{(k+1)} \leftarrow \mu_k R \rho^{(k)} R$, where μ is the normalization constant and *R* is an operator defined as $R = \sum_i f_i / p_i M_i$. Here, f_i are the measured frequencies and p_i are the theoretical probabilities given by the Born's rule $p_i = \mathrm{Tr}\{\rho M_i\}$. In the Results section of this paper, we address how the measurement being incomplete affects the quantification accuracy of the concurrence and the mutual

information. In such a case, the closure relation, $\sum_i M_i = 1$, is no longer fulfilled. The optimal strategy is to map the set of projectors $\{M_i\}$ into a new set $\{M_{i'}\}$ via $M_{i'} = G^{-1/2} M_i G^{-1/2}$ with $G = \sum_i M_i$. One can easily check that the $\{M_{i'}\}$ now fulfill the completeness relation, $\sum_i M_{i'} = 1$. The iterative map then updates to

$$\rho^{(k+1)} \leftarrow \mu_k G^{-1/2} R \rho^{(k)} R G^{-1/2} \tag{3}$$

and represents a procedure that we follow in the main text. We consider measurement settings to be the Pauli projectors, i.e., projectors into eigenstates of Pauli operators $\{\sigma_x, \sigma_y, \sigma_z\}$. The MaxLik estimator ρ^{MaxLik} is defined as a fixed point of the iterative map (Eq. 3). The iteration process starts from the completely mixed state $\rho_{\text{init}} = 1/D$ and is stopped when the Hilbert-Schmidt distance between the subsequent iterations reaches 10^{-16}. In the case of encoding qubit states into the polarization degrees of freedom, Pauli measurement consists of projectors onto three mutually unbiased basis sets $\{|H\rangle\langle H|, |V\rangle\langle V|, |D\rangle\langle D|, |A\rangle\langle A|, |R\rangle\langle R|, |L\rangle\langle L|\}$.

There are other methods for quantum state tomography such as maximum likelihood–maximum entropy (*83*), SDP (*84*), or compressed sensing (*30*). These methods and their comparison to MaxLik and DNNs are presented in the Supplementary Materials.

DNN methods

Neural networks are machine learning models that learn to perform tasks by analyzing data. A DNN model consists of multiple layers of interconnected artificial neurons and acts as a highly nonlinear transformation parameterized by a large number of trainable parameters (*85*). DNNs have the ability to generalize from learning stage, i.e., once trained they can perform unexpectedly well even for inputs that were not observed during the learning stage. The basic principles of DNNs operation are well known, but the full span of their generalization ability is the subject of current research (*86*, *87*). In science and technology, neural networks have been successfully applied to a wide range of problems, including predicting the behavior of complex systems and analyzing large datasets from experiments and simulations (*88*, *89*).

Let us first consider the DNN quantification of entanglement in a two-qubit system. The measurement-specific DNNs are fully connected networks. The network has seven fully connected layers with a few dozens of thousands trainable parameters in total. The exact number of the free parameters differs between the networks that have different length of the input vector, dependent on the number of projectors measured. We trained 193 measurement-specific DNNs (12 per point except of full 36 projectors) with varying length of the input layer, starting with the full 36 input neurons down to 4 (with increment of 2).

We construct a set of quantum states as follows: We generate 10^6 random quantum states ρ of which four-fifths are randomly distributed according to the Bures measure induced by the Bures metric (*90*)

$$\rho = \frac{(1+U^\dagger) G G^\dagger (1+U)}{\text{Tr}(1+U^\dagger) G G^\dagger (1+U)} \tag{4}$$

To achieve this, we generate a Ginibre matrix G with complex entries sampled from the standard normal distribution, $G_{ij} \sim \mathcal{N}(0,1) + i\mathcal{N}(0,1)$, together with a random unitary U distributed according to the Haar measure (*91*). The remaining one-fifth of the set consists of random Haar pure states mixed with white noise. The generation of the set aims at the most uniform and broadest coverage of partially mixed quantum states. The set of quantum states is randomly shuffled and split to two parts, i.e., the training and validation sets containing 800,000 and 200,000 samples, respectively. The test set has 5000 states generated according to the Eq. 4.

For the quantum states, we prepare the corresponding datasets by computing the probability distribution with elements $p_i \equiv \text{Tr}\{\rho M_i\}$ and evaluate the quantum correlation measure (concurrence or mutual information) using Eqs. 1 and 2. We trained the measurement-specific DNNs to predict the quantum correlations from the probability distribution p. The training and validation datasets have the following structures

$$\begin{aligned} \mathscr{D}^{\text{input}} &= \{\text{Tr}\{\rho M_1\}, \ldots, \text{Tr}\{\rho M_{36}\}\} \\ \mathscr{D}^{\text{output}} &= \{\mathscr{Q}(\rho)\} \end{aligned} \tag{5}$$

where the length of the input vector $\mathscr{D}^{\text{input}}$ is different for various measurement-specific neural networks, ranging from full 36 projectors down to 4. The output $\mathscr{Q}(\rho)$ stands for either the concurrence or the mutual information.

We achieve the learning of the neural networks by backpropagating the error through the use of chain rule of derivation. It minimizes the loss function defined as the mean absolute difference between the true values of the quantum correlations measure $\mathscr{Q}_{\text{true}}$ and the values $\mathscr{Q}^{\theta}_{\text{predicted}}$ predicted by the networks. The loss function thus takes a form

$$\mathscr{L} = |\mathscr{Q}_{\text{true}} - \mathscr{Q}^{\theta}_{\text{predicted}}| \tag{6}$$

and the minimum is found by minimizing the $\mathscr{L}$ over all components of a training dataset to update weights and biases $\{\theta\}$ using the Nesterov-accelerated adaptive moment estimation (NAdam) algorithm. At the step t, the NAdam procedure updates parameters

$$\theta_t \leftarrow \theta_{t-1} - \eta \frac{\overline{\mathbf{m}}_t}{\sqrt{\hat{\mathbf{n}}_t} + \varepsilon} \tag{7}$$

with

$$\begin{aligned} \mathbf{g}_t &\leftarrow \nabla_{\theta_{t-1}} \mathscr{L}(\theta_{t-1}), \\ \hat{\mathbf{g}} &\leftarrow \frac{\mathbf{g}_t}{1 - \prod_{j=1}^{t} \mu_j}, \\ \mathbf{m}_t &\leftarrow \mu \mathbf{m}_{t-1} + (1-\mu)\mathbf{g}_t, \\ \hat{\mathbf{m}}_t &\leftarrow \frac{\mathbf{m}_t}{1 - \prod_{j=1}^{t+1} \mu_j}, \\ \mathbf{n}_t &\leftarrow \nu \mathbf{n}_{t-1} + (1-\nu)\mathbf{g}_t^2, \\ \hat{\mathbf{n}}_t &\leftarrow \frac{\mathbf{n}_t}{1-\nu^t}, \\ \overline{\mathbf{m}}_t &\leftarrow (1-\mu_t)\hat{\mathbf{g}}_t + \mu_{t+1}\hat{\mathbf{m}}_t \end{aligned} \tag{8}$$

The parameter η represents the learning rate, parameter μ represents the exponential decay rate for the first moment estimates $\hat{\mathbf{m}}$, the parameter ν is the exponential decay rate for the weighted norm $\mathbf{g}_t^2$, and ε is a parameter that ensures the numerical stability of the NAdam optimization procedure. In our work, we set the numerical values of parameters $\{\eta, \mu, \nu, \varepsilon\}$ to $\{0.001, 0.9, 0.999, 10^{-7}\}$. The training takes over 2000 epochs with data further divided into 100 batches to optimize the learning time and accuracy of the

predictions on the validation dataset. All above is implemented using Keras and Tensorflow libraries for Python.

The measurement-independent DNN is a generalization to the measurement-specific DNNs, and, therefore, it consists of single network that predicts the concurrence and the mutual information from any set of projectors that we chose to work with. This functionality is accomplished by a restructuralization of the input layer that inputs not only the vector of probabilities p but also the description of the measurement $\{M_i\}$ itself

$$\mathscr{D}^{\text{input}} = \{M_1, \text{Tr}\{\rho M_1\}, \ldots, M_{36}, \text{Tr}\{\rho M_{36}\}\} \quad (9)$$

The kernel of the first convolutional layer has a stride length equal to the length of the pair $\{M_i, \text{Tr}\{\rho M_i\}\}$ to prevent the network to see cross-talk between the adjacent input pairs. Each projector M_i is vectorized using d^2 trace orthonormal basis operators $\{\Gamma_i|\Gamma_i \geq 0, \text{Tr}\{\Gamma_i\Gamma_j\} = \delta_{ij}\ \forall\ i, j\}$. For incomplete measurements containing less that 36 projectors, we set the values of missing measurement probabilities and projectors to zero.

For the three-, four-, and five-qubit systems, the structure, the loss function, and the optimization procedure of the measurement-specific DNNs are the same as for the systems of two qubits. We trained 44 measurement-specific DNNs for three qubits. The length of the input vector is different for each measurement-specific neural network, ranging from $6^3 = 216$ down to 5. We also trained two specific networks for the four- and five-qubit systems with 325 and 1944 input measurements, respectively. The training and validation datasets are divided in a ratio of 4:1. They contain 100,000 measurement probability distributions (input) and values of the mutual information (output) computed from 100,000 quantum states generated using the same process as for the two-qubit states. The number of training points in the dataset is lower than that in the two-qubit case due to memory limitations. For this reason, we adopted the incremental learning method (*92*). After the loss function on the validation dataset reaches minimum, which is not updated in the next 200 epochs, the training is stopped, and the best model is saved. Next, we generate different 100,000 data points and continue training. The test set consists of 500 states generated according to Eq. 4.

The complexity of the developed DNNs is rather low, and their scaling to higher-dimensional systems is feasible. The largest network presented (two-qubit device-independent DNN) has almost 460,000 trainable parameters. The five-qubit DNN has slightly more than 230,000 trainable parameters. Its training on 2 million data samples takes 45 hours on a single consumer-grade graphics processing unit. With larger computational resources (available today), we believe that training the networks for entanglement quantification in systems with dozens of qubits should be feasible. The conventional methods, such as MaxLik, are also computationally demanding and have to be evaluated for every new data. In contrast, our approach is computationally demanding only in the training stage. The forward evaluation (from data to entanglement) is computationally easy. Specifically, the DNN entanglement quantification is, on average, four orders of magnitude faster than the MaxLik and two orders of magnitude faster than the SDP.

Experiment

The spontaneous parametric downconversion source consists of a beta barium borate (BBO) crystal cut for type II colinear generation of two correlated orthogonally polarized photons with the central wavelength of 810 nm. The BBO crystal was pumped by a continuous laser. An entangled singlet polarization state was conditionally generated by interfering the correlated photons at a balanced beamsplitter.

To achieve the complete set of data, we performed the full quantum state tomography. This was performed by measuring all 36 projective measurements as combinations of local projections to horizontal, vertical, diagonal, anti-diagonal, right-hand, and left-hand circular polarizations. The polarization analyzer consists of a sequence of half-wave and quarter-wave plates followed by a polarizer, single-mode fiber coupling, and a single-photon detector. The detection events from the two detectors were taken in coincidence basis. To obtain the datasets where the entanglement was reduced by noise, one of the pair photons was propagated through a noisy channel. The noise was implemented by injecting a weak classical signal from an attenuated laser diode. The concurrence of the entangled state reached 0.98 for a short coincidence window and no injected noise. However, for larger coincidence windows and higher levels of injected noise, the concurrence of the detected state decreased. The experimental data for the entangled states with the concurrence of 0.985 ± 0.001 and 0.201 ± 0.002 used in this work were acquired in (*93*).

Semiconductor quantum dot source consists of a quantum dot embedded in a circular Bragg grating cavity (*94*) that enables high photon collection efficiency. The quantum dot was excited via two-photon resonant excitation of the biexciton (*95*). The excitation pulses were derived from a pulsed 80-MHz repetition rate Ti:Sapphire laser. The laser scattering was spectrally filtered, and the exciton and biexciton emission were separated ahead of single-mode fiber coupling. The polarization state of the generated entangled state was analyzed using two polarization analyzers in the process of full quantum state tomography in the same way as it was performed for the parametric downconversion source. The observable degree of entanglement was predominantly limited by the nonzero fine structure splitting.

Supplementary Materials

This PDF file includes:

Supplementary Text
Figs. S1 and S2

REFERENCES AND NOTES

1. R. Horodecki, P. Horodecki, M. Horodecki, K. Horodecki, Quantum entanglement. *Rev. Mod. Phys.* **81**, 865–942 (2009).
2. A. Einstein, B. Podolsky, N. Rosen, Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.* **47**, 777–780 (1935).
3. J. S. Bell, On the Einstein Podolsky Rosen paradox. *Phys. Phys. Fiz.* **1**, 195–200 (1964).
4. B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, R. Hanson, Loophole free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature* **526**, 682–686 (2015).
5. L. K. Shalm, E. Meyer-Scott, B. G. Christensen, P. Bierhorst, M. A. Wayne, M. J. Stevens, T. Gerrits, S. Glancy, D. R. Hamel, M. S. Allman, K. J. Coakley, S. D. Dyer, C. Hodge, A. E. Lita, V. B. Verma, C. Lambrocco, E. Tortorici, A. L. Migdall, Y. Zhang, D. R. Kumor, W. H. Farr,

F. Marsili, M. D. Shaw, J. A. Stern, C. Abellán, W. Amaya, V. Pruneri, T. Jennewein, M. W. Mitchell, P. G. Kwiat, J. C. Bienfang, R. P. Mirin, E. Knill, S. W. Nam, Strong loophole-free test of local realism. *Phys. Rev. Lett.* **115**, 250402 (2015).
6. M. Giustina, M. A. M. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J.-Å. Larsson, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, J. Beyer, T. Gerrits, A. E. Lita, L. K. Shalm, S. W. Nam, T. Scheidl, R. Ursin, B. Wittmann, Significant-loophole-free test of Bell's theorem with entangled photons. *Phys. Rev. Lett.* **115**, 250401 (2015).
7. A. Osterloh, L. Amico, G. Falci, R. Fazio, Scaling of entanglement close to a quantum phase transition. *Nature* **416**, 608–610 (2002).
8. T. J. Osborne, M. A. Nielsen, Entanglement in a simple quantum phase transition. *Phys. Rev. A* **66**, 032110 (2002).
9. L. Amico, R. Fazio, A. Osterloh, V. Vedral, Entanglement in many-body systems. *Rev. Mod. Phys.* **80**, 517–576 (2008).
10. H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu, P. Hu, X.-Y. Yang, W.-J. Zhang, H. Li, Y. Li, X. Jiang, L. Gan, G. Yang, L. You, Z. Wang, L. Li, N.-L. Liu, C.-Y. Lu, J.-W. Pan, Quantum computational advantage using photons. *Science* **370**, 1460–1463 (2020).
11. L. S. Madsen, F. Laudenbach, M. Falamarzi. Askarani, F. Rortais, T. Vincent, J. F. F. Bulmer, F. M. Miatto, L. Neuhaus, L. G. Helt, M. J. Collins, A. E. Lita, T. Gerrits, S. W. Nam, V. D. Vaidya, M. Menotti, I. Dhand, Z. Vernon, N. Quesada, J. Lavoie, Quantum computational advantage with a programmable photonic processor. *Nature* **606**, 75–81 (2022).
12. S. Wengerowsky, S. K. Joshi, F. Steinlechner, H. Hübel, R. Ursin, An entanglement-based wavelength-multiplexed quantum communication network. *Nature* **564**, 225–228 (2018).
13. J. Yin, Y. H. Li, S. K. Liao, M. Yang, Y. Cao, L. Zhang, J. G. Ren, W. Q. Cai, W. Y. Liu, S. L. Li, R. Shu, Y. M. Huang, L. Deng, L. Li, Q. Zhang, N. L. Liu, Y. A. Chen, C. Y. Lu, X. B. Wang, F. Xu, J. Y. Wang, C. Z. Peng, A. K. Ekert, J. W. Pan, Entanglement-based secure quantum cryptography over 1,120 kilometres. *Nature* **582**, 501–505 (2020).
14. G. Guccione, T. Darras, H. le Jeannic, V. B. Verma, S. W. Nam, A. Cavaillès, J. Laurat, Connecting heterogeneous quantum networks by hybrid entanglement swapping. *Sci. Adv.* **6**, eaba4508 (2020).
15. F. B. Basset, M. Valeri, E. Roccia, V. Muredda, D. Poderini, J. Neuwirth, N. Spagnolo, M. B. Rota, G. Carvacho, F. Sciarrino, R. Trotta, Quantum key distribution with entangled photons generated on demand by a quantum dot. *Sci. Adv.* **7**, eabe6379 (2021).
16. M. Erhard, M. Krenn, A. Zeilinger, Advances in high-dimensional quantum entanglement. *Nat. Rev. Phys.* **2**, 365–381 (2020).
17. J. Eisert, D. Hangleiter, N. Walk, I. Roth, D. Markham, R. Parekh, U. Chabaud, E. Kashefi, Quantum certification and benchmarking. *Nat. Rev. Phys.* **2**, 382–390 (2020).
18. P. Horodecki, Measuring quantum entanglement without prior state reconstruction. *Phys. Rev. Lett.* **910**, 167901 (2003).
19. J. Fiurášek, N. J. Cerf, How to measure squeezing and entanglement of gaussian states without homodyning. *Phys. Rev. Lett.* **93**, 063601 (2004).
20. S. P. Walborn, P. H. S. Ribeiro, L. Davidovich, F. Mintert, A. Buchleitner, Experimental determination of entanglement with a single measurement. *Nature* **440**, 1022–1024 (2006).
21. R. Islam, R. Ma, P. M. Preiss, M. Eric Tai, A. Lukin, M. Rispoli, M. Greiner, Measuring entanglement entropy in a quantum many-body system. *Nature* **528**, 77–83 (2015).
22. A. M. Kaufman, M. E. Tai, A. Lukin, M. Rispoli, R. Schittko, P. M. Preiss, M. Greiner, Quantum thermalization through entanglement in an isolated many-body system. *Science* **353**, 794–800 (2016).
23. R. Filip, Overlap and entanglement-witness measurements. *Phys. Rev. A* **65**, 062320 (2002).
24. A. K. Ekert, C. M. Alves, D. K. L. Oi, M. Horodecki, P. Horodecki, L. C. Kwek, Direct estimations of linear and nonlinear functionals of a quantum state. *Phys. Rev. Lett.* **88**, 217901 (2002).
25. K. Vogel, H. Risken, Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase. *Phys. Rev. A Gen. Phys.* **40**, 2847–2849 (1989).
26. M. Paris, J. Rehacek, Quantum State Estimation (Springer-Verlag, 2004).
27. J. T. Barreiro, N. K. Langford, N. A. Peters, P. G. Kwiat, Generation of hyperentangled photon pairs. *Phys. Rev. Lett.* **95**, 260501 (2005).
28. T. Monz, P. Schindler, J. T. Barreiro, M. Chwalla, D. Nigg, W. A. Coish, M. Harlander, W. Hänsel, M. Hennrich, R. Blatt, 14-Qubit entanglement: Creation and coherence. *Phys. Rev. Lett.* **106**, 130506 (2011).
29. X.-L. Wang, Y. H. Luo, H. L. Huang, M. C. Chen, Z. E. Su, C. Liu, C. Chen, W. Li, Y. Q. Fang, X. Jiang, J. Zhang, L. Li, N. L. Liu, C. Y. Lu, J. W. Pan, 18-Qubit entanglement with six photons' three degrees of freedom. *Phys. Rev. Lett.* **120**, 260502 (2018).
30. D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, J. Eisert, Quantum state tomography via compressed sensing. *Phys. Rev. Lett.* **105**, 150401 (2010).
31. M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, Y.-K. Liu, Efficient quantum state tomography. *Nat. Commun.* **1**, 149 (2010).
32. G. Tóth, W. Wieczorek, D. Gross, R. Krischek, C. Schwemmer, H. Weinfurter, Permutationally invariant quantum tomography. *Phys. Rev. Lett.* **105**, 250403 (2010).
33. B. P. Lanyon, C. Maier, M. Holzäpfel, T. Baumgratz, C. Hempel, P. Jurcevic, I. Dhand, A. S. Buyskikh, A. J. Daley, M. Cramer, M. B. Plenio, R. Blatt, C. F. Roos, Efficient tomography of a quantum many-body system. *Nat. Phys.* **13**, 1158–1162 (2017).
34. G. Brida, I. P. Degiovanni, A. Florio, M. Genovese, P. Giorda, A. Meda, M. G. A. Paris, A. Shurupov, Experimental estimation of entanglement at the quantum limit. *Phys. Rev. Lett.* **104**, 100501 (2010).
35. G. Brida, I. P. Degiovanni, A. Florio, M. Genovese, P. Giorda, A. Meda, M. G. A. Paris, A. P. Shurupov, Optimal estimation of entanglement in optical qubit systems. *Phys. Rev. A* **83**, 052301 (2011).
36. C. Benedetti, A. P. Shurupov, M. G. A. Paris, G. Brida, M. Genovese, Experimental estimation of quantum discord for a polarization qubit and the use of fidelity to assess quantum correlations. *Phys. Rev. A* **87**, 052136 (2013).
37. D.-L. Deng, X. Li, S. D. Sarma, Quantum entanglement in neural network states. *Phys. Rev. X* **7**, 021021 (2017).
38. G. Carleo, M. Troyer, Solving the quantum many-body problem with artificial neural networks. *Science* **355**, 602–606 (2017).
39. G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, G. Carleo, Neural-network quantum state tomography. *Nat. Phys.* **14**, 447–450 (2018).
40. G. Carleo, Y. Nomura, M. Imada, Constructing exact representations of quantum many body systems with deep neural networks. *Nat. Commun.* **9**, 5322 (2018).
41. M. J. Hartmann, G. Carleo, Neural-network approach to dissipative quantum many-body dynamics. *Phys. Rev. Lett.* **122**, 250502 (2019).
42. M. Reh, M. Schmitt, M. Gärttner, Time-dependent variational principle for open quantum systems with artificial neural networks. *Phys. Rev. Lett.* **127**, 230501 (2021).
43. A. Szabó, C. Castelnovo, Neural network wave functions and the sign problem. *Phys. Rev. Res.* **2**, 033075 (2020).
44. T. Westerhout, N. Astrakhantsev, K. S. Tikhonov, M. I. Katsnelson, A. A. Bagrov, Generalization properties of neural network approximations to frustrated magnet ground states. *Nat. Commun.* **11**, 1593 (2020).
45. G. Torlai, R. G. Melko, Latent space purification via neural density operators. *Phys. Rev. Lett.* **120**, 240503 (2018).
46. G. Torlai, B. Timar, E. P. L. van Nieuwenburg, H. Levine, A. Omran, A. Keesling, H. Bernien, M. Greiner, V. Vuletić, M. D. Lukin, R. G. Melko, M. Endres, Integrating neural networks with a quantum simulator for state reconstruction. *Phys. Rev. Lett.* **123**, 230504 (2019).
47. E. S. Tiunov, V. V. Tiunova (Vyborova), A. E. Ulanov, A. I. Lvovsky, A. K. Fedorov, Experimental quantum homodyne tomography via machine learning. *Optica* **7**, 448–454 (2020).
48. A. M. Palmieri, E. Kovlakov, F. Bianchi, D. Yudin, S. Straupe, J. D. Biamonte, S. Kulik, Experimental neural network enhanced quantum tomography. *npj Quantum Inf.* **6**, 20 (2020).
49. O. Danaci, S. Lohani, B. T. Kirby, R. T. Glasser, Machine learning pipeline for quantum state estimation with incomplete measurements. *Mach. Learn. Sci. Technol.* **2**, 035014 (2021).
50. A. Rocchetto, S. Aaronson, S. Severini, G. Carvacho, D. Poderini, I. Agresti, M. Bentivegna, F. Sciarrino, Experimental learning of quantum states. *Sci. Adv.* **5**, eaau1946 (2019).
51. V. Gebhart, R. Santagati, A. A. Gentile, E. M. Gauger, D. Craig, N. Ares, L. Banchi, F. Marquardt, L. Pezzè, C. Bonato, Learning quantum systems. *Nat. Rev. Phys. Ther.* **5**, 141–156 (2023).
52. S. Aaronson, in STOC 2018: Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing (Association for Computing Machinery, 2018), pp. 325–338.
53. H.-Y. Huang, R. Kueng, J. Preskill, Predicting many properties of a quantum system from very few measurements. *Nat. Phys.* **16**, 1050–1057 (2020).
54. G. I. Struchalin, Y. A. Zagorovskii, E. V. Kovlakov, S. S. Straupe, S. P. Kulik, Experimental estimation of quantum state properties from classical shadows. *PRX Quantum* **2**, 010307 (2021).
55. S. J. van Enk, C. W. J. Beenakker, Measuring $\mathrm{Tr}\rho^n$ on single copies of ρ using random measurements. *Phys. Rev. Lett.* **108**, 110503 (2012).
56. A. Elben, B. Vermersch, M. Dalmonte, J. I. Cirac, P. Zoller, Rényi entropies from random quenches in atomic hubbard and spin models. *Phys. Rev. Lett.* **120**, 050406 (2018).
57. T. Brydges, A. Elben, P. Jurcevic, B. Vermersch, C. Maier, B. P. Lanyon, P. Zoller, R. Blatt, C. F. Roos, Probing Rényi entanglement entropy via randomized measurements. *Science* **364**, 260–263 (2019).
58. A. Elben, S. T. Flammia, H.-Y. Huang, R. Kueng, J. Preskill, B. Vermersch, P. Zoller, The randomized measurement toolbox. *Nat. Rev. Phys.* **5**, 9–24 (2023).
59. P. Horodecki, A. Ekert, Method for direct detection of quantum entanglement. *Phys. Rev. Lett.* **89**, 127902 (2002).
60. J. Fiurášek, Structural physical approximations of unphysical maps and generalized quantum measurements. *Phys. Rev. A* **66**, 052315 (2002).

61. C. Carmeli, T. Heinosaari, A. Karlsson, J. Schultz, A. Toigo, Verifying the quantumness of bipartite correlations. *Phys. Rev. Lett.* **116**, 230403 (2016).
62. O. Gühne, P. Hyllus, D. Bruß, A. Ekert, M. Lewenstein, C. Macchiavello, A. Sanpera, Detection of entanglement with few local measurements. *Phys. Rev. A* **66**, 062305 (2002).
63. M. Barbieri, F. de Martini, G. di Nepi, P. Mataloni, G. M. D'Ariano, C. Macchiavello, Detection of entanglement with polarized photons: Experimental realization of an entanglement witness. *Phys. Rev. Lett.* **91**, 227901 (2003).
64. K. M. R. Audenaert, M. B. Plenio, When are correlations quantum?—Verification and quantification of entanglement by simple measurements. *New J. Phys.* **8**, 266–266 (2006).
65. B. Jungnitsch, T. Moroder, O. Guühne, Taming multiparticle entanglement. *Phys. Rev. Lett.* **106**, 190502 (2011).
66. M. Walter, B. Doran, D. Gross, M. Christandl, Entanglement polytopes: Multiparticle entanglement from single-particle information. *Science* **340**, 1205–1208 (2013).
67. A. Ketterer, N. Wyderka, O. Gühne, Characterizing multipartite entanglement with moments of random correlations. *Phys. Rev. Lett.* **122**, 120505 (2019).
68. L. Knips, J. Dziewior, W. Kłobus, W. Laskowski, T. Paterek, P. J. Shadbolt, H. Weinfurter, J. D. A. Meinecke, Multipartite entanglement analysis from random correlations. *npj Quantum Inf.* **6**, 51 (2020).
69. J. Gao, L. F. Qiao, Z. Q. Jiao, Y. C. Ma, C. Q. Hu, R. J. Ren, A. L. Yang, H. Tang, M. H. Yung, X. M. Jin, Experimental machine learning of quantum states. *Phys. Rev. Lett.* **120**, 240501 (2018).
70. C. Harney, S. Pirandola, A. Ferraro, M. Paternostro, Entanglement classification via neural network quantum states. *New J. Phys.* **22**, 045001 (2020).
71. J. Roik, K. Bartkiewicz, A. Černoch, K. Lemr, Accuracy of entanglement detection via artificial neural networks and human-designed entanglement witnesses. *Phys. Rev. Appl.* **15**, 054006 (2021).
72. D. Kaszlikowski, A. Sen(De), U. Sen, V. Vedral, A. Winter, Quantum correlation without classical correlations. *Phys. Rev. Lett.* **101**, 070502 (2008).
73. M. Mi˘uda, R. Stárek, J. Provazník, O. Leskovjanová, L. Mišta, Verifying genuine multipartite entanglement of the whole from its separable parts. *Optica* **6**, 896–901 (2019).
74. D. Lu, T. Xin, N. Yu, Z. Ji, J. Chen, G. Long, J. Baugh, X. Peng, B. Zeng, R. Laflamme, Tomography is necessary for universal entanglement detection with single-copy observables. *Phys. Rev. Lett.* **116**, 230501 (2016).
75. N. Yu, Multipartite entanglement certification, with or without tomography. *IEEE Trans. Inf. Theory* **66**, 6369–6377 (2020).
76. M. A. Valdez, D. Jaschke, D. L. Vargas, L. D. Carr, Quantifying complexity in quantum phase transitions via mutual information complex networks. *Phys. Rev. Lett.* **119**, 225301 (2017).
77. Z. Hradil, J. Řeháˇek, J. Fiurášek, M. Ježek, in Quantum State Estimation (Springer, 2004), pp. 59–112.
78. C. Schwemmer, L. Knips, D. Richart, H. Weinfurter, T. Moroder, M. Kleinmann, O. Gühne, Systematic errors in current quantum state tomography tools. *Phys. Rev. Lett.* **114**, 080403 (2015).
79. G. B. Silva, S. Glancy, H. M. Vasconcelos, Investigating bias in maximum-likelihood quantum-state tomography. *Phys. Rev. A* **95**, 022107 (2017).
80. A. Kalev, R. L. Kosut, I. H. Deutsch, Quantum tomography protocols with positivity are compressed sensing protocols. *npj Quantum Inf.* **1**, 15018f (2015).
81. Z. Hradil, Quantum-state estimation. *Phys. Rev. A* **55**, R1561–R1564 (1997).
82. M. Ježek, J. Fiurášek, Z. Hradil, Quantum inference of states and processes. *Phys. Rev. A* **68**, 012305 (2003).
83. Y. S. Teo, H. Zhu, B.-G. Englert, J. Řeháˇek, Z. Hradil, Quantum-state reconstruction by maximizing likelihood and entropy. *Phys. Rev. Lett.* **107**, 020404 (2011).
84. L. Vandenberghe, S. Boyd, Semidefinite programming. *SIAM Rev.* **38**, 49–95 (1996).
85. Y. LeCun, Y. Bengio, G. Hinton, Deep learning. *Nature* **521**, 436–444 (2015).
86. M. Belkin, D. Hsu, S. Ma, S. Mandal, Reconciling modern machine-learning practice and the classical bias-variance trade-off. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 15849–15854 (2019).
87. K. Kawaguchi, Y. Bengio, L. Kaelbling, in Mathematical Aspects of Deep Learning (Cambridge Univ. Press, 2022), pp. 112–148.
88. G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, L. Zdeborová, Machine learning and the physical sciences. *Rev. Mod. Phys.* **91**, 045002 (2019).
89. T. Neupert, M. H. Fischer, E. Greplová, K. Choo, M. M. Denner, Introduction to machine learning for the sciences. arxiv:2102.04883 (2021).
90. D. Bures, An extension of Kakutani's theorem on infinite product measures to the tensor product of semifinite w^* -algebras. *Trans. Am. Math. Soc.* **135**, 199 (1969).
91. F. Mezzadri, How to generate random matrices from the classical compact groups. *Not. Am. Math. Soc.* **54**, 592–604 (2007).
92. G. A. Carpenter, S. Grossberg, D. B. Rosen, Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system. *Neural Netw.* **4**, 759–771 (1991).
93. I. Straka, M. Miková, M. Mi˘uda, M. Dušek, M. Ježek, R. Filip, Conditional cooling limit for a quantum channel going through an incoherent environment. *Sci. Rep.* **5**, 16721 (2015).
94. S. Ates, L. Sapienza, M. Davanco, A. Badolato, K. Srinivasan, Bright single-photon emission from a quantum dot in a circular bragg grating microcavity. *IEEE J. Sel. Top. Quantum Electron.* **18**, 1711–1721 (2012).
95. H. Jayakumar, A. Predojević, T. Huber, T. Kauten, G. S. Solomon, G. Weihs, Deterministic photon pairs and coherent optical control of a single quantum dot. *Phys. Rev. Lett.* **110**, 135505 (2013).

Acknowledgments: We thank J. Fiurášek for discussion. We acknowledge the use of cluster computing resources provided by the Department of Optics, Palacký University Olomouc. We thank J. Provazník for maintaining the cluster and providing support. **Funding:** This work was supported by the Czech Science Foundation (grant no. 21-18545S), the Ministry of Education, Youth and Sports of the Czech Republic (grant no. 8C18002), and European Union's Horizon 2020 (2014–2020) research and innovation framework programme (project HYPER-U-P-S). D.K. acknowledges support by the Palacký University (grants IGA-PrF-2022-001 and IGA-PrF-2023-002). A.P. would like to acknowledge the Swedish Research Council (grants 2017-04467 and 2018-00848). Project HYPER-U-P-S has received funding from the QuantERA ERA-NET Cofund in Quantum Technologies implemented within the European Union's Horizon 2020 Programme. L.G. was supported by the Knut and Alice Wallenberg Foundation [through the Wallenberg Centre for Quantum Technology (WACQT)]. S.H. acknowledges financial support by the State of Bavaria. **Author contributions:** D.K. and M.J. provided the theoretical analysis. D.K. performed the numerical simulations and experimental data analysis. L.G. and A.P. performed the measurements on quantum dot source. M.M.-D., S.H., and C.S. fabricated the sample. M.J. initiated and coordinated the project. D.K., A.P., and M.J. wrote the manuscript. All authors were involved in revising the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 29 June 2022
Accepted 16 June 2023
Published 19 July 2023
10.1126/sciadv.add7131

From Stars to Molecules: AI Guided Device-Agnostic Super-Resolution Imaging

Dominik Vašinka, Filip Juráň, Jaromír Běhal, and Miroslav Ježek*
Department of Optics, Faculty of Science, Palacký University, 17. listopadu 12, 77900 Olomouc, Czechia

Super-resolution imaging has revolutionized the study of systems ranging from molecular structures to distant galaxies. However, existing super-resolution methods require extensive calibration and retraining for each imaging setup, limiting their practical deployment. We introduce a device-agnostic deep-learning framework for super-resolution imaging of point-like emitters that eliminates the need for calibration data or explicit knowledge of optical system parameters. Our model is trained on a diverse, numerically simulated dataset encompassing a broad range of imaging conditions, enabling generalization across different optical setups. Once trained, it reconstructs super-resolved images directly from a single resolution-limited camera frame with superior accuracy and computational efficiency compared to state-of-the-art methods. We experimentally validate our approach using a custom microscopy setup with ground-truth emitter positions. We also demonstrate its versatility on astronomical and single-molecule localization microscopy datasets, achieving unprecedented resolution without prior information. Our findings establish a pathway toward universal, calibration-free super-resolution imaging, expanding its applicability across scientific disciplines.

MAIN

Optical resolution is fundamentally constrained by diffraction. This limits the ability to observe structures comparable to the wavelength of light divided by the numerical aperture of the imaging system. Super-resolution imaging has emerged as a transformative technique in disciplines ranging from biomedical [1, 2] and material sciences [3] to astronomy [4, 5], by circumventing this limit and revealing previously inaccessible structural details. Numerous applications rely specifically on imaging point-like or single-emitter sources. In biology, single-molecule localization microscopy enables nanoscale visualization of cellular structures [6–11]. Quantum physics leverages super-resolution for characterizing quantum dots [12] and precisely imaging cold atoms in optical lattices [13, 14]. Astronomy benefits through resolving individual stars and galaxies [15, 16].

Despite these successes, super-resolution methods require precise calibration and extensive knowledge of optical parameters. This requirement presents a substantial barrier, often involving laborious measurements, calibration data acquisition, and computationally expensive model retraining. Furthermore, practical applications frequently struggle with inhomogeneities and instabilities in imaging conditions, which further limit the versatility of existing algorithms.

To address these fundamental challenges, we propose a novel deep-learning framework that is inherently device-agnostic, eliminating the need for calibration or prior system-specific information. Our method exploits numerically simulated data encompassing an unprecedented range of optical conditions, enabling extreme generalization and adaptability across different imaging setups. Once trained, the model delivers rapid and accurate super-resolution reconstructions directly from a single, diffraction-limited image frame.

In this work, we demonstrate the efficacy of our approach through computational evaluations and experimental validations across multiple benchmarks, outperforming statistical Bayesian and deep-learning-based state-of-the-art methods. We further confirm the practical advantages and wide applicability of our approach using diverse datasets from molecular localization microscopy to astronomical imaging. The developed framework paves the way towards universal, calibration-free super-resolution imaging, significantly enhancing its accessibility and impact across scientific and technological disciplines.

Super-resolution and device-dependence

Super-resolution imaging can be achieved through various approaches, such as linear inverse and statistical Bayesian algorithms [17], sparse representation [18], tomographic image synthesis [10, 19, 20], and methods based on blinking emitters [6–8]. Many deep learning super-resolution methods have been developed in the last decade [21–25]. Recently, artificial intelligence has been used to identify novel super-resolution microscopy setups [26].

Regardless of the specific approach, super-resolution imaging can generally be classified into two categories: reconstruction and parameter estimation. Reconstruction aims to directly restore the whole super-resolved image [17, 22], while estimation focuses on extracting key parameters and features, such as emitter localization [24, 27, 28]. The localization performance rapidly decreases for images with higher number of overlapping emitters, which typically limits its application to stochastically blinking or photoactivated samples. In the following discussion, we will focus exclusively on the former, more universal approach: image reconstruction.

Achieving accurate reconstruction of an object based on a single intensity image requires substantial knowledge

* Correspondence email address: jezek@optics.upol.cz

about the imaging system under consideration. Central to this knowledge is an accurate estimation of the point spread function, which characterizes the response of the system to a point source of light. The reconstruction quality is inherently linked to the accuracy of the point spread function estimate; any inaccuracies will lead to poor reconstruction results [29]. However, its precise estimation entails additional calibration measurements involving an isolated emitting point source. This condition can prove very demanding, especially in systems with high concentration of emitters, low signal-to-noise ratio, or temporally developing systems. Additionally, imaging systems commonly exhibit variations in the shape of the point spread function across the image. Consequently, reconstructing larger areas may require separate calibration for each segment [30].

Some approaches can be adapted to perform a blind reconstruction, operating without explicit prior knowledge of the point spread function [31]. Instead, these methods impose additional assumptions on the imaging system, typically in the form of a predefined point spread function shape. During reconstruction, parameters of the point spread function, such as its width or other defining characteristics, are iteratively estimated directly from the input image. As a result, these methods introduce imperfections by simultaneously reconstructing the observed object and the point spread function of the imaging system. Altogether, blind reconstruction remains a non-convex optimization problem, posing a challenging task in real-life scenarios and often yielding inferior results compared to non-blind algorithms.

Similar device-dependent constraints arise with reconstruction using deep learning. While these approaches often outperform classical algorithms in both reconstruction and parameter estimation tasks [21, 22, 24, 25], their calibration presents an even more significant obstacle. Unlike traditional approaches, deep learning model operates on a data-driven paradigm, inferring the reconstruction mapping by extracting relevant information from observed samples. However, due to the device-dependent nature of a typical training process, the model can only be applied to a particular point spread function profile and additional conditions such as emitter power, background noise distribution, and concentration of emitters, each requiring prior estimation from the calibration datasets [32]. Modifying these parameters requires remeasuring the calibration set and retraining the model, making it a resource-expensive procedure.

RESULTS

Device-agnostic modeling

We developed a device-agnostic modeling network (DAMN) designed to reconstruct an intensity image of point-like emitting sources, as illustrated in Fig. 1. Unlike traditional approaches, the DAMN model operates using a single intensity frame without requiring additional information about the optical parameters of the imaging system. Without the need for calibration measurements or retraining, it can directly process images from diverse applications. Our model leverages the convolutional architecture of a neural network (refer to the Methods section), enabling it to reconstruct frames of varying shapes and sizes. Moreover, the reconstruction process is performed by a single-pass propagation through the neural network, resulting in rapid reconstruction of super-resolved images.

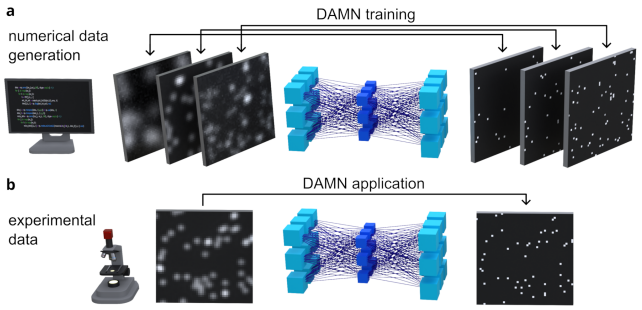


FIG. 1. Schematic representation of the DAMN approach and its application to intensity images of point-like emitting sources. **a** The model is trained using numerically simulated data pairs comprising resolution-limited noisy images alongside their super-resolved counterparts. Each training sample represents a unique combination of underlying optical parameters, such as the width of the point spread function and the signal-to-noise ratio. **b** Following training, this model is applied to enhance the resolution of experimental images acquired using a real-life imaging system.

We train the DAMN model using pairs of targeted ideal objects and their corresponding resolution-limited images generated through numerical simulation. The ensemble of data pairs covers various optical parameters, such as emitter power and concentration, background noise distribution, and even different shapes and widths of the point spread function. Their ranges were selected to cover the majority of realistic situations (additional details in the Methods section). The extent of the synthetic training data is unprecedented in the astronomy or emitter microscopy fields. Subsequently, we utilized incremental learning techniques [33] for the model to acquire knowledge about the widest possible parameter combinations. As a result of the optimization and training process, this single model accurately reconstructs super-resolved images independently of the underlying parameter values. Consequently, without the need for retraining or calibration data, the model can be applied to diverse imaging systems and is resistant to their parameter changes, including inhomogeneity across a field of view and instability, such as drifts and fluctuations in time.

Evaluation using simulated datasets

To characterize the effectiveness of the DAMN model, we conduct a comparative analysis with the Richardson-Lucy deconvolution based on Bayesian estimation with a uniform prior [34, 35]. This widely used iterative algorithm implements image reconstruction by repeatedly utilizing a known point spread function. We evaluate the results using a separate test set of image pairs not included in the training data to assess the generalization ability for new samples. We quantify the performance of both methods using the mean absolute error, calculated between the reconstructed image and its target object, averaged over the whole test dataset. In contrast to the Richardson-Lucy algorithm, which requires (and was supplied with) prior knowledge of the imaging system, our DAMN model operates independently of any device-dependent information. Despite this distinction, the device-agnostic model (without the prior information) outperforms the classical Richardson-Lucy deconvolution (with complete information on the point spread function) in terms of reconstruction quality, as discussed in the following text. Additionally, due to the iterative nature of the Richardson-Lucy algorithm, it requires three orders of magnitude higher computational time on average, further emphasize the advantages of the DAMN approach, see the Methods section.

Panels **a** - **c** in Fig. 2 illustrate the dependence of mean absolute error on the emitter power, the width of the point spread function, and the concentration of emitters, respectively. The results of these log-log graphs are characterized using an average value over the test dataset with a corresponding 90% confidence interval. Panel **a** contains dual horizontal axes representing the varying emitter power using both the signal-to-noise ratio (SNR) and the peak-to-noise ratio (PNR). For additional details regarding the explored parameters and evaluation of both approaches, see the Methods section. The DAMN model, represented by the red color, exhibits significantly better performance than the green-colored Richardson-Lucy algorithm. Similar results can also be observed in panels **b** and **c**. Panel **b** depicts the error dependence on the point spread function width (consult the Methods section). The remaining panel **c** characterizes the dependence on the emitter concentration, referring to the collective number of emitters in a single image. As seen across all three panels, our device-agnostic approach provides up to two orders of magnitude more accurate reconstructions than the device-dependent Richardson-Lucy deconvolution. The depicted errors were evaluated across the whole 50×50 pixel image and can be, alternatively, expressed as per-pixel errors by dividing the values by the number of pixels.

The panel **d** in Fig. 2 visualizes the error dependence on the point spread function shape, gradually transitioning from an Airy pattern to a Gaussian profile of the same width. As seen from this semi-logarithmic graph, the performance of both methods remains approximately

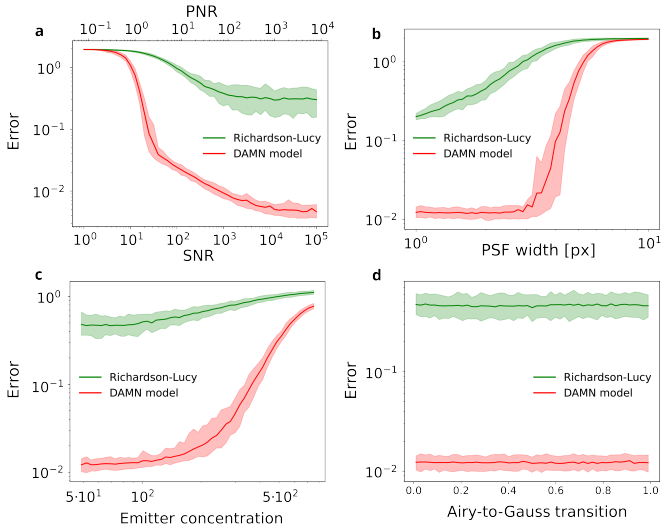


FIG. 2. The dependence of the mean absolute error on **a** the signal-to-noise ratio (SNR), **b** the width of a Gaussian point spread function (PSF), **c** the number of emitters in the image (concentration), and **d** the continuous transition between an Airy and a Gaussian PSF, respectively. The resulting averages of the Richardson-Lucy algorithm (green) and the DAMN model (red) are accompanied by their 90% confidence intervals over the test set. Panel **a** is provided with a secondary horizontal axis recalculating the SNR values to the peak-to-noise ratio (PNR). Across all panels, the DAMN model consistently outperforms the Richardson-Lucy deconvolution by up to two orders of magnitude. The optical parameters not investigated in a given panel have the following values: SNR = 500, the average noise intensity = 10, the concentration = 50, and the Gaussian PSF with $\sigma = 2$ px.

constant during this transition. Such behavior is expected from the Richardson-Lucy deconvolution, as we always provide the correct point spread function profile. On the other hand, the DAMN model was trained using solely the exact Gaussian and Airy point spread functions constituting half of the dataset each. Despite this simplification in the training process, the graph unambiguously demonstrates the adaptive ability of the DAMN approach to generalize to previously unseen point spread function shapes. This adaptability is beneficial during training, as it significantly reduces the data and simulation complexity required for training device-agnostic models. Altogether, the presented panels display numerous benefits of the DAMN approach, proving it superior to the state-of-the-art algorithm.

Optical microscopy experimental validation

Demonstrating the validity of our approach beyond the simulated data, we applied it to reconstruct experimentally acquired images. Explicitly for this, we developed a custom-built optical microscopy setup illustrated in Fig. 3 and shown in detail in Extended data Fig. 1.

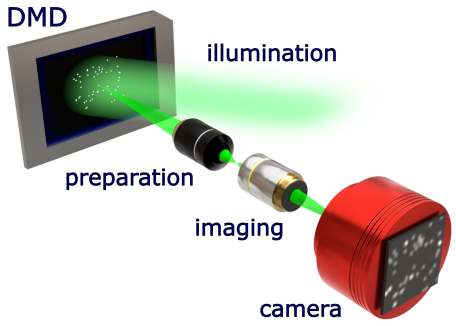


FIG. 3. Schematic illustration of the optical setup used to collect experimental data pairs. A ground-truth mask is imposed on the digital micromirror device (DMD) by configuring its mirrors. An incoherent illumination light reflected by these mirrors impinges a high-resolution preparation system. The DMD-imposed mask is re-imaged into the front sample plane of the preparation system, creating point-like emitters with the intended spatial distribution. The imaging part of the setup, comprised of a low-resolution microscope objective, images the sample-plane emitters onto a camera. The resulting camera-captured intensity image and the DMD-imposed mask represent the experimental data pairs.

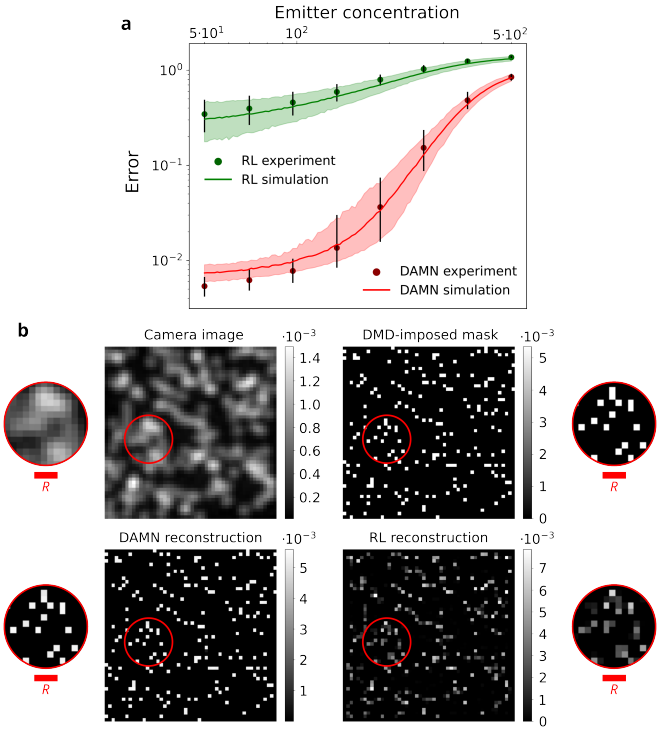


FIG. 4. **a** The red and green dots represent the mean absolute error between the DMD-imposed masks and the super-resolved images reconstructed by each method. These errors were evaluated across various emitter concentrations. The accompanying continuous lines depict error values derived from simulated data using optical parameters estimated for our imaging system. **b** A typical camera image containing nearly 200 emitters, alongside its corresponding DMD-imposed mask and each method reconstructions. The circled areas contain a magnified region for easier visual comparison. It is evident that the DAMN model significantly outperforms the Richardson-Lucy (RL) algorithm even in regions where mutual emitter distance is well-below the Rayleigh resolution limit $R = 3.9$ px (inset scalebars).

This microscope provides complete control over the spatial distribution of emitting point-like sources, achieving this for the first time in super-resolution benchmarking. An incoherent light illuminates a digital micromirror device (DMD), onto which a mask representing the targeted ground-truth image is imposed. By configuring the micromirrors, we prepare a mask containing an arbitrary number of sources positioned at desired locations. This mask is then re-imaged by a high-demagnification and high-resolution preparation system, creating point-like emitters and thus forming a sample object. Subsequently, a low-resolution imaging microscope projects the emitters onto a camera. The resulting intensity profile captured by the camera serves as the resolution-limited input for reconstruction methods. Together with the ground-truth mask, this setup provides experimental test data pairs, which allow for exact metric quantification of each reconstruction method performance. For details, see the Methods section.

Using this setup, we investigated the performance of both methods while varying the concentration of emitters. Panel **a** in Fig. 4 illustrates the mean absolute error between reconstructed images and their corresponding target objects using the experimental data, shown as dot markers. For comparison, the continuous lines represent the results obtained using simulated data with corresponding optical parameters. The DAMN model outperforms the Richardson-Lucy deconvolution by orders of magnitude, even when applied to real-world images acquired from experimental measurements. This improvement is achieved despite operating significantly below the resolution limit of our optical system. To further emphasize these findings, panel **b** in Fig. 4 shows a representative experimental image containing nearly 200 emitters and its reconstruction by each method. As observed, the DAMN model offers a near-perfect reconstruction of the original image without the need for any calibration. The Richardson-Lucy deconvolution exhibits severe deviations and artifacts despite knowing the point spread function of the imaging setup. This experimental demonstration highlights the benefits of our device-agnostic approach, which significantly outperforms the established state-of-the-art algorithm requiring prior knowledge of the point spread function.

Astronomical and localization microscopy demonstration

We have also explored the benefits of incorporating upsampling layers into the DAMN convolutional architecture, further extending its capabilities. The modified

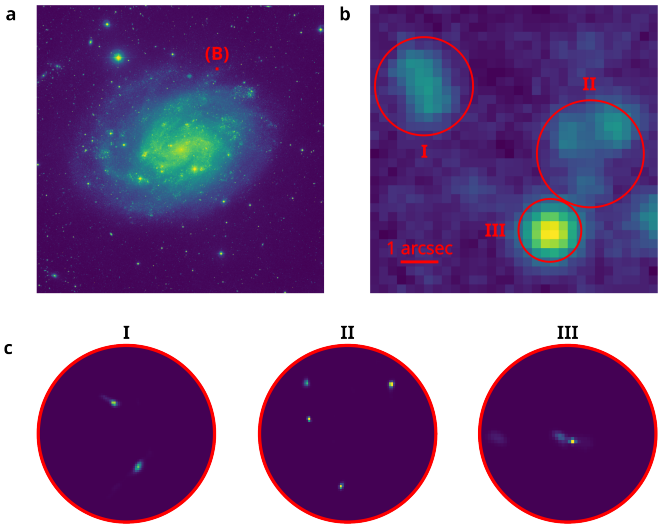


FIG. 5. Demonstration of the DAMN model super-resolving capabilities on an astronomy image of a dense star cluster. **a** Intensity image of the spiral galaxy NGC 300 acquired by European Southern Observatory. **b** Magnified region of the galaxy, highlighting key areas of interest. **c** High-resolution reconstructions of these regions generated by the DAMN model, revealing details of individual stars. These reconstructions do not use any calibration data or prior information on the employed imaging system. The region III reconstruction indicates the presence of a secondary source completely beyond the optical resolution of the telescope.

network reconstructs images with an eightfold increase in dimensions, leading to an even more prominent improvement in the resolution. We applied this model to an astronomy image of a dense star cluster from observations of the southern spiral galaxy NGC 300 in the Sculptor constellation by the European Southern Observatory [36], see Fig. 5 **a**. Panel **b** depicts a magnified segment of the galaxy with marked regions of interest, while panel **c** presents their high-resolution reconstruction produced by the DAMN model. Despite the stars in all three regions being below the resolution limit of the imaging system, the DAMN reconstruction clearly reveals them. Notably, a detailed examination of region III shows two distinct intensity peaks. While the original data resemble a single structure, our reconstruction indicates that it might be two proximal stars, with an approximate 0.11 arcsec angular distance, hidden deep below the resolution limit. Confirming this hypothesis would require further analysis using additional data. While the DAMN model can process the entire 7600×7600 pixel image, saturation artifacts in the central part of the original data (Fig. 5 **a**) led us to demonstrate the reconstruction of smaller outer regions of the galaxy. Moreover, as shown in **c**, the resolution enhancement is so significant that proper visualization requires considerable zoom into the reconstructed areas.

Additionally, we applied the same DAMN model, without retraining, to a publicly available dataset [37, 38] acquired for the single-molecule localization microscopy challenge. Fig. 6 depicts the performance evaluation using these 500 tubulin images of 128×128 pixels with high molecule concentration **a**. Panel **b** shows a reference image provided by the SOSplugin [39], a least-square localization method assuming a Gaussian point spread function, which participated in the challenge and is available as a plugin for ImageJ. Panel **c** depicts a tubulin reconstruction from Deep-STORM [22], a state-of-the-art deep learning model for image reconstruction in localization microscopy. Unlike our approach, Deep-STORM operates in a device-dependent regime, requiring extensive optical system information, such as camera specification, point-spread-function model, approximate SNR, and expected emitter density. In comparison, the DAMN model directly processes the dataset without calibration or prior knowledge of the imaging system and provides the super-resolved tubulin image, see panel **d**. The highlighted areas show a close-up comparison of all three approaches. Our DAMN model accurately captures the underlying tubulin structure, achieving resolution significantly surpassing that of the SOSplugin localization method. Unlike the localization method, the DAMN model clearly resolves the two vertically oriented, overlapping microtubules and reconstructs the horizontal one without gaps. Moreover, a close inspection reveals that our model provides a sharper image with a cleaner and thinner microtubule structure compared to the device-dependent Deep-STORM. Panel **e** visualizes the projection of the microtubule profile over the yellow segment, demonstrating the resolution improvement. This result ultimately underlines the reconstruction ability of device-agnostic models in real-world applications. Although the presented model provides super-resolved reconstructions, the same paradigm of device-agnostic deep learning can be applied to localization tasks, i.e., directly predicting the positions and intensities of the emitting sources. Such a model could be highly advantageous for single-molecule localization microscopy of densely populated samples.

DISCUSSION

We demonstrated a device-agnostic approach to super-resolution imaging of point-like emitting sources that utilizes deep learning techniques. The presented approach solves the long-standing problems of measuring calibration data and estimating parameters accompanying the training of a deep learning model. Using only numerically simulated data, we develop a model capable of reconstructing intensity images using a single frame without explicit knowledge of the imaging system. This model can be applied to images of arbitrary shape and size originating from various optical systems, and the calibration-free nature grants it resistance against non-uniformity and drifts of the optical parameters. We compared this device-agnostic model with the well-established Richardson-Lucy deconvolution al-

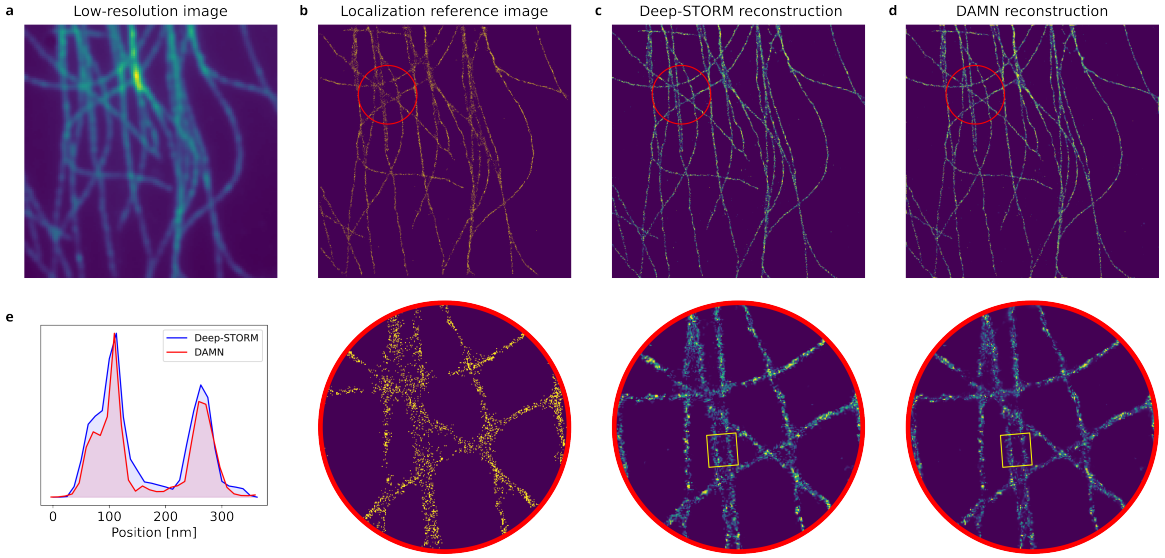


FIG. 6. Demonstration of the DAMN model on a high-concentration tubulin dataset from the single-molecule localization microscopy challenge. **a** A low-resolution image integrated from the set of 500 measured images, containing numerous emitting molecules. **b** A reference binary map of localized molecule positions derived from the dataset. **c** A reconstruction generated by Deep-STORM, a device-dependent deep learning approach calibrated using detailed information on the imaging setup. **d** A super-resolved image produced from the calibration-free, device-agnostic model without prior information. Panels **b-d** include inset images with a magnified region highlighting the resolution details. **e** Projection of the microtubule profile over the yellow rectangle segment.

gorithm. The analysis results undeniably show the superior performance of our approach, outperforming the Richardson-Lucy algorithm by orders of magnitude in terms of both reconstruction accuracy and speed. Moreover, we designed an advanced optical setup for acquiring experimental images of emitting sources together with their precise ground-truth references. Providing full control over the spatial distribution of emitters, this setup enables exact performance quantification of any reconstruction method. Using these experimental data pairs, we verified the superior performance of our approach.

To further demonstrate the universality of our approach, we applied the model to astronomy and microscopy data, achieving significant resolution improvements in both domains. We reconstructed a high-resolution tubulin image from single-molecule localization microscopy data. While in astronomy, we enhanced the resolution of a densely packed star cluster image. Notably, all DAMN reconstructions were performed without any prior knowledge of the optical setup, data preprocessing, or parameter estimation. Despite this, the obtained results surpass the state-of-the-art methods. With sufficient computation resources, even further improvements can be achieved by incorporating more complex data simulations and larger network architectures. By taking this first step, our work lays a foundation for universal image reconstruction, entirely independent of optical settings, thus contributing to the advancement of image processing in various fields.

METHODS

Data Simulation

A resolution-limited image can be characterized by its underlying optical parameters. In our case, these parameters include the shape and width of the point spread function, the background noise intensity, the emitter power representing the number of emitted photons, and the number of emitters present within a 50×50 pixels field of view of an image, which we term emitter concentration. Consequently, all possible resolution-limited images form a high-dimensional space. To implement the device-agnostic framework with a data-driven deep learning algorithm, the model must observe data samples covering the space during training. Therefore, the approach highly benefits from using simulated data pairs, as collecting a sufficient amount of experimentally acquired samples would be exceptionally time-consuming. Additionally, data simulation allows using incremental learning techniques (see Deep Learning Model), which are ideal for applications with large datasets. The following ranges of optical parameters were used for the data simulation: the emitter power $\in \left[1, 10^5\right]$, the average noise $\in [1, 100]$, the emitter concentration $\in [5, 500]$, and the point spread function width $\in \left[10^{-0.25}, 10^{1.25}\right]$ px $\approx [0.5, 17.75]$ px. For the first time in the emitter visualization community, data with such a broad scale of imaging parameters have been generated and used to train an artificial neural network.

The generation process of a single simulated data sample follows these steps. First, we generate the emitter concentration, followed by assigning each emitter its power and pixel position in the image. Next, we perform a convolution with the point spread function to simulate the effects of finite resolution. Subsequently, shot noise is added to each pixel of the image. With knowledge of the average noise intensity and emitter power, we can calculate the signal-to-noise ratio of the generated sample as $\mathrm{SNR} = \frac{\text{emitter power}}{\text{average noise intensity}}$. Alternatively, we can calculate the peak-to-noise ratio, $\mathrm{PNR} = \frac{\text{peak emitter intensity}}{\text{average noise intensity}}$, using the maximum value of the convolved emitter intensity. It is worth noting that the simulation process can be further enhanced by incorporating additional features and parameters, such as point spread function asymmetry and aberrations [40, 41]. For this showcase study, we opted to simplify the simulation to reduce the computational demand during training. Despite this simplification, the DAMN model achieves excellent results on experimental data and demonstrates significant generalization ability beyond the expected regime, as shown in Fig. 2 **d**.

To perform the convolution, we generate the convolution matrix of the point spread function using two distinct parameters - shape and width. The shape parameter distinguishes between an Airy pattern A, simulating low numerical aperture scenarios, and a two-dimensional Gaussian pattern G, typically used for cases with higher numerical aperture values [42]. The convolution matrix is normalized to a unit sum to conserve the emitter power. The pattern functions can be expressed as

$$A(r,\sigma) \sim \left[\frac{\mathrm{J}_1(\frac{2r}{\sigma})}{\frac{r}{\sigma}}\right]^2,$$
$$G(r,\sigma) \sim \exp\left(-\frac{r^2}{\sigma^2}\right),$$

where r is the distance from the center, and J_1 is the first-order Bessel function of the first kind. The variable σ is the width parameter, which dictates the full width at half maximum of the point spread function as $2\sigma\sqrt{\ln 2}$.

Deep Learning Model

The DAMN model is a deep convolutional neural network comprising 35 hidden layers with 71 channels each and a final output layer with a single channel. Information flows through each hidden layer via trainable 7×7 filters that apply local convolutional operations, allowing the network to detect localized patterns. These transformations are followed by a LeakyReLU activation function with a 0.05 negative slope, which introduces a computationally efficient non-linearity without causing the dying ReLU problem [43]. Since we normalize the input samples to a unit sum before processing, the output layer utilizes a softmax activation function [44] to preserve this condition and ensure non-negative values. To prevent overfitting, each hidden layer is paired with dropout regularization at a 0.01 rate, randomly setting a fraction of input units to zero during training [45]. This regularization encourages the network to learn robust features that do not rely on any specific neurons. As the DAMN model consists solely of convolutional layers, it can process images of arbitrary shapes and sizes without retraining. Additional resolution improvement is achieved by incorporating upsampling layers into the convolutional architecture. These layers increase the image dimension by repeating its rows and columns. The modified model applied in Fig. 5 and 6 contains three upsampling layers that double the input dimensions, resulting in their eightfold increase. This architecture makes the network highly flexible and effective for image-processing applications.

The presented model architecture results from extensive optimization utilizing a mesh adaptive direct search using the Nomad library [46], as well as numerous additional manual adjustments. Altogether, our model contains over 8 million trainable parameters and was trained incrementally on nearly 3 million simulated images, with a 25% validation set split. Incremental learning techniques gradually present the training data to the network. These techniques involve training the model using randomly generated data, which are continuously replaced with newly generated samples throughout the training process [33]. The level of network complexity and dataset size is unprecedented in the field of emitter reconstruction and localization. Nevertheless, the model remains relatively modest compared to modern large language models, suggesting further potential to integrate additional model features by increasing complexity.

The training of the DAMN model follows the backpropagation algorithm, which computes the gradient of the loss function with respect to each weight in the network using the chain rule. We use a mean squared error (MSE) loss, calculated between the predicted I_1 and target I_2 images and averaged over a batch of data samples. Given a set of two-dimensional image pairs, we calculate the loss as

$$\mathrm{MSE} = \frac{1}{J}\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{l=1}^{L}\left[I_1(j,k,l) - I_2(j,k,l)\right]^2,$$

where J is the batch size, K and L are the image dimensions, and $I(j,k,l)$ is the intensity of the j-th image in the batch at the pixel position (k,l). The gradients obtained from this loss are used to iteratively update the weights with each batch of 128 training samples. To improve convergence towards the minimum loss, we implement the Adam optimizer [47], which incorporates adaptive moment estimation. Adam updates the weights using both the first and second moments of the gradient, making the descent more efficient. A portion of the training data is reserved as a validation set to actively monitor the con-

vergence using the mean absolute error (MAE) metric

$$\mathrm{MAE} = \frac{1}{J}\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{l=1}^{L}|I_1(j,k,l) - I_2(j,k,l)|.$$

The metric provides an additional measure of model performance that is, unlike a loss function, not directly optimized during training.

Richardson-Lucy Algorithm

The Richardson-Lucy algorithm is an iterative method used to restore an image blurred by a known point spread function [34, 35]. This sophisticated and highly flexible algorithm is derived from the Bayesian probability theory with a flat prior. It maximizes the posterior probability that the observed image results from the estimated image convolved with the given point spread function. Assuming a multinomial distribution of detection events, the Richardson-Lucy is equivalent to the expectation-maximization algorithm for likelihood maximization in positive linear inverse problems [48, 49]. The quality of the reconstructed image relies heavily on accurate knowledge of the point spread function; discrepancies can lead to artifacts and inaccurate reconstructions. Since we simulate the testing data, the Richardson-Lucy algorithm can use the precise point spread function, making it an excellent method for providing a state-of-the-art baseline for evaluating the DAMN model.

In addition to the point spread function P, the iterative algorithm requires an initial guess for the reconstruction, $I^{(0)}$. A common practice is to set this guess as the observed resolution-limited image I_{blurred} or to start with a uniform image. We explored both approaches and chose to set $I^{(0)} = I_{\mathrm{blurred}}$, as there was a negligible difference in accuracy and computational speed. Then, an iterative procedure updates the estimation at the $k+1$ step

$$I^{(k+1)} = I^{(k)} \cdot \left(\frac{I_{\mathrm{blurred}}}{I^{(k)} \circledast P} \circledast P^*\right),$$

where $\circledast$ is convolution, and P^* denotes the flipped point spread function. Both the initial guess $I^{(0)}$ and the point spread function P need to be normalized to a unit sum. We implement two stopping criteria for the algorithm: either the mean absolute error between successive iterations becomes smaller than 10^{-10} per pixel, or the number of iterations exceeds 10^6. The first criterion is typically met in the majority of cases.

Furthermore, we evaluated the computation time required by the Richardson-Lucy algorithm. As the time depends on the underlying optical parameters of an image, we have simulated an extensive test dataset containing approximately 300,000 randomly generated samples. Using an AMD EPYC 7443P 24-core processor, the Richardson-Lucy needs almost five days to reconstruct this dataset. In comparison, the DAMN model utilizing the same CPU resources reconstructs the data in a little over two hours. Moreover, the DAMN model with a GPU-enabled operation requires only approximately eight minutes for the same task when using an NVIDIA RTX A5500 GPU.

Experiment

The optical setup, which provides resolution-limited experimental images with their corresponding target objects, consists of three segments, see Extended Data Fig. 1. The first segment prepares the illumination of the digital micromirror device (DMD). An incoherent broadband light with a central wavelength of 555 nm, produced by a laser-induced fluorescence light source (Crytur MonaLIGHT B01), is coupled to a 4 mm diameter light guide. The emerging light is collimated by a microscope objective $\mathrm{MO_s}$ (OLYMPUS 20x/0.4). An iris diaphragm D is positioned proximal to the objective. This diaphragm is re-imaged onto the DMD (Texas Instruments DLP LightCrafter 6500) by a telescope consisting of two achromatic doublets, $\mathrm{L_{s1}}$ and $\mathrm{L_{s2}}$, with focal lengths of 50 mm and 150 mm, respectively. The telescope improves the illumination homogeneity, and re-imaging the diaphragm reduces reflections from the passive parts of the DMD.

In the second setup segment, point-like emitters with the targeted spatial distribution are prepared. By addressing the DMD, we impose a ground-truth mask with each pixel comprising a 5×5 micromirror grid with a 7.6 μm micromirror pitch. As we address only the central micromirror of the grid, the closest distance between two emitting sources is 38 μm in the DMD plane, see Extended Data Fig. 1. This mask is re-imaged into the front focal plane of $\mathrm{MO_p}$ by a preparation telescope consisting of a 200 mm focal-length achromatic doublet $\mathrm{L_p}$ and a high-resolution microscope objective $\mathrm{MO_p}$ (OLYMPUS 100x/0.9). As the point-like emitters, of the approximate 310 nm full width at half maximum and 340 nm minimal distance, are created in this plane, we term it the emitter plane.

Unlike the first two segments, which prepare the emitters and provide complete control over their spatial distributions, the final segment represents an imaging setup with optical parameters unknown to the DAMN model. A low-resolution microscope objective $\mathrm{MO_I}$ (OLYMPUS 10x/0.25) projects the emitter plane onto a CMOS camera (ZWO-ASI 178MM, 2.4 μm square pixels) with a mounted spectral filter (central wavelength of 532 nm with 10 nm spectral width). The $\mathrm{MO_I}$ resolution is approximately 1.13 μm, regarding the full width at half maximum in the emitter plane. In comparison, the 340 nm minimal emitter distance is 3.3× below this resolution limit. The emitter plane is imaged with an approximate 230× magnification to match the major part of the camera chip. The exposure time for each measurement is

set to 5.7 seconds to utilize the majority of the 14-bit dynamic range of the camera while avoiding its saturation. The captured image is down-scaled by a factor of 32 to match the 50×50 ground-truth mask size to enable metric evaluations. Using this imaging setup, we collected the experimental dataset of resolution-limited camera images and their corresponding ground-truth masks for various emitter concentrations, see Fig. 4 **a**. Additionally, we measured a high-SNR calibration set for precise estimation of the point spread function, which we provided only to the Richardson-Lucy algorithm. Besides the concentration ranging from 50 to 500 emitters, the optical parameters in the collected dataset exhibit the following estimated values: SNR = 2300, average noise intensity = 10, and an Airy-shaped kernel with $\sigma = 2.05$ px, equivalent to the 3.4 px full width at half maximum. The Rayleigh resolution limit of our imaging system, given by the radius R of the first dark ring in the Airy pattern, is approximately 3.9 px, see Fig. 4 **b**.

DATA AVAILABILITY

The code and data supporting the results of this study are publicly available on GitHub [50].

ACKNOWLEDGMENTS

We acknowledge the massive use of cluster computing resources provided by the Department of Optics, Palacký University Olomouc. We thank J. Provazník for maintaining the cluster and providing support.

AUTHOR CONTRIBUTIONS

M.J. conceived the idea of a device-agnostic approach to image super-resolution and supervised the project. D.V. developed deep-learning models and performed numerical simulations and data processing. F.J. and J.B. developed the optical experiment, performed data acquisition, and participated in data processing. J.B. supervised the experimental part of the project. D.V. wrote the manuscript, and all authors were involved in revising the manuscript.

FUNDING

This work was supported by the Czech Science Foundation (project 21-18545S), and the Ministry of Education, Youth, and Sports of the Czech Republic (project OP JAC CZ.02.01.01/00/23_021/0008790). DV acknowledges the support by Palacký University (projects IGA-PrF-2024-008 and IGA-PrF-2025-010).

DISCLOSURES

The authors declare no conflict of interest.

[1] Sahl, S. J., Hell, S. W. & Jakobs, S. Fluorescence nanoscopy in cell biology. *Nature Reviews Molecular Cell Biology* **18**, 685–701 (2017).

[2] Schermelleh, L. *et al.* Super-resolution microscopy demystified. *Nature Cell Biology* **21**, 72–84 (2019).

[3] Pujals, S., Feiner-Gracia, N., Delcanale, P., Voets, I. & Albertazzi, L. Super-resolution microscopy as a powerful tool to study complex synthetic materials. *Nature Reviews Chemistry* **3**, 68–84 (2019).

[4] Starck, J. L., Pantin, E. & Murtagh, F. Deconvolution in astronomy: A review. *Publications of the Astronomical Society of the Pacific* **114**, 1051–1069 (2002).

[5] Bertero, M., Boccacci, P., Desiderà, G. & Vicidomini, G. Image deblurring with Poisson data: from cells to galaxies. *Inverse Problems* **25**, 123006 (2009).

[6] Rust, M. J., Bates, M. & Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods* **3**, 793–796 (2006).

[7] Betzig, E. *et al.* Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313**, 1642–1645 (2006).

[8] Balzarotti, F. *et al.* Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes. *Science* **355**, 606–612 (2017).

[9] Lelek, M. *et al.* Single-molecule localization microscopy. *Nature Reviews Methods Primers* **1**, 39 (2021).

[10] Chen, B. *et al.* Resolution doubling in light-sheet microscopy via oblique plane structured illumination. *Nature Methods* **19**, 1419–1426 (2022).

[11] Bodén, A., Ollech, D., York, A. G., Millett-Sikking, A. & Testa, I. Super-sectioning with multi-sheet reversible saturable optical fluorescence transitions (resolft) microscopy. *Nature Methods* **21**, 882–888 (2024).

[12] Lidke, K. A., Rieger, B., Jovin, T. M. & Heintzmann, R. Superresolution by localization of quantum dots using blinking statistics. *Optics Express* **13**, 7052–7062 (2005).

[13] McDonald, M., Trisnadi, J., Yao, K.-X. & Chin, C. Superresolution microscopy of cold atoms in an optical lattice. *Physical Review X* **9**, 021001 (2019).

[14] Impertro, A. *et al.* An unsupervised deep learning algorithm for single-site reconstruction in quantum gas microscopes. *Communications Physics* **6**, 166 (2023).

[15] Cava, A. *et al.* The nature of giant clumps in distant galaxies probed by the anatomy of the cosmic snake. *Nature Astronomy* **2**, 76–82 (2017).

[16] Zhang, K. *et al.* An earth-mass planet and a brown dwarf in orbit around a white dwarf. *Nature Astronomy* **8**, 1575–1582 (2024).

[17] Bertero, M., Boccacci, P. & De MoI, C. *Introduction to Inverse Problems in Imaging* (CRC Press, 2021).

[18] Yang, J., Wright, J., Huang, T. S. & Ma, Y. Image super-resolution via sparse representation. *IEEE Transactions on Image Processing* **19**, 2861–2873 (2010).

[19] Luo, W., Greenbaum, A., Zhang, Y. & Ozcan, A. Synthetic aperture-based on-chip microscopy. *Light: Science & Applications* **4**, e261–e261 (2015).

[20] Bianco, V. *et al.* Label-free intracellular multi-specificity in yeast cells by phase-contrast tomographic flow cytometry. *Small Methods* **7**, 2300447 (2023).

[21] Rivenson, Y. *et al.* Deep learning microscopy. *Optica* **4**, 1437 (2017).

[22] Nehme, E., Weiss, L. E., Michaeli, T. & Shechtman, Y. Deep-storm: super-resolution single-molecule microscopy by deep learning. *Optica* **5**, 458 (2018).

[23] Moen, E. *et al.* Deep learning for cellular image analysis. *Nature Methods* **16**, 1233–1246 (2019).

[24] Speiser, A. *et al.* Deep learning enables fast and dense single-molecule localization with high accuracy. *Nature Methods* **18**, 1082–1090 (2021).

[25] von Chamier, L. *et al.* Democratising deep learning for microscopy with zerocostdl4mic. *Nature Communications* **12** (2021).

[26] Rodríguez, C., Arlt, S., Möckl, L. & Krenn, M. Automated discovery of experimental designs in super-resolution microscopy with XLuminA. *Nature Communication* **15**, 10658 (2024).

[27] Smith, C. S., Joseph, N., Rieger, B. & Lidke, K. A. Fast, single-molecule localization that achieves theoretically minimum uncertainty. *Nature Methods* **7**, 373–375 (2010).

[28] Hekrdla, M. *et al.* Optimized molecule detection in localization microscopy with selected false positive probability. *Nature Communications* **16** (2025).

[29] Booth, M., Andrade, D., Burke, D., Patton, B. & Zurauskas, M. Aberrations and adaptive optics in super-resolution microscopy. *Microscopy* **64**, 251–261 (2015).

[30] Xu, F. *et al.* Three-dimensional nanoscopy of whole cells and tissues with in situ point spread function retrieval. *Nature Methods* **17**, 531–540 (2020).

[31] Campisi, P. & Egiazarian, K. *Blind Image Deconvolution: Theory and Applications* (CRC Press, 2007).

[32] Belthangady, C. & Royer, L. A. Applications, promises, and pitfalls of deep learning for fluorescence image reconstruction. *Nature Methods* **16**, 1215–1225 (2019).

[33] Van de Ven, G. M., Tuytelaars, T. & Tolias, A. S. Three types of incremental learning. *Nature Machine Intelligence* **4**, 1185–1197 (2022).

[34] Richardson, W. H. Bayesian-based iterative method of image restoration. *Journal of the Optical Society of America* **62**, 55 (1972).

[35] Lucy, L. B. An iterative technique for the rectification of observed distributions. *The Astronomical Journal* **79**, 745 (1974).

[36] European Southern Observatory. Wide Field Imager view of the southern spiral NGC 300. `https://www.eso.org/public/images/eso1037a/` (2010).

[37] Olivier, N., Keller, D. & Manley, S. Super-resolution microscopy hub. `https://srm.epfl.ch/srm/dataset/challenge-2D-real/Real_High_Density/index.html` (2013).

[38] Sage, D. *et al.* Quantitative evaluation of software packages for single-molecule localization microscopy. *Nature Methods* **12**, 717–724 (2015).

[39] Reuter, M. *et al.* BRCA2 diffuses as oligomeric clusters with RAD51 and changes mobility after DNA damage in live cells. *Journal of Cell Biology* **207**, 599–613 (2014).

[40] Li, J., Xue, F. & Blu, T. Fast and accurate three-dimensional point spread function computation for fluorescence microscopy. *Journal of the Optical Society of America A* **34**, 1029 (2017).

[41] Mikš, A. & Novák, J. Point spread function of an optical system with defocus and spherical aberration—analytical formulas. *Applied Optics* **58**, 5823 (2019).

[42] Stallinga, S. & Rieger, B. Accuracy of the Gaussian point spread function model in 2D localization microscopy. *Optics Express* **18**, 24461 (2010).

[43] Maas, A. L., Hannun, A. Y. & Ng, A. Y. Rectifier nonlinearities improve neural network acoustic models. *Proceedings of the 30th International Conference on Machine Learning* **28**, 6 (2013).

[44] Goodfellow, I., Bengio, Y. & Courville, A. *Deep Learning* (MIT Press, 2016).

[45] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research* **15**, 1929–1958 (2014).

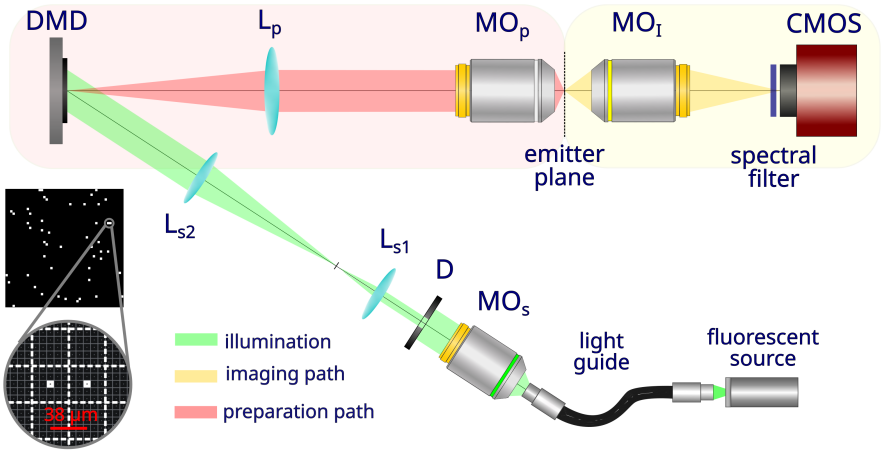
[46] Audet, C., Le Digabel, S., Rochon Montplaisir, V. & Tribes, C. Algorithm 1027: NOMAD version 4: Nonlinear optimization with the MADS algorithm. *ACM Transactions on Mathematical Software* **48**, 35:1–35:22 (2022).

[47] Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. *Proceedings of the 3rd International Conference on Learning Representations* (2015). arXiv:1412.6980.

[48] Dempster, A. P., Laird, N. M. & Rubin, D. B. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)* **39**, 1–22 (1977).

[49] Vardi, Y. & Lee, D. From image deblurring to optimal investments: Maximum likelihood solutions for positive linear inverse problems. *Journal of the Royal Statistical Society: Series B (Methodological)* **55**, 569–612 (1993).

[50] Vašinka, D. Github repository. `https://github.com/VasinkaD/DAMN` (2024).



Extended Data Fig. 1. A detailed visualization of the optical setup. In the illumination path, we generate an incoherent light collimated by a microscopic objective MO_s. The subsequent iris diaphragm D and two achromatic doublets, L_{s1} and L_{s2}, improve the illumination homogeneity and reduce undesired reflections from the passive DMD parts. In the preparation path, the DMD-induced ground-truth mask is re-imaged by an achromatic doublet L_p and a high-resolution microscope objective MO_p, creating point-like emitters in the emitter plane. Lastly, this plane is projected by a low-resolution microscope objective MO_I onto a camera with a mounted spectral filter.