

UNIVERZITA PALACKÉHO V OLOMOUCI
PŘÍRODOVĚDECKÁ FAKULTA

Katedra optiky



**Pokročilé metody detekce a estimace
kvantových a klasických stavů
optických polí**

Mgr. Dominik Koutný

Disertační práce předložena k získání titulu
Doktor Filosofie

Školitel: prof. Mgr. Jaroslav Řeháček, Ph.D.

Olomouc 2024

PALACKÝ UNIVERSITY OLOMOUC
FACULTY OF SCIENCE

Department of optics



**Advanced detection and estimation
methods of quantum and classical
states of optical fields**

Mgr. Dominik Koutný

A dissertation thesis presented for the degree of
Doctor of Philosophy

Supervisor: prof. Mgr. Jaroslav Řeháček, Ph.D.

Olomouc 2024

Abstrakt a Klíčová slova

Tato disertační práce se zabývá studiem moderních technik a teoretických výpočtů převážně z oblastí multiparametrických odhadů ze zašuměných dat, aplikaci neuronových sítí pro estimaci parametrů kvantových stavů a kvantovou tomografií. Začátek práce je věnován úvodu do studovaných problematik, po kterém následuje teoretický základ pro metody použité v hlavních publikacích autora. Dále potom práce představí hlavní výsledky základního výzkumu, kterých autor dosáhl během svých doktorských studií.

V kapitolách 4 a 5 autor prezentuje výsledky dosažené v oblasti multiparametrické estimace a estimace axiálního posunutí z intenzitní detekce. V těchto kapitolách byly odvozeny vztahy pro kvantové meze na přesnost odhadu daných parametrů.

V kapitolách 6 a 7 jsme použili moderní metody neuronových sítí pro kvantovou tomografií ze zašuměných dat a také pro odhad kvantové provázanosti z neúplného měření. Náš teoretický výzkum ukázal, že neuronové sítě dokáží i z neúplných a zašuměných dat estimovat kvantové stavy či integrální veličiny s porovnatelnou nebo lepší přesností než standardní tomografické metody. Teoretické výsledky dosažené v kapitole 7 se podařilo ověřit experimentálně.

Klíčová slova

Klasické dvou-bodové rozlišení, Rayleighovo kritérium, superrozišení, Fisherova informace, kvantová Fisherova informace, kvantová tomografie, neuronové sítě, kvantová provázanost

Abstract and Key words

This dissertation studies modern techniques and theoretical calculations, primarily in multiparametric estimation from noisy data, using neural networks to estimate quantum state parameters and quantum tomography. The beginning of the dissertation is dedicated to an introduction to the studied issues, followed by the theoretical basis for the methods used in the author's main publications. The dissertation then presents the main results of the fundamental research achieved by the author during his doctoral studies.

In chapters 4 and 5, the author presents results achieved in multiparametric and axial displacement estimations from intensity detection. In these chapters, relationships were derived for the quantum limits on the accuracy of the estimation of given parameters.

In chapters 6 and 7, we used modern neural network methods for quantum tomography from noisy data and for estimating quantum entanglement from incomplete measurements. Our theoretical research has shown that neural networks can estimate quantum states or entanglement measures from incomplete and noisy data with comparable or better accuracy than standard tomographic methods. The theoretical results achieved in chapter 7 were experimentally verified.

Key words

Classical two-point resolution, Rayleigh criterion, superresolution, Fisher information, quantum Fisher information, quantum tomography, neural networks, quantum entanglement

Acknowledgements

I would like to express my appreciation to prof. Jaroslav Řeháček for his guidance, insightful advice, and support throughout my Ph.D. studies. I am grateful to prof. Zdeněk Hradil, for his valuable feedback on my work. I also want to thank Miroslav Ježek for introducing me to the world of neural networks and for his infinite patience with me.

I am also thankful to all my collaborators and colleagues for their valuable discussions and insights. Special thanks go to Martin Paúr, Bohumil Stoklasa, Libor Mořka, Marek Vitek, and Michal Peterek, my close colleagues and friends, who helped me immensely during my studies.

Finally, my biggest thanks belong to my family for their support and love.

Thank you.

Dominik

Declaration

I hereby declare that this thesis is my own work and is based on five original publications [1–4] which summarized my most significant results reached during my Ph.D. studies. My contribution to these publications was mainly in theoretical calculations and numerical simulations and furthermore analysis of the experimental data in Chapter 7. The main goal of this work is to apply novel theoretical methods to parameter estimations and quantum state reconstruction.

I declare that I wrote this Ph.D. thesis entitled *Advanced detection and estimation methods of quantum and classical states of optical fields* on my own under the guidance of my supervisor prof. Mgr. Jaroslav Řeháček, Ph.D. This Thesis uses only the resources cited in the Bibliography section.

I agree with the further usage of this thesis according to the requirements of the Palacký University Olomouc and the Department of optics.

In Olomouc, 23rd August, 2024

.....
Dominik Koutný

List of Abbreviations

3D	Three-dimensional
CRLB	Cramér-Rao lower bound
qCRLB	quantum Cramér-Rao lower bound
MSE	Mean square error
PDF	Probability density function
POVM	Positive-operator valued measurement
PSF	Point-spread function
FWHM	Full width half maximum
FI	Fisher information
qFI	quantum Fisher information
qFIM	quantum Fisher information matrix
SLD	Symmetric logarithmic derivative
NN	Neural Network
DNN	Deep Neural Network
CNN	Convolutional Neural Network
SDP	Semidefinite program
MLE	Maximum likelihood estimator

Contents

Abstrakt a Klíčová slova	i
Abstract and Key words	ii
Acknowledgements	iii
Declaration	iv
List of Abbreviations	v
Contents	vii
1 Goal of the Thesis	1
2 Contemporary state of research	3
3 Methods and Tools	12
3.1 Rayleigh's criterion	12
3.1.1 Rectangular aperture	12
3.2 Parameter estimation in signal processing	14
3.2.1 Fisher information and Cramér-Rao bound	15
Quantum estimation theory	15
3.2.2 Quantum Fisher information and quantum Cramér-Rao bound	16
3.3 Quantum tomography	17
3.4 Deep neural networks	18
3.4.1 Fully connected neural networks	19
3.4.1.1 The layout of a neural network	19
3.4.1.2 Training of a neural network	21
3.4.1.3 Backpropagation	22
4 Optimal measurement for quantum spatial superresolution	24
5 Axial superlocalization with vortex beams	32
6 Neural network quantum state tomography	43
7 Deep learning of quantum entanglement	55

CONTENTS

8 Conclusions	71
Stručné shrnutí v češtině	73
Publications of the author and a list of citations	75
Bibliography	82

Chapter 1

Goal of the Thesis

This dissertation thesis aims to present my theoretical research, which I have done during my Ph.D. studies. The first part of the thesis seeks theoretical calculations concerning optimal quantum measurements and calculating quantum and classical limits on the precision with which we can estimate parameters from noisy data. The second part presents the advantages of using deep neural networks for quantum tomography and estimating quantum correlations from incomplete measurements.

This chapter aims to give you a short overview of the theoretical work I have done with my colleagues, based on four publications [1–4]. Chapter 2 briefly describes a contemporary state of the research, mainly dedicated to progress in superresolution and neural networks. Chapter 3 introduces the main theoretical tools and methods used in my work. Chapters 4, 5, 6, and 7 introduce achieved theoretical results. A brief summary of the main results is at the beginning of each chapter, followed by a published paper. Finally, Chapter 8 reviews all results this thesis contains. The list of my publications, citation index, and bibliography sections are appended at the end of this thesis.

In Chapter 4, we investigate the ultimate limits on the precision of simultaneous estimation of centroid, separation, and relative intensities of two incoherent point sources. By employing quantum estimation theory and the quantum Fisher information matrix, we construct optimal measurements that achieve the precision predicted by quantum theory. Our results indicate that the optimal sub-Rayleigh resolution limit can be achieved for any real-valued amplitude point spread function provided the system output is projected onto a suitable complete set of modes. This chapter includes publication [1] J. Řeháček, Z. Hradil, D. Koutný, J. Grover, A. Krzic and L. L. Sánchez-Soto. ‘Optimal measurements for quantum spatial superresolution’. In: *Physical Review A* 98.1 (2018), p. 012103.

Chapter 5 investigates the ultimate precision in axial localization using vortex beams, focusing on Laguerre-Gauss (LG) beams. We demonstrated that the quantum limit for axial localization can be achieved with an intensity scan, but only if the detector is optimally placed. We showed that traditional intensity sensors are inadequate for estimating the axial distance from superpositions of LG beams, particularly those with rotating intensity profiles. We proposed replacing these sensors with advanced mode-sorting techniques to fully exploit the potential of vortex beams in microscopy, thereby significantly enhancing axial resolution and making 3D superresolution imaging more feasible. This chapter includes publication [2] [D. Koutný, Z. Hradil, J. Řeháček and L. L. Sánchez-Soto. ‘Axial superlocalization with vortex beams’. In: *Quantum Science and Technology* 6.2 \(2021\), p. 025021.](#)

In chapter 6, we explored the application of neural networks to quantum state tomography, demonstrating that the positivity constraint can be effectively implemented with trained networks, enabling state-of-the-art deep learning methods for quantum state reconstruction under various types of noise. The neural network-based approach provides significant speed advantages over traditional methods like semidefinite programming and maximum likelihood estimation. Our work highlights the potential of deep learning techniques in enhancing the efficiency and accuracy of quantum state tomography. This chapter includes publication [3] [D. Koutný, L. Motka, Z. Hradil, J. Řeháček and L. L. Sánchez-Soto. ‘Neural-network quantum state tomography’. In: *Physical Review A* 106.1 \(2022\), p. 012409.](#)

Chapter 7 presents a novel approach for quantifying quantum entanglement using neural networks without requiring full knowledge of the quantum state. Our method significantly reduces the error of entanglement quantification compared to traditional quantum tomography techniques by training networks on simulated data. The networks can independently process data from various measurement scenarios, making the approach versatile and robust. We demonstrated the effectiveness of our neural network-based methods using experimental data from nonlinear parametric processes and semiconductor quantum dots, showcasing its practical applicability in real-world quantum systems. This chapter includes publication [4] [D. Koutný, L. Ginés, M. Moczala-Dusanowska, S. Höfling, Ch. Schneider, A. Predojević and M. Ježek. ‘Deep learning of quantum entanglement from incomplete measurements’. In: *Science Advances* 9.29 \(2023\).](#)

Chapter 2

Contemporary state of research

Optical resolution

Optical resolution is an important feature not only in our lives but also in physical sciences. One can easily understand the concept of optical resolution with everyday experiences with optical imaging through the human eye. This is the most powerful tool for humans and animals to observe the surroundings to make appropriate decisions based on these observations. People spent centuries of theoretical and experimental advancements to improve the resolving power [5] of what we can see to understand our surroundings better, from visualizing small particles to extraterrestrial planets. In other words, the optical resolution limits the level of detail we can see. We cannot distinguish between bright atoms that act like point sources of light on surfaces surrounding us since the gap between them is far below the resolution limit of the human eye. As an optical example, let us consider a single-point source and a diffraction-limited optical imaging system described by a rectangular aperture, which we will discuss in more detail in the following chapter. The impulse response is described by a sinc function, resulting in the main central lobe and an infinity of side lobes with decreasing intensity. The concept of single-point resolution translates to the determination of the position of the point source. This scenario can be applied to estimating arbitrary parameters from noisy data, forming a baseline example for the two-point resolution criteria in the following text.

Classical two-point resolution criteria

Classical two-point resolution criteria in optics are fundamental principles that determine the minimum distance at which two closely spaced points can be distinguished as separate entities. The broad summary can be found in den Dekker and van den Bros [6] and Ramsay et al. [7]. The most well-known criterion is the Rayleigh criterion [8, 9], established by Lord

Rayleigh in the late 19th century. According to the Rayleigh criterion, two-point sources are resolvable when the principal maximum of one diffraction pattern coincides with the first minimum of the other diffraction pattern. This criterion is mathematically expressed as $r = 1.22 \frac{\lambda}{NA}$, where r is the minimum resolvable distance, λ is the wavelength of light, and NA is the numerical aperture of the imaging system.

Another important criterion is the Abbe limit [10], named after Ernst Abbe, which relates the resolution to the wavelength of light and the numerical aperture of the optical system. The Abbe limit is given by $d = \frac{\lambda}{2NA}$ where d is the minimum resolvable distance, λ is the wavelength, and again NA is the numerical aperture. These criteria emphasize that higher resolution can be achieved with shorter wavelengths and larger numerical apertures.

In addition to the Rayleigh and Abbe criteria, several other resolution criteria are used to define the resolving power of optical systems. One notable criterion is the Sparrow criterion [11], which states that two-point sources are resolvable when the combined intensity distribution has a dip at the midpoint that is just perceptible. Mathematically, this criterion implies that the second derivative of the intensity distribution at the midpoint between the two sources is zero. The Sparrow criterion is particularly useful in cases where the sources are very close to each other, providing a more stringent resolution limit than the Rayleigh criterion.

Another important criterion is the Houston criterion [12], which defines resolution based on the image's contrast. According to this criterion, two points are resolved if the contrast at the midpoint of the combined intensity distribution is reduced to a specific fraction of the peak intensity, namely to the full width at half maximum (FWHM) of either point source. The Houston criterion is often used in systems where contrast plays a crucial role, such as in imaging with limited signal-to-noise ratios.

The Dawes criterion [13] is primarily used in astronomy. An empirical formula gives the minimum angular separation at which two stars of equal brightness can be resolved using a telescope. The Dawes limit is given by $\theta = \frac{116}{D}$, where θ is the angular separation in arcseconds and D is the diameter of the telescope's aperture in millimeters. This criterion is particularly useful for astronomers because it provides a practical guideline for the resolving power of telescopes based on real-world observations.

These classical criteria have significantly guided the design and improvement of optical instruments such as microscopes and telescopes. They highlight the intrinsic limitations of diffraction, which causes light from a point source to spread out and form an Airy disk pattern in the image plane. The central bright spot of this pattern, surrounded by concentric rings of decreasing intensity, represents the fundamental limit to the precision with which the position of a point source can be determined. Understanding and applying these diverse criteria can lead to more precise and effective imaging solutions across various scientific

and technical fields.

Moreover, advancements in parameter estimation theory, such as Fisher information and the Cramér-Rao lower bound, have provided deeper insights into the fundamental limits of optical resolution. These methods allow researchers to quantify the maximum achievable precision for estimating the positions of closely spaced point sources, revealing that the Rayleigh limit is not an insurmountable barrier but a consequence of suboptimal measurement strategies. By optimizing the imaging process and employing advanced statistical techniques, it is possible to enhance the resolution beyond the classical diffraction limits, paving the way for new discoveries in fields ranging from biology to nanotechnology.

Parameter estimation aided by quantum theory

As discussed in the previous paragraph, a mathematical approach based on estimation theory is needed to precisely retrieve parameter values from noisy data. The parameter estimation theory provides us with tools to quantify the precision with which we can infer, for example, the separation between two sources. Modern methods introduce the concept of *superresolution*, which states that with a given measurement, we can surpass the classical limits on resolution and thus provide more accurate estimators of parameters of interest.

The choice of estimator will determine the precision of estimated values. For example, one of the most common unbiased estimators, the maximum likelihood estimator, can be used, given that we know the probability density function for each detection. Every unbiased estimator's variance is bounded from below by the Cramér-Rao lower bound (CRLB), which is a reciprocal value of the well-known Fisher information [14, 15]. The greater the value of the Fisher information, the greater the dependence of the measured statistics is on the estimated parameter. This allows us to distinguish between different measurement schemes to determine which allows us to determine parameter values precisely. For two separated but equally bright sources, the Fisher information about small separation goes quadratically to zero, and the CRLB grows beyond all limits if we work with the standard intensity detection-based schemes. This behavior is called the *Rayleigh's curse*.

However, Tsang et al. [16] showed that if we treat the problem of separation estimation from the viewpoint of quantum theory, the *Rayleigh's curse* is dispelled. Authors derived the fundamental limit based on the quantum Fisher information, first introduced by Helström [17], whose reciprocal value consequently forms a quantum CRLB (QCRLB). The classical Fisher information is bounded from above by quantum Fisher information, which is optimized among all possible measurement schemes. The quantum Fisher information for two equally bright sources is constant for arbitrary small separations. The authors also proposed a measurement scheme based on spatial-mode demultiplexing that attains the

quantum bound on precision on separation [18, 19]. The first experimental realization of the SPADE measurement scheme was performed by Paúr et al. [20], where the authors considered two displaced Gaussian beams and used a digital holography setup to show that the Rayleigh criterion is violated.

The task of simultaneously estimating multiple parameters is vastly more challenging. Realistically, considering two sources of light, we would have to determine not only the separation, but also the positions of centroids as well as the ratio of intensities of the sources. Generally, multiparametric estimation brings tradeoffs between precision with which we estimate different parameters. Řeháček et al. [1] showed that if the intensities of the two sources are even very slightly unbalanced, the *Rayleigh's course* once again is retrieved. The information about separation goes to zero even in the quantum regime. Moreover, the authors also calculated the quantum Fisher information matrix (qFIM), the central quantity for multiparametric estimation strategies, for separation centroid, and the ratio of intensities, which shows the interconnection between precisions of estimation of different parameters.

Axial superresolution

Axial distance is a fundamental parameter in optical imaging systems, particularly in three-dimensional microscopy, defining the ability to distinguish between two points along the optical axis. Traditional limits on axial resolution are governed by the diffraction limit, as described by the Abbe-Rayleigh criterion [9, 10], which sets a theoretical resolution limit based on the wavelength of light and the numerical aperture of the system. However, advancements in superresolution microscopy techniques have significantly extended these boundaries. Methods such as stimulated-emission-depletion microscopy [21], photoactivated localization microscopy [22], and point spread function engineering [23] have achieved resolutions that surpass conventional diffraction limits by orders of magnitude. These techniques are particularly valuable in high-precision imaging fields, such as cellular biology, where understanding the fine structure of cells and tissues at the nanoscale is critical [24, 25].

In the work [26] by Řeháček et al., the authors analyzed the estimation of axial displacement from the point of view of quantum Fisher information. They showed that the quantum limit on precision can be attained using simple direct intensity detection if placed in one of the two optimal transversal detection planes. Our work, discussed in chapter 5, further advances the analysis of vortex beams [2].

Neural networks

Before going through the applications of various types of neural networks in the quantum world, let us first talk about their history. The history of neural networks dates back to the 1940s with the pioneering work [27] of Warren McCulloch and Walter Pitts, who proposed the first mathematical model of a neural network in 1943. Their model laid the foundation for understanding how neurons could be used to perform logical operations. In 1958, Frank Rosenblatt developed the perception [28], an early type of neural network designed for image recognition tasks, which marked a significant milestone in the field. However, the limitations of the perceptron, particularly its inability to solve non-linear problems, as highlighted by Marvin Minsky and Seymour Papert in their 1969 book "Perceptrons," [29] led to a temporary decline in research interest. The field saw a resurgence in the late 1970s with the introduction of backpropagation [30] by Paul Werbos in 1975, which allowed neural networks to adjust their weights more effectively through multiple layers, setting the stage for the future development of deep learning. This period up to 1980 established the foundational concepts and challenges that would shape the evolution of neural networks in the subsequent decades.

The period from 1980 to 2000 witnessed significant advancements in the field of neural networks, driven largely by the re-discovery and popularization of backpropagation. In 1986, David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams published a seminal paper [31] that demonstrated the effectiveness of backpropagation in training multi-layer neural networks, leading to a resurgence of interest in the field. Researchers developed various network architectures during the 1980s and 1990s, such as convolutional neural networks by Yann LeCun and his colleagues in the late 1980s [32], which proved highly effective for image recognition tasks. The introduction of recurrent neural networks by John Hopfield [33] and their subsequent improvements by Jürgen Schmidhuber and Sepp Hochreiter [34], notably the development of Long Short-Term Memory networks in 1997, advanced sequential data processing. These innovations, coupled with increasing computational power and the availability of larger datasets, set the stage for the neural network boom in the 21st century, establishing foundational techniques and architectures still in use today.

Neural networks in quantum physics

In recent years, deep neural networks and convolutional neural networks have significantly contributed to quantum physics, addressing complex challenges in studying and manipulating quantum systems. One of the notable advancements was the application of DNNs

to quantum many-body physics, where researchers have used these networks to recognize and classify different phases of matter. For instance, Carleo and Troyer [35] proposed the neural network quantum state approach, utilizing a restricted Boltzmann machine to approximate the wavefunction of quantum systems, thereby enabling efficient simulation of complex quantum states. Additionally, CNNs have been effectively used in quantum state tomography, reconstructing quantum states from measurement data. Torlai and Melko [36] demonstrated the use of neural networks to perform quantum state reconstruction, significantly improving the accuracy and efficiency of the process. Moreover, CNNs have been employed in quantum error correction, where they help identify and correct errors in quantum computations [37], as shown by Krastanov and Jiang. These advancements highlight the transformative impact of DNNs and CNNs in quantum physics, offering powerful tools for analyzing and interpreting high-dimensional quantum data.

Neural networks for quantum tomography

The application of Boltzmann machines and other neural network architectures in quantum tomography has seen substantial progress over the past decade. Quantum tomography, which involves reconstructing quantum states from measurement data, has greatly benefited from these advancements. Boltzmann machines, particularly restricted Boltzmann machines, have been employed to represent quantum states and facilitate their reconstruction efficiently. For example, Torlai et al. [38] demonstrated the use of restricted Boltzmann machines in quantum state tomography, where the network learned to represent quantum states from limited measurement data, providing accurate and scalable solutions to the reconstruction problem. Moreover, advanced deep learning techniques such as generative adversarial networks and variational autoencoders have shown promise in modeling and reconstructing quantum states. Gao and Duan [39] proposed a method using generative adversarial networks for quantum state tomography, showcasing the network's ability to learn complex quantum distributions. Additionally, Rocchetto [40] presented a variational autoencoder-based approach for quantum state reconstruction, highlighting its effectiveness in dealing with high-dimensional quantum data. Neural network-based methods, including feedforward neural networks and recurrent neural networks, have also been explored for their potential to enhance the precision and efficiency of quantum state tomography. Carrasquilla et al. [41] utilized a combination of feedforward neural networks and RNNs to improve the accuracy of quantum state reconstruction. These advancements underscore the powerful role of neural networks in advancing the field of quantum tomography, enabling more precise and efficient reconstruction of quantum states.

Quantum entanglement

Quantum entanglement, a fundamental concept in quantum mechanics, was first introduced by Albert Einstein, Boris Podolsky, and Nathan Rosen in their famous 1935 paper [42], aiming to demonstrate what they considered the incompleteness of quantum mechanics by describing a scenario where two particles could instantaneously affect each other regardless of distance. This phenomenon, which Einstein famously referred to as "spooky action at a distance," was further explored by Erwin Schrödinger [43], who coined the term "entanglement" and emphasized its significance as a unique feature of quantum mechanics. The theoretical groundwork laid by these pioneers remained largely untested until John Bell formulated Bell's Theorem in 1964 [44], providing a concrete way to test the EPR paradox experimentally and showing that no local hidden variable theories could replicate the predictions of quantum mechanics.

The first significant experimental tests of Bell's inequalities were conducted by John Clauser and Alain Aspect in the 1970s and 1980s. Clauser and his colleagues developed the CHSH inequality [45], a practical version of Bell's inequality, which they used to test and confirm quantum entanglement. Aspect's experiments [46] in 1982 provided even more convincing evidence by demonstrating the violation of Bell's inequalities under strict conditions, thus reinforcing the reality of entanglement. In the 1990s, quantum entanglement became a cornerstone of quantum information science, enabling groundbreaking developments in quantum cryptography and quantum computing. Charles Bennett and Gilles Brassard introduced the first quantum key distribution protocol [47], which utilized entanglement for secure communication. Meanwhile, Peter Shor's algorithm [48] demonstrated the computational advantages of quantum entanglement, showing how it could solve problems intractable for classical computers. Advances continued into the 21st century, with researchers like Anton Zeilinger achieving quantum teleportation [49], a process that relies on entanglement to transfer quantum information between particles over a distance. Today, the study of entanglement continues to be a vibrant and critical area of research, exploring its potential in various applications such as quantum networks and quantum metrology [50, 51].

Entanglement witnessing

Witnessing entanglement involves using specific criteria or measurements to verify the presence of entanglement in a quantum system. One of the most fundamental methods is the violation of Bell's inequalities, which provides a clear signature of entanglement by demonstrating correlations that classical physics cannot explain. Clauser, Horne, Shimony,

and Holt [45] formulated a variant known as the CHSH inequality, which has been widely used in experiments to test for entanglement. Another important tool is entanglement witnesses, specific observables designed to detect entanglement. An observable W is an entanglement witness if $\text{Tr}(W\rho) < 0$ for some entangled state ρ , while $\text{Tr}(W\sigma) \geq 0$ for all separable states σ . This method was notably advanced by Terhal [52], who provided a framework for constructing and utilizing such witnesses. These techniques have been crucial in experimental quantum physics, allowing researchers to identify and quantify entanglement in various systems, from photonic qubits to trapped ions. Recently, neural networks have also been used for witnessing quantum entanglement [53–55]

Direct detection of entanglement

Direct detection of entanglement, including multicopy measurements, provides more nuanced and quantitative approaches to assessing entanglement. Multicopy measurements involve using multiple copies of the same entangled state to estimate entanglement measures directly. This method allows for the computation of entanglement entropy and other measures without a complete state tomography. One prominent approach is using the SWAP test [56], which compares multiple copies of a quantum state to estimate the overlap and infer entanglement properties. Additionally, techniques such as quantum state tomography have been refined to utilize multicopy measurements effectively, as demonstrated by Audenaert et al. [57], who showed how these measurements could be used to estimate the logarithmic negativity, a quantifiable measure of entanglement, based on the partial transposition criterion [58, 59]. More recently, methods involving machine learning and neural networks have been developed to analyze multicopy measurements [38], further enhancing the precision and efficiency of entanglement detection. These advancements highlight the importance of direct detection techniques in providing deeper insights into the nature and extent of entanglement in complex quantum systems.

Entanglement measures

Entanglement measures are essential for quantifying the degree of entanglement in quantum systems, and various methods have been developed to address this challenge. One prominent measure is concurrence [60], introduced by Wootters in 1998. Concurrence is used to quantify entanglement in two-qubit systems by analyzing the density matrix of the system and comparing specific derived values. This measure helps determine how intertwined the quantum states are, providing a clear metric for entanglement. For multi-qubit systems, the measurement of entanglement becomes more complex. One commonly used approach

is the multipartite entanglement measure, which includes methods such as the geometric measure of entanglement and negativity. Negativity, introduced by Vidal and Werner in 2002 [61], is a practical way to quantify entanglement by evaluating the properties of the system's density matrix after a certain mathematical transformation. Additionally, entanglement entropy, particularly the von Neumann entropy, is used to measure entanglement across different parts of a multi-qubit system [62], allowing researchers to analyze how entanglement is distributed within the system. These measures are crucial for advancing quantum information science, providing the necessary tools to assess and utilize entanglement in increasingly complex quantum systems [56, 63–66].

Chapter 3

Methods and Tools

This chapter encompasses an introduction to the theory of parameter estimation, followed by basic concepts of topics related to quantum tomography associated with the main results of this thesis. We will discuss the theory and applications of classical and quantum Fisher information, a standard mathematical tool for adequately dealing with parameter estimation from noisy data. Finally, we will review the basic concepts behind deep learning as deep neural networks provide a different approach to parameter and quantum state estimation, a subject of the last two chapters.

3.1 Rayleigh's criterion

In 1903, lord Rayleigh published his famous discussion concerning the resolution of two incoherent sources. He stated that if the light sources are separated at least by a diffraction-limited spot size on the image plane, one can resolve them. This result found its applications in many fields of applied optics, mainly in microscopy and astronomy. However, Rayleigh's criterion is rather heuristical. In the early nineteenth century, the framework for adequately describing the errors in the estimation process from the noisy data, called Fisher information, was established. This technique allowed us to surpass Rayleigh's criterion in distinguishing the separation of two light sources. For clarity, we will briefly review the simple diffraction pattern that stems from illuminating the rectangular aperture. It is an excellent historical detour to how scientists initially approached two-point resolution problems.

3.1.1 Rectangular aperture

Let us consider an intensity distribution caused by uniformly irradiating a 1-D aperture with a half-width W . The transmittance function of the aperture is defined as

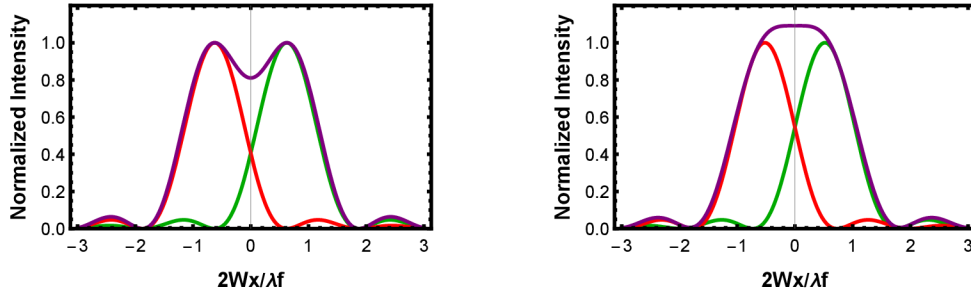


Figure 3.1: Sketch of intensity profiles of two displaced incoherent sources propagated through rectangular aperture. In both panels, the red and the green curves represent the intensity patterns of individual sources, whereas the purple curve represents the sum of the intensities of those two sources. In the left panel, the sources are separated by a Rayleigh distance (3.4), whereas the right panel shows intensity profiles and their sum of sources that are no longer resolvable according to Rayleigh's criterion.

$$t_A(x) = \text{rect}\left(\frac{x}{2W}\right), \quad (3.1)$$

where $\text{rect}(x) = 1$ for $|x| < 1/2$ and $\text{rect}(x) = 0$ for $|x| \geq 1/2$. In this simple example, the amplitude in the image plane can be computed using the Fraunhofer diffraction integral and takes the following form

$$U(x) = e^{ikf} e^{i\frac{kx^2}{2f}} \frac{2W}{i\lambda f} \text{sinc}\left(\frac{2kWx}{f}\right) \quad (3.2)$$

with the intensity distribution

$$I(x) = U^*(x)U(x) = \left(\frac{2W}{\lambda f}\right)^2 \text{sinc}^2\left(\frac{2kWx}{f}\right). \quad (3.3)$$

The width of the central lobe of the diffraction pattern is defined as

$$d = \frac{\lambda f}{W}. \quad (3.4)$$

This defines Rayleigh's criterion on the minimum resolvable distance between two incoherent sources. If the two sources are separated by a distance less than Rayleigh's diffraction limit, they become unresolvable. However, this criterion is rather heuristic, and in the next chapters, we will see how we can surpass Rayleigh's limit and distinguish between two sources by applying proper statistical methods.

3.2 Parameter estimation in signal processing

Estimation theory, a branch of mathematical statistics, aims to determine the values of parameters based on measured empirical data, which often contains noise. Suppose that we want to estimate a set of parameters $\theta = \{\theta_1, \theta_2, \dots, \theta_M\}$ given the data \mathcal{D} . We start by constructing a function $f(\mathcal{D})$, known as an estimator, which takes the measured data \mathcal{D} as input and produces the estimated parameters $\hat{\theta}$ as output. However, not all estimators have the same properties. For instance, some estimators may be biased, or the variance of the inferred values may vary. In the following discussion, we will explore the theory of Fisher information and its quantum counterpart. These concepts help us identify the most accurate unbiased estimator.

Classical estimation theory

Suppose that we collect N data points from the experiment $x = \{x[0], x[1], \dots, x[N-1]\}$, which are dependent on the true values of parameters θ . It is convenient to write estimated quantities via the formulae

$$\hat{\theta} = f(x[0], x[1], \dots, x[N-1]), \quad (3.5)$$

with the function f is called an estimator. Moreover, the elements of random vector x are possessed by an underlying probability density function (PDF) $p(x|\theta)$, which can be deduced from the physics that underlies the experiment. The conditional probability $p(x|\theta)$ tells us the probability of measuring x given the fixed set of parameters θ . Having the PDF at hand, we can define the unbiasedness of an estimator, the difference between the expected value of the estimator and the true value, as

$$E[\hat{\theta}]_{p(x|\theta)} = \theta, \quad (3.6)$$

where $E[\cdot]_{p(x|\theta)}$ denotes the expected value over the distribution $p(x|\theta)$. Unbiased estimators play a crucial role in the theory of Fisher information, which gives us a mathematical apparatus for how to differ between estimators based on the variance of $\hat{\theta}$. For example, the well-known maximum likelihood estimator is unbiased as the number of data points N approaches infinity.

3.2.1 Fisher information and Cramér-Rao bound

Naturally, as the measured data x is burdened by randomness, estimates $\hat{\theta}$ will have a spread $\Delta\hat{\theta} = \hat{\theta} - \theta$. The standard measure of deviation is the mean squared error (MSE), defined as

$$\begin{aligned}\text{MSE}(\hat{\theta}) &= E[(\Delta\hat{\theta})^2] \\ &= \text{Var}(\hat{\theta}) + b^2(\hat{\theta}),\end{aligned}\tag{3.7}$$

where $\text{Var}(\hat{\theta}) = E[\hat{\theta}^2] - E[\hat{\theta}]^2$ is the variance of the estimator and $b = E[\hat{\theta}] - \theta$ is a bias. As we can see, finding an estimator that minimizes MSE is not convenient, as it depends on the bias. However, if we restrict ourselves to unbiased estimators, the task now translates to minimizing the variance, and one can show [67] that

$$\text{Var}(\hat{\theta}) \geq \frac{1}{F(\theta)},\tag{3.8}$$

also known as the Cramér-Rao lower bound for one detection event. The denominator of the righthand side of eq. (3.8) is the Fisher information, defined as

$$F(\theta) = E\left[\left(\frac{\partial \ln p(x|\theta)}{\partial \theta}\right)^2\right],\tag{3.9}$$

and it serves as a method to determine the total amount of information measured data contains about parameter θ .

Quantum estimation theory

The mathematics behind quantum mechanics gives us interesting insights into the parameter estimation theory. First, in a quantum world, many quantities are not directly observable. For example, the measure of entanglement between two parties, represented by a nonphysical quantum map, is not a directly measurable quantity. We need to adopt techniques of quantum tomography to surpass these obstacles, which allow us to determine the whole quantum state. Quantum theory provides us with a tool for mathematically describing the measurement process. It is possible to optimize further the classical CRLB (3.8) over all possible measurement settings to find the ultimate quantum CRLB (qCRLB).

First, let us introduce a general description of quantum states and quantum measurements. A general quantum state is represented as a statistical mixture of basis pure states

$\{|\psi_i\rangle\}$ and therefore can be written as

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|, \quad (3.10)$$

where p_i 's are probabilities that the system can be found in the state $|\psi_i\rangle$. The relation between measurement, represented by positive operator-valued measures (POVMs) with elements $\{\Pi_i\}$ is given by a Born's rule

$$p(x) = \text{Tr}(\rho\Pi_x). \quad (3.11)$$

Considering that the quantum state ρ carries information about parameter θ , we can also write the probabilities given by eq. (3.11) are dependent on the parameter θ , $p(x|\theta)$. Recalling the definition of the fisher information (3.9), we have to deal with the differentiation of a density operator ρ_θ with respect to parameter θ . We can do this by introducing the symmetric logarithmic derivative

$$\frac{\partial\rho_\theta}{\partial\theta} = \frac{L_\theta\rho_\theta + \rho_\theta L_\theta}{2}, \quad (3.12)$$

which allows us to write the derivative of the probability as

$$\partial_\theta p(x|\theta) = \text{Re}(\text{Tr}(\rho_\theta L_\theta \Pi_x)). \quad (3.13)$$

Consequently, if we plug this result into the eq. (3.9), we arrive with the formulae

$$F(\theta) = \int dx \frac{\text{Re}(\text{Tr}(\rho_\theta L_\theta \Pi_x))^2}{\text{Tr}(\rho_\theta \Pi_x)}. \quad (3.14)$$

3.2.2 Quantum Fisher information and quantum Cramér-Rao bound

Now, we are ready to optimize the eq. (3.14) over all possible measurements $\{\Pi_x\}$.

$$\begin{aligned} F(\theta) &\leq \int dx \left| \frac{\text{Tr}(\rho_\theta \Pi_x L_\theta)}{\sqrt{\text{Tr}(\rho_\theta \Pi_x)}} \right|^2 \\ &= \int dx \left| \text{Tr} \left(\frac{\sqrt{\rho_\theta} \sqrt{\Pi_x}}{\sqrt{\text{Tr}(\rho_\theta \Pi_x)}} \sqrt{\Pi_x} L_\theta \sqrt{\rho_\theta} \right) \right|^2 \\ &\leq \int dx \text{Tr}(\Pi_x L_\theta \rho_\theta L_\theta) \\ &= \text{Tr}(\rho_\theta L_\theta^2), \end{aligned} \quad (3.15)$$

where we used the famous Cauchy-Schwartz inequality $|\text{Tr}(A^\dagger B)|^2 \leq \text{Tr}(A^\dagger A)\text{Tr}(B^\dagger B)$. Therefore, the classical FI is always constrained by a quantity Q :

$$F(\theta) \leq Q(\theta) = \text{Tr}(\rho_\theta L_\theta^2), \quad (3.16)$$

called the quantum Fisher information (QFI). This also defines the quantum Cramér-Rao lower bound for variance of an unbiased estimator

$$\text{Var}(\hat{\theta}) \geq \frac{1}{NQ(\theta)}, \quad (3.17)$$

where N is the number of detection events. Even though calculating QFI can be mathematically very challenging, let us consider the simple example of a single parameter θ being encoded into the pure state $\rho_\theta = |\psi(\theta)\rangle\langle\psi(\theta)|$. the Fisher information can be written as

$$F(\theta) = \text{tr}(\rho_\theta (\partial_\theta \rho_\theta)^2) = 4\langle\psi(\theta)| \left(\frac{\partial \rho_\theta}{\partial \theta}\right)^2 |\psi(\theta)\rangle. \quad (3.18)$$

If we now use the fact the $\partial_\theta \rho_\theta = |\partial_\theta \psi(\theta)\rangle\langle\psi(\theta)| + |\psi(\theta)\rangle\langle\partial_\theta \psi(\theta)|$, the result simplifies to

$$Q(\theta) = 4(\langle\partial\psi(\theta)|\partial\psi(\theta)\rangle + |\langle\psi(\theta)|\partial\psi(\theta)\rangle|^2). \quad (3.19)$$

3.3 Quantum tomography

So far, we have been interested in estimating signal parameters from noisy data. However, if we would like to know the whole description of the quantum system, it is no longer sufficient to estimate only a few parameters. Quantum systems are mathematically described as density matrix ρ with the following properties; quantum states are positive semidefinite operators $\rho \geq 0$ with unit trace, $\text{Tr}(\rho) = 1$. From the positivity condition, density matrices are also hermitian, $\rho^\dagger = \rho$. Considering that the source's state spans d -dimensional Hilbert space, the density matrix ρ is fully described by $d^2 - 1$ independent parameters. The goal of quantum state tomography is to estimate ρ from measurements performed on identical copies of the system. Generally, the measurement is represented by positive operator-valued measures (POVMs). They are a set of positive Hermitian operator $\{\Pi_i\}$ with the properties

$$\Pi_i \geq 0, \quad \Pi_i^\dagger = \Pi_i, \quad \sum_i \Pi_i = 1 \quad \forall i. \quad (3.20)$$

Each POVM element represents a single output of a measurement apparatus. To describe the whole quantum state of the source, the number of linearly independent meas-

measurements should be larger than the number of free parameters describing the state. Such a measurement is then informationally complete measurement. The connection between POVMs and the output probabilities p_i is given by a Born's rule defined in eq. (3.11). After choosing a set of traceless Hermitian operators $\{\Gamma_j\}$ forming a basis in a d -dimensional Hilbert space, we can rewrite the Born's rule into a set of linear equations

$$\mathbf{p} = C\mathbf{r}, \quad (3.21)$$

where now the state is characterized by a real Bloch vector \mathbf{r} with entries $r_k = \text{Tr}(\rho\Gamma_k)$, whereas the $C_{lk} = \text{Tr}(\Pi_l\Gamma_k)$ is a measurement matrix describing the relation between quantum state ρ and the theoretical probabilities \mathbf{p} . In the presence of noise and with a finite number of copies of a quantum state, the collected data \mathbf{f} will differ from the expected values \mathbf{p} . A naive solution would be to perform a linear inversion of the measurement matrix C . However, the resulting estimated vector $\hat{\mathbf{r}}$ is no longer guaranteed to represent a positive semidefinite operator. To bypass this obstruction, another possibility how to solve the eq. (3.21) is to adopt a semidefinite program with the positivity constraint. The resulting estimator of \mathbf{r} is then a solution of

$$\text{minimize } \|\mathbf{f} - C\mathbf{r}\|^2; \quad \rho \geq 0 \quad \text{and} \quad \text{Tr}(\rho) = 1. \quad (3.22)$$

Last, let us briefly discuss the maximum likelihood estimator (MLE). As the name of the procedure suggests, the quantum state is reconstructed via maximizing the log-likelihood function $\mathcal{L}(\rho) \approx \sum_i f_i \log p_i(\rho)$, which can be written [68] as the iterative map

$$\rho^{(k+1)} \leftarrow \mu_k R \rho^{(k)} R, \quad (3.23)$$

where μ_k is the normalization constant and the operator R is defined as $R = \sum_i \frac{f_i}{p_i} \Pi_i$. Usually, we start the iteration process from the maximally mixed state $\rho^{(0)} = 1/d$. The positivity constraint persists throughout the interaction process, and typically, we need to perform thousands of iterations to observe the stationary point of the map (3.23).

3.4 Deep neural networks

Artificial neural networks have recently revolutionized science, technology, and everyday life. As a few examples, they are used to classify images, translate texts, drive cars, or play complex games such as chess at a superhuman level. In our work, we explore the ability of neural networks to infer general quantum states from noisy data and their ability

to predict entanglement from informationally incomplete measurements. Let us introduce the basic concepts and theory behind the fully connected deep neural network, which forms a structure used in our scientific papers.

3.4.1 Fully connected neural networks

Essentially, deep neural networks can be viewed as powerful approximators of general functions that can be trained using extensive data. Let's suppose that we want to approximate a function F ,

$$\mathbf{y} = F(\mathbf{x}), \quad (3.24)$$

which input a vector \mathbf{x} and outputs vector \mathbf{y} . In practice, we might have very little or no information about the function F , and thus we can only get a gist of it from a large number of samples, where each sample is given by an input-output combination (\mathbf{x}, \mathbf{y}) . The neural network presents an efficient way of approximating the function F via

$$\mathbf{y} = F_{\theta}(\mathbf{x}), \quad (3.25)$$

where the parameters $\theta = \{\theta_1, \theta_2, \dots\}$ are adjusted during the training process to match outputs with inputs precisely. Let us consider the following example. The vector \mathbf{x} represents the intensity data of a beam that passes through an optical system. For example, the vector \mathbf{y} stands for the first ten values of aberrations. The neural network then learns to predict imperfections caused by the system from just intensity scans, which allows us to calibrate the experiment more precisely.

3.4.1.1 The layout of a neural network

As in the human brain, neurons form the basis of artificial networks. In the case of a fully connected neural network, neurons are arranged in layers, where each neuron in a given layer is connected to every neuron in the next layer. To be more precise, suppose that the value of the k^{th} neuron in the $(n)^{\text{th}}$ layer is $y_k^{(n)}$. Then, the transformation of the signal from the $(n)^{\text{th}}$ layer to the $(n+1)^{\text{th}}$ layer is described as

$$\begin{aligned} z_j^{(n+1)} &= \sum_k w_{jk}^{(n+1)} y_k^{(n)} + b_j^{(n)}, \\ y_j^{(n+1)} &= f\left(z_j^{(n+1)}\right). \end{aligned} \quad (3.26)$$

The weight matrix $(W^{(n+1)})_{jk} = w_{jk}^{(n+1)}$ together with the bias vector $(B^{(n+1)})_j = b_j^{(n+1)}$ forms a linear transformation from layer to layer. However, if a deep neural network would

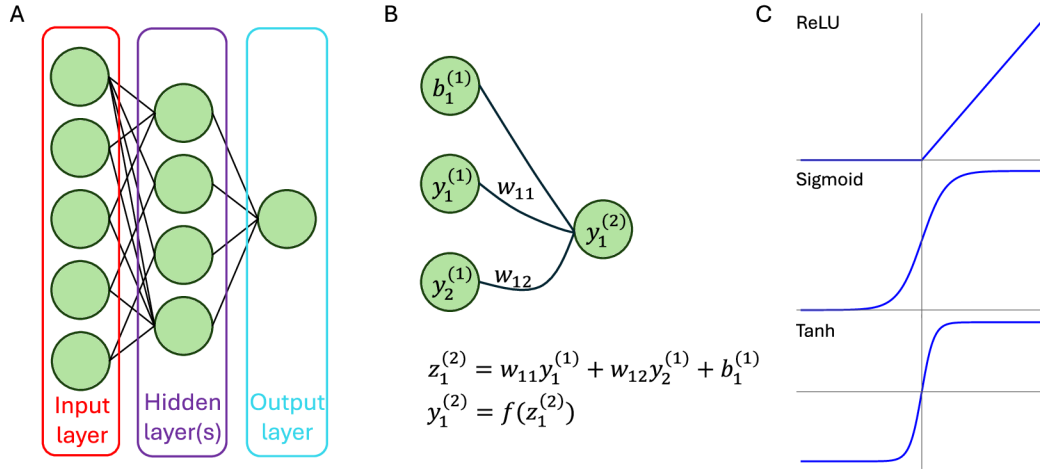


Figure 3.2: Schematic of a layout of a neural network. Panel (A) represents a sketch of a neural network with one deep layer. Each neuron from a deep layer is connected to every neuron in the next layer. Panel (B) shows the transition of information in a neural network, also called forward propagation. The value of a neuron in the second layer is a weighted sum of the values of all neurons in the previous layer plus bias. Furthermore, this number is fed to the nonlinear function of our choice, which finally gives us the value of the neuron in the second layer. Panel (C) shows the three most used nonlinear activation functions. The first one is the Rectified Linear Unit (ReLU) (3.27), the second one is the sigmoid function (3.28), and the third one is the hyperbolic tangent function (3.29).

perform only subsequent linear transformations, the network would again perform only a linear transformation as a whole. For this reason, applying a nonlinear function $f(\cdot)$, called the activation function, is key. As for a few examples of commonly used activation functions, let us mention the Rectified Linear Unit (ReLU) function, the sigmoid function, or the hyperbolic tangent function. The ReLU activation function is the simplest example of a nonlinear function defined as

$$f_{\text{ReLU}}(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } x > 0, \end{cases} \quad (3.27)$$

and it is commonly used in hidden layers of the deep neural network. The second example is the sigmoid activation function defined by

$$f_{\text{sigmoid}}(x) = \frac{1}{1 + e^{-x}}. \quad (3.28)$$

The sigmoid activation function maps the real axis to the interval $[0, 1]$, so we can use it as an activation function in the last layer of a deep neural network if we estimate a quantity restricted to the same interval. In our work, we then use it in neural networks in

the last layer to estimate the quantum mutual information and the concurrence, measures of quantum entanglement. Finally, let us mention the hyperbolic tangent function

$$f_{\tanh}(x) = \tanh(x), \quad (3.29)$$

which is bounded from below by -1 and from above by 1 . Once again, this activation function can find its place in the last layer of the deep neural network if the estimated quantity has the same bounds.

3.4.1.2 Training of a neural network

The number of deep layers, the number of neurons in each layer, and the choice of activation functions on neurons determine the structure of the neural network. However, each neural network needs to be trained on prelabeled data. To do this, we will need to choose a measure of deviation, or cost function, between the true data $F(\mathbf{x})$ and the predictions of the neural network $F_\theta(\mathbf{x})$. In the simplest case, we might measure the quadratic deviation between the true answer $F(\mathbf{x})$ and the network output $F_\theta(\mathbf{x})$. First, we define a sample-specific cost function $C_{\mathbf{x}}$ as

$$C_{\mathbf{x}}(\theta) = |F(\mathbf{x}) - F_\theta(\mathbf{x})|^2. \quad (3.30)$$

Then, by averaging over all data, we define the cost function as

$$C(\theta) = \langle C_{\mathbf{x}}(\theta) \rangle_{\mathbf{x}}. \quad (3.31)$$

The choice of a cost function is essential for adequately training the neural network, and it varies depending on the data type and structure we are trying to predict. With the cost function at hand, we aim to find its global minimum. A simple method would be to follow the negative gradient of the cost function and, in the k -th step of the training process, update the parameters of the neural network via

$$\theta_{k+1} = \theta_k - \eta \frac{\partial C(\theta)}{\partial \theta_k}, \quad (3.32)$$

where η is the learning rate. Once again, the value of η has to be chosen wisely to get the best performance of your neural network. The immediate challenge stems from the fact that the cost function is averaged over a large amount of data points, which is too expensive to compute in each step given that the number of parameters θ describing our network is usually 10^5 and more. This issue can be resolved by averaging over only some randomly chosen subset of all data points. Suppose we average the cost function over N

training samples, also called a batch:

$$C(\theta) \approx \frac{1}{N} \sum_{j=1}^N (C_{\mathbf{x}_j}(\theta)) = \langle C_{\mathbf{x}}(\theta) \rangle_{\text{batch}}. \quad (3.33)$$

This defines the stochastic gradient descent method

$$\theta_{k+1} = \theta_k - \eta \left\langle \frac{\partial C_{\mathbf{x}}(\theta)}{\partial \theta_k} \right\rangle_{\text{batch}} = \theta_k - \eta \frac{\partial C(\theta)}{\partial \theta_k} + \text{noise}. \quad (3.34)$$

The basic idea is that the noise averages out after sufficiently many learning steps, providing that η is small enough.

3.4.1.3 Backpropagation

Finally, we are left with calculating the gradient of the cost function (3.34) with respect to its parameters θ . In the early days of neural networks, numerical differentiation was applied. However, this approach is time-consuming and inefficient. Here, we will show that we can arrive at the formula for updating the parameters θ involving only matrix multiplication. For a quadratic cost function, we have

$$\frac{\partial C_x(\theta)}{\partial \theta_k} = 2 \sum_l ([F_\theta(\mathbf{x})]_l - [F(\mathbf{x})]_l) \frac{\partial [F_\theta(\mathbf{x})]_l}{\partial \theta_k}, \quad (3.35)$$

where $[F_\theta(\mathbf{x})]_l = y_l^{(n)}$ is the value of the neuron l in the output layer (n). In other words, the task was reduced to finding the gradient of the neuron value for all possible parameters θ in previous layers. Recalling (3.26), we get

$$\frac{\partial y_l^{(n)}}{\partial \theta_k} = f'(z_l^{(n)}) \frac{\partial z_l^{(n)}}{\partial \theta_k}, \quad (3.36)$$

where we have

$$\frac{\partial z_l^{(n)}}{\partial \theta_k} = \sum_m w_{l,m}^{(n,n-1)} \frac{\partial y_m^{(n-1)}}{\partial \theta_k}. \quad (3.37)$$

If we define the matrix $M_{l,m}^{(n,n-1)} = w_{l,m}^{(n,n-1)} f'(z_m^{(n-1)})$, the equation (3.37) can be viewed as a matrix-vector product. Combining the above equations, we finally arrive at the final formula

$$\frac{\partial z_l^{(n)}}{\partial \theta_k} = \left[M^{(n,n-1)} M^{(n-1,n-2)} \dots M^{(n'+1,n')} \right] \frac{\partial z_l^{(n')}}{\partial \theta_k}, \quad (3.38)$$

where the parameter θ_k is found in the (n') -th layer. This leads to the famous backpropagation algorithm [31]. Let us point out that backpropagation is equally demanding as a forward pass of a signal through the network. Thus, the backpropagation forms an effective algorithm for updating the neural network parameters by minimizing the cost function.

Chapter 4

Optimal measurement for quantum spatial superresolution

In this work, I helped carry out the calculations regarding the quantum Fisher information matrix (qFIM) and provided the subsequent analysis of the optimal measurement in the regime of small separations. I was also involved in the preparation of the manuscript.

Metrology, the science of precise parameter estimation, relies on schemes that extract the most accurate estimates possible. The quantum Cramér-Rao lower bound (qCRLB) limits estimation uncertainty and guides us in finding the optimal measurements. This work [1] explores the simultaneous estimation of centroid, separation, and relative intensities of two incoherent point sources using a linear optical system, addressing both theoretical and practical aspects of quantum-inspired imaging.

The motivation behind this research is two-fold. First, it aims to provide a theoretical foundation for quantum-enhanced imaging techniques that can be practically implemented. Second, it aims to demonstrate that these techniques can achieve superresolution, significantly improving the precision of measurements in various applications, such as observational astronomy, biological microscopy, and quantum information processing.

This work paves the way for future practical implementations of quantum-inspired imaging by addressing the challenges of multiparameter estimation in quantum metrology. The results presented here have the potential to revolutionize high-resolution imaging, providing new tools for scientific discovery and technological innovation.

The density operator for two incoherent sources of different intensities is given by

$$\rho_\theta = q\rho_+ + (1 - q)\rho_-, \quad (4.1)$$

where q and $1 - q$ are normalised intensities of the sources, and $\rho_\pm = |\psi_\pm\rangle\langle\psi_\pm|$ are the x -displaced PSF states, $|\psi_\pm\rangle = e^{-i(s_0 - s/2)P}|\psi\rangle$. In this multiparameter scenario, the central

quantity is the quantum Fisher information matrix [17], defined as $Q_{\alpha\beta}(\theta) = \frac{1}{2}\text{Tr}(\rho_\theta\{L_\alpha, L_\beta\})$. After a lengthy calculation, we arrived with a compact expression

$$Q = 4 \begin{pmatrix} p^2 + 4q(1-q)\wp^2 & (q-1/2)p^2 & -iw\wp \\ (q-1/2)p^2 & p^2/4 & 0 \\ -iw\wp & 0 & \frac{1-w^2}{4q(1-q)} \end{pmatrix} \quad (4.2)$$

where $p^2 = \langle\psi|P^2|\psi\rangle$, $\wp = \langle\psi|e^{isP}P|\psi\rangle$, and $w = \langle\psi|e^{isP}|\psi\rangle$. As one might expect, the measurement of separation is not correlated with the other two parameters for equally bright sources. In general and realistic scenarios, when $q \neq 1/2$, all three parameters are correlated. In general, finding an optimal measurement scheme in closed form is unattainable. However, considering only small separations, we can proceed as follows. First, we consider a specific basis in the signal space made up of spatial derivatives of the amplitude PSF $\langle x|\Psi_n\rangle = \frac{\partial^n}{\partial x^n}\Psi(x-x_0)$ for $n = 0, 1, 2, \dots$, where x_0 is an arbitrary displacement in the x representation. The Gramm-Schmidt process can convert Such a basis into an orthogonal one. Using this basis, we can construct four linearly independent projectors $\Pi_j = |\pi_j\rangle\langle\pi_j|$ by expanding the signal components in the small parameter. Lastly, we showed that if we set the displacement precisely as

$$x_0^{\text{opt}} = (1-q)(s_0 - s/2) + q(s_0 + s/2), \quad (4.3)$$

We can saturate the quantum bounds with a resolution of all three parameters. This can be shown by evaluating the classical Fisher information matrix (3.9) with $p_j = q\langle\Psi_+|\Pi_j|\Psi_+\rangle + (1-q)\langle\Psi_-|\Pi_j|\Psi_-\rangle$. The result (4.3) shows that the optimal choice of displacement is the weighted centroid rather than the geometrical centroid. In practice, we would have to use adaptive schemes to achieve multiparameter superresolution in the regime of small separations.

In summary, we studied the ultimate limits for the simultaneous estimation of centroid, separation, and relative intensities of two incoherent light sources. We demonstrated that the quantum spatial superresolution of two incoherent unbalanced sources can be achieved by projecting the signal onto a suitable complete set of modes. These results pave the way for future experimental implementations and innovations in quantum-inspired imaging, particularly in observational astronomy and microscopy.

Optimal measurements for quantum spatial superresolution

J. Řeháček,¹ Z. Hradil,¹ D. Koutný,¹ J. Grover,² A. Krzic,² and L. L. Sánchez-Soto^{3,4}

¹Department of Optics, Palacký University, 17. listopadu 12, 771 46 Olomouc, Czech Republic

²ESA—Advanced Concepts Team, European Space Research Technology Centre (ESTEC), Keplerlaan 1, Postbus 299, NL-2200AG Noordwijk, Netherlands

³Departamento de Óptica, Facultad de Física, Universidad Complutense, 28040 Madrid, Spain

⁴Max-Planck-Institut für die Physik des Lichts, Staudtstraße 2, 91058 Erlangen, Germany



(Received 22 December 2017; published 3 July 2018)

We construct optimal measurements, achieving the ultimate precision predicted by quantum theory, for the simultaneous estimation of centroid, separation, and relative intensities of two incoherent point sources using a linear optical system. We discuss the physical feasibility of the scheme, which could pave the way for future practical implementations of quantum-inspired imaging.

DOI: [10.1103/PhysRevA.98.012103](https://doi.org/10.1103/PhysRevA.98.012103)

I. INTRODUCTION

Metrology is the science of devising schemes that extract as precise as possible an estimate of the parameters associated with a system. The quantum foundations of this field were laid years ago [1,2]; since then, most of the efforts have been devoted to single-parameter estimation, with a special emphasis in the prominent example of phase [3,4]. The quantum Cramér-Rao lower bound (qCRLB) then provides a saturable bound on the estimation uncertainty, and recipes for finding the optimal measurement attaining that limit are known [5].

The case of multiparameter estimation is considerably more involved [6–8]. Although the equivalent qCRLB was formulated long time ago [9], this bound is not always saturable. The intuitive reason for this is the incompatibility of the measurements for different parameters. The conditions under which the qCRLB can be saturated have been determined [10,11]. The associated optimal measurements have been worked out for pure states [12], but for mixed states the results are fragmentary [13–15].

In this work we will address these problems in the context of the two-point resolution limit for an optical system. In classical optics several criteria exist [16–18] to quantitatively determine these limits, the most famous of which is due to Rayleigh [19].

Most of these criteria exploit properties of the point spread function (PSF) that specifies the intensity response to a point light source. This provides an intuitive picture of the mechanisms limiting resolution, but also has several shortcomings. These mainly stem from the fact that these criteria were developed for the human eye as the main detector. For example, the Rayleigh limit is defined as the distance from the center to the first minimum of the PSF, which can be made arbitrarily small with ordinary linear optics, although at the expense of the side lobes becoming much higher than the central maximum [20]. This confirms that determining the position of the two points becomes also a question of photon statistics rather than being solely described by the Rayleigh limit.

A careful reconsideration of this conundrum has been performed in the framework of quantum estimation

theory [21–28]. This work showed that, in the case of two identical incoherent point sources with *a priori* knowledge of their centroid, the precision of an optimal measurement stays constant at all separations. As a consequence, the Rayleigh limit is subsidiary to the problem and arises because standard direct imaging discards all the phase information contained in the field. These predictions fuelled a number of proof-of-principle experiments [29–32].

While remarkable, this result does not hold in the more general case of two unequally bright sources. In a suitable multiparameter scenario [33], where simultaneous estimation of centroid, separation, and relative brightness was considered, it was found that their estimation precisions decreased with separation [34]. Nonetheless, an appropriate strategy was shown to lead to a significant improvement in precision at small separations over direct imaging for any fixed number of photons. The measurements attaining the ultimate quantum limits for this case are relevant to a number of applications, for example, observational astronomy and microscopy.

II. MODEL AND ASSOCIATED MULTIPARAMETER QUANTUM CRAMÉR-RAO BOUND

To be as self-contained as possible, we first set the stage for our analysis. We assume a linear spatially invariant system illuminated with quasimonochromatic paraxial waves with one specified polarization. We consider one spatial dimension, x denoting the image-plane coordinate.

We phrase what follows in a quantum language that will simplify the following calculations. To a field of complex amplitude $U(x)$ we assign a ket $|U\rangle$, such that $U(x) = \langle x|U\rangle$, $|x\rangle$ being a pointlike source at x . The system PSF is denoted by $I(x) = |\langle x|\Psi\rangle|^2 = |\Psi(x)|^2$, so that $\Psi(x)$ can be interpreted as the amplitude PSF.

Two point sources, of different intensities and separated by a distance s , are imaged by that system. Since they are incoherent with respect to each other, the total signal must be depicted as a density operator,

$$\varrho_{\theta} = q \varrho_{+} + (1 - q) \varrho_{-}, \quad (1)$$

where q and $1 - q$ are the intensities of the sources (the total intensity is normalized to unity). The individual components $\varrho_{\pm} = |\Psi_{\pm}\rangle\langle\Psi_{\pm}|$ are just x -displaced PSF states; that is, $\langle x|\Psi_{\pm}\rangle = \langle x - s_0 \mp s/2|\Psi\rangle$, so that they are symmetrically located around the geometric centroid s_0 . Note that

$$|\Psi_{\pm}\rangle = \exp[-i(s_0 \pm s/2)P]|\Psi\rangle, \quad (2)$$

where P is the momentum operator, which generates displacements in the x variable, and acts as a derivative $P \mapsto -i\partial_x$.

The measured density matrix depends on the centroid s_0 , the separation s , and the relative intensities of the sources q . This is indicated by the vector $\theta = (s_0, s, q)^t$. Our task is to estimate the values of θ through the measurement of some observables on ϱ_{θ} .

In this multiparameter estimation scenario, the central quantity is the quantum Fisher information matrix (qFIM) [35]. This is a natural generalization of the classical Fisher information, which is a mathematical measure of the sensitivity of a quantity to changes in its underlying parameters. However, the qFIM is optimized over all the possible measurements. It is defined as

$$Q_{\alpha\beta}(\theta) = \frac{1}{2}\text{Tr}(\varrho_{\theta}\{L_{\alpha}, L_{\beta}\}), \quad (3)$$

where the Greek indices run over the components of the vector θ and $\{\cdot, \cdot\}$ denotes the anticommutator. Here, L_{α} stands for the symmetric logarithmic derivative (SLD) [9] with respect to the parameter θ_{α} :

$$\frac{1}{2}(L_{\alpha}\varrho_{\theta} + \varrho_{\theta}L_{\alpha}) = \partial_{\alpha}\varrho_{\theta}, \quad (4)$$

with $\partial_{\alpha} = \partial/\partial\theta_{\alpha}$.

The qFIM is a distinguishability metric on the space of quantum states and leads to the multiparameter qCRLB [5,8] for a single detection event:

$$\text{Cov}(\hat{\theta}) \geq Q^{-1}(\theta), \quad (5)$$

where $\text{Cov}(\hat{\theta})$ is the covariance matrix for a locally unbiased estimator $\hat{\theta}$ of the quantity θ . Its matrix elements are $\text{Cov}_{\alpha\beta}(\hat{\theta}) = \mathbb{E}[(\hat{\theta}_{\alpha} - \theta_{\alpha})(\hat{\theta}_{\beta} - \theta_{\beta})]$, $\mathbb{E}[Y]$ being the expectation value of the random variable Y . The above inequality should be understood as a matrix inequality. In general, we can write $\text{Tr}[\mathcal{C}\text{Cov}(\hat{\theta})] \geq \text{Tr}[\mathcal{C}Q^{-1}(\theta)]$, where \mathcal{C} is some positive cost matrix, which allows us to asymmetrically prioritise the uncertainty cost of different parameters.

Unlike for a single parameter, the collective bound in Eq. (5) is not always saturable, as the measurements for different parameters may be incompatible [2]. The multiparameter qCRLB can be saturated provided

$$\text{Tr}(\varrho_{\theta}[L_{\alpha}, L_{\beta}]) = 0, \quad (6)$$

where $[\cdot, \cdot]$ is the commutator. This condition is necessary and sufficient for pure states [10,11], upon which the criterion is equivalent to the existence of some pair of SLDs that commute. It is then possible to find an optimal measurement as the common eigenbasis of these SLDs. For mixed states, this criterion has been discussed by a number of authors [36] and has met some small inconsistencies in its usage, being variously identified as sufficient [37] or necessary and sufficient [38]. Reference [39] offers a clear account of this question. For our particular case, Eq. (6) is fulfilled whenever the PSF amplitude

is real [34], $\Psi(x)^* = \Psi(x)$, which will be assumed henceforth ensuring that the parameters are therefore compatible.

For the model we are considering, and after a lengthy calculation [34], we obtain a compact expression for the qFIM; viz

$$Q = 4 \begin{pmatrix} p^2 + 4q(1-q)\wp^2 & (q-1/2)p^2 & -iw\wp \\ (q-1/2)p^2 & p^2/4 & 0 \\ -iw\wp & 0 & \frac{1-w^2}{4q(1-q)} \end{pmatrix}, \quad (7)$$

which depends solely on the quantities

$$\begin{aligned} w &\equiv \langle\Psi_{\pm}|\Psi_{\mp}\rangle = \langle\Psi|\exp(i\wp P)|\Psi\rangle, \\ p^2 &\equiv \langle\Psi_{\pm}|P^2|\Psi_{\pm}\rangle = \langle\Psi|P^2|\Psi\rangle, \\ \wp &\equiv \pm\langle\Psi_{\pm}|P|\Psi_{\mp}\rangle = \langle\Psi|\exp(i\wp P)P|\Psi\rangle. \end{aligned} \quad (8)$$

The quantity p^2 is determined by the shape of the PSF, whereas both w and \wp (which is purely imaginary) depend on the separation s .

Only for equally bright sources, $q = 1/2$, the measurement of s is uncorrelated with the other parameters. In general, when $q \neq 1/2$ the separation is correlated with the centroid (via the intensity term $q - 1/2$) and the centroid is correlated with the intensity (via p^2).

The individual parameter θ_{α} can be estimated with a variance satisfying $\text{Var}(\hat{\theta}_{\alpha}) \geq (Q^{-1})_{\alpha\alpha}(\theta)$. It is convenient to use the inverses of the variances $H_{\alpha} = 1/\text{Var}(\theta_{\alpha})$, usually called the precisions [40]. By inverting the QFIM and taking the limit $s \rightarrow 0$, they turn out to be [34]

$$\begin{aligned} H_{s_0}^Q &\simeq \mathcal{Q}^2 G_{22} s^2 + O(s^4), \\ H_s^Q &\simeq \frac{\mathcal{Q}^2}{4(1-\mathcal{Q}^2)} G_{22} s^2 + O(s^4), \\ H_q^Q &\simeq \frac{1}{\mathcal{Q}^2} G_{22} s^4 + O(s^6), \end{aligned} \quad (9)$$

where

$$\begin{aligned} \mathcal{Q}^2 &= 4q(1-q) < 1, \\ G_{22}^2 &= \text{Var}(P^2) = \langle\Psi|P^4|\Psi\rangle - \langle\Psi|P^2|\Psi\rangle^2. \end{aligned} \quad (10)$$

The superscript Q indicates that the quantities are evaluated from the quantum matrix Q .

III. OPTIMAL MEASUREMENTS

We shall focus on finding measurements attaining the quantum limit, thus offering significant advantages with respect to conventional direct intensity measurements. In the general case of unequally bright sources ($q \neq 1/2$), the lack of symmetry makes this issue challenging and one cannot expect to find closed-form expressions for the optimal positive operator valued measures (POVMs) for all the values of the source parameters. However, this becomes viable when separations get very small. As already discussed, this is the most interesting regime, where conventional imaging techniques fail.

We start by specifying a basis in the signal space. A suitable choice is the set $\{|\Psi_n\rangle\}$ defined in terms of the spatial

derivatives of the amplitude PSF:

$$\langle x|\Psi_n\rangle = \frac{\partial^n}{\partial x^n}\Psi(x-x_0), \quad n=0,1,2,\dots, \quad (11)$$

where x_0 is an arbitrary displacement in the x representation. We convert this set into an orthonormal basis $\{|\Phi_n\rangle\}$ by the standard Gram-Schmidt process. In this basis, all results can be expressed in a PSF-independent form. Moreover, signals well centered on the origin and with small separation, are represented by low-dimensional states; i.e., $\varrho_\theta \rightarrow |\Phi_0\rangle\langle\Phi_0|$ for $\mathfrak{s}_0 \rightarrow x_0$, and $\mathfrak{s} \rightarrow 0$.

To estimate three independent parameters, the required POVM must have at least four elements. We therefore consider the following class of measurements $\Pi_j = |\pi_j\rangle\langle\pi_j|$, $j=0,\dots,2$ and $\Pi_3 = \mathbb{1} - \Pi_0 - \Pi_1 - \Pi_2$, so only three of these are independent. The first three POVM elements are defined in a four-dimensional subspace, with basis $\{|\Phi_0\rangle,\dots,|\Phi_3\rangle\}$, wherein we expand $|\pi_j\rangle$ ($j=0,\dots,2$) as

$$|\pi_j\rangle = \sum_{k=0}^3 C_{jk}|\Phi_k\rangle. \quad (12)$$

Obviously, the projectors $|\pi_j\rangle\langle\pi_j|$ must be linearly independent. In addition, we impose the following set of conditions

$$\begin{array}{cc} & |\Phi_0\rangle & |\Phi_1\rangle, \\ |\pi_0\rangle & C_{00} = 0 & C_{01} \neq 0, \\ |\pi_1\rangle & C_{10} = 0 & C_{11} \neq 0, \\ |\pi_2\rangle & C_{20} \neq 0 & C_{21} \neq 0, \end{array} \quad (13)$$

where the row index can be permuted. In this way, two of the three rank-one projectors are orthogonal to the signal PSF Φ_0 , a crucial factor boosting the performance of the measurement. We stress that, by changing the displacement x_0 , the basis and the measurement itself is displaced.

Next, we expand the signal components in the small parameter. We define $a_\pm = \mathfrak{s}_0 \pm \mathfrak{s} - x_0$, so we have $\langle x|\Psi_\pm\rangle = \Psi(x-x_0-a_\pm)$, and the expansion in a_\pm gives

$$\begin{aligned} \langle x|\Psi_\pm\rangle &= \sum_m \frac{(-a_\pm)^m}{m!} \frac{\partial^m}{\partial x^m}\Psi(x-x_0) = \sum_m \frac{(-a_\pm)^m}{m!} \langle x|\Psi_m\rangle \\ &= \langle x|\sum_n |\Phi_n\rangle \sum_m \frac{(-a_\pm)^m}{m!} G_{nm}, \end{aligned} \quad (14)$$

where $G_{nm} = \langle\Phi_n|\Psi_m\rangle$ [note that G_{22} in Eq. (10) is consistent with this general definition]. Keeping terms up to the fourth power, we get

$$\begin{aligned} |\Psi_\pm\rangle &\simeq \left(G_{00} + \frac{a_\pm^2}{2}G_{02} + \frac{a_\pm^4}{24}G_{04}\right)|\Phi_0\rangle \\ &+ \left(a_\pm G_{11} + \frac{a_\pm^3}{6}G_{13}\right)|\Phi_1\rangle \\ &+ \left(\frac{a_\pm^2}{2}G_{22} + \frac{a_\pm^4}{24}G_{24}\right)|\Phi_2\rangle + \frac{a_\pm^3}{6}G_{33}|\Phi_3\rangle. \end{aligned} \quad (15)$$

Notice that for real amplitude PSFs, all G s carrying both odd and even subscripts are zero. This follows from the fact that $\langle\Psi_n|\Psi_m\rangle = 0$, and hence $\langle\Psi|P^{m+n}|\Psi\rangle = 0$ for any combination of odd and even subscripts, whenever the wavefunction is real. We also have $G_{nm} = 0$ for all $n > m$, by

construction of the basis set, which makes a basis function orthogonal to all lower-order nonorthogonal functions, as the latter span a subspace that the former is orthogonal to.

We are set to evaluate the probabilities,

$$p_j = q\langle\Psi_+|\Pi_j|\Psi_+\rangle + (1-q)\langle\Psi_-|\Pi_j|\Psi_-\rangle, \quad (16)$$

and the corresponding classical Fisher information matrix per detection event,

$$F_{\alpha\beta} = \sum_{j=0}^3 \frac{(\partial_\alpha p_j)(\partial_\beta p_j)}{p_j}. \quad (17)$$

The maximum of the classical Fisher information F is its quantum version Q , as Q is optimized over all POVMs. The corresponding precisions are thus related by $H_\alpha^Q \geq H_\alpha$.

Our initial strategy is to align the center of the measurement Eq. (12) with the signal centroid by letting $x_0 = \mathfrak{s}_0$. The calculation of the precisions turns out to be a very lengthy task, yet the final result is surprisingly simple:

$$H_\alpha = \lambda H_\alpha^Q. \quad (18)$$

Therefore, H_α differs from the quantum limit precision by a factor

$$\lambda = \mathcal{Q}^2 \mathcal{A}, \quad \mathcal{A} = \frac{(C_{01}C_{12} - C_{02}C_{11})^2}{C_{01}^2 + C_{11}^2} < 1. \quad (19)$$

The coefficient λ consists of the product of two factors: one depending solely on the intensities [as defined in Eq. (10)], the other depending on the measurement. The latter one will be called the quality factor of the measurement. Conditions Eq. (13) are crucial for deriving relations Eqs. (18) and (19): Violating them makes the dominant terms of H_α disappear and kills the superresolution. One pertinent example would be projection on the basis set $|\Phi_k\rangle$: $C_{jk} = \delta_{jk}$ as for example projections on a set of Hermite-Gauss modes for a Gauss PSF advocated in Refs. [21] and [25], among others. Such projections can be optimal for estimating separation, but ultimately fail when separation, centroid, and intensity are to be estimated together in a multiparameter scenario considered here.

Going back to our result, two remarks are in order here. First, the performance of the measurement Eq. (12), when aligned with the centroid, scales with the same power of \mathfrak{s} as the quantum limit does. The quantum limit is attained, but for a separation independent factor. This is true for all real-valued PSFs, no matter how we set the remaining free parameters of the measurement. Second, by optimizing those free parameters, the separation-independent factor λ can be made arbitrarily close to $\lambda_{\max} = \mathcal{Q}^2$. Hence, for balanced signals ($q = 1/2$), $\lambda_{\max} \rightarrow 1$ and the measurement Eq. (12) becomes optimal. Conversely, for unbalanced signals, the measurement is suboptimal and its performance worsens with q , approaching the limit $\lambda \rightarrow 0$ when $q \rightarrow 0$ and $q \rightarrow 1$.

Next, we show that quantum limits can be saturated for any q by optimizing the displacement x_0 . The key point is that in the limit $\mathfrak{s} \ll 1$, the precisions $H_\alpha(x_0)$, when considered as a function of the measurement displacement x_0 , take a Lorentzian shape, as can be appreciated in Fig. 1 for the particular case of $H_\mathfrak{s}(x_0)$. On decreasing the signal separation, the Lorentzian narrows down, with its center approaching the

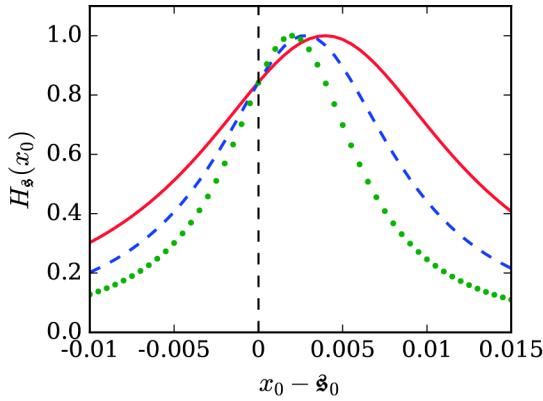


FIG. 1. The precision H_s of the separation for point objects with relative intensities $q = 0.3$ and $s = 0.02$ (red solid line), $s = 0.014$ (blue broken line), and $s = 0.01$ (green dots) as a function of misalignment $x_0 - s_0$ between the measurement displacement and the centroid. The maxima of the Lorentzians are normalized to unity to make the changes in widths and centers apparent.

signal centroid. We therefore adopt the model

$$H_s(x_0) = \frac{\ell_1 s^2}{1 + \frac{\ell_2(x_0 - s_0 + \ell_3 s)^2}{s^2}}. \quad (20)$$

The parameters can be identified by expanding H_s in s and $x_0 - s_0$:

$$\ell_1 s^2 = \mathcal{A} H_s^Q, \quad \ell_2 = \frac{1}{q(1-q)}, \quad \ell_3 = \frac{1}{2}(1-2q). \quad (21)$$

This uncovers the optimal displacement and precisions

$$x_0^{\text{opt}} = \arg \max_{x_0} H_s(x_0) = s_0 - \frac{1}{2}s(1-2q), \quad (22)$$

$$H_s(x_0^{\text{opt}}) = \mathcal{A} H_s^Q.$$

This is the central result of this paper. The optimal choice of displacement is precisely

$$x_0^{\text{opt}} = (1-q)(s_0 - s/2) + q(s_0 + s/2), \quad (23)$$

so that the weighted centroid, rather than the geometrical centroid, is relevant to align the measurement. Note that the weighted centroid only coincides with the center of mass of the PSF when the PSFs are symmetric. By optimizing the measurement displacement x_0 , the intensity dependent \mathcal{Q}^2 term is removed from Eqs. (18) and (19) and the qCRLBs are saturated for all the signal parameters simply by letting $\mathcal{A} \rightarrow 1$. As this can be done in infinitely many ways, we conclude there are infinitely many measurements attaining the quantum limit in multiparameter superresolution imaging. They can be constructed following our recipe for any real-valued amplitude PSF.

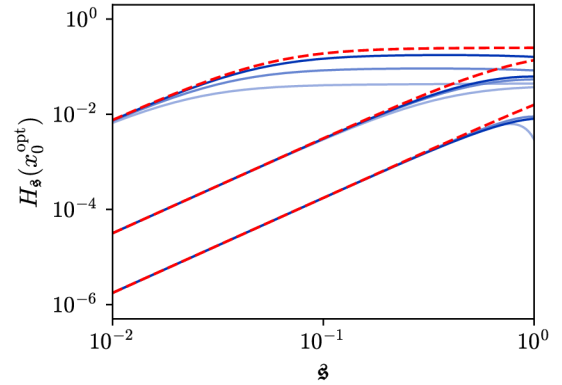


FIG. 2. The precision $H_s(x_0)$ for an optimally displaced measurement Eq. (24) (blue lines) as compared to the quantum limit (9) (red broken lines). The lines are grouped by the intensity difference: $q = 0.49$ (top), $q = 0.35$ (middle), and $q = 0.1$ (bottom). Within each group (light to dark) $\phi = \pi/4, 7\pi/20$, and $9\pi/20$, respectively. Notice the fast convergence towards the quantum limit as $s \rightarrow 0$. A Gaussian PSF of a unit width $\sigma = 1$ is assumed.

IV. EXAMPLES

To illustrate our result with a concrete example, we construct three orthogonal vectors through

$$|\pi_{0,1}\rangle = \begin{pmatrix} 0 \\ \frac{\sin(\phi/2)}{\sqrt{1+\cos\phi}} \\ \pm \frac{\cos(\phi/2)}{\sqrt{1+\cos\phi}} \\ -\sqrt{\frac{\cos\phi}{1+\cos\phi}} \end{pmatrix}, \quad |\pi_2\rangle = \begin{pmatrix} \sqrt{\frac{2\cos\phi}{1+3\cos\phi}} \\ \sqrt{\frac{2\cos\phi}{1+3\cos\phi}} \\ 0 \\ \sqrt{\frac{1-\cos\phi}{1+3\cos\phi}} \end{pmatrix}, \quad (24)$$

with $0 < \phi < \pi/2$ in the $|\Phi_k\rangle$ -representation, to build a family of POVMs according to the recipe Eq. (12). This measurement satisfies all the requirements, and the quality factor becomes $\mathcal{A} = 1$, so that the quantum limit is attained for any real-valued PSF as long as $s \ll \sigma$.

The theory thus far is largely independent of the actual form of the PSF. To be more specific, we adopt a Gaussian PSF, with unit width $\sigma = 1$, which will serve from now on as our basis unit length. The associated orthonormal basis is then a set of displaced Hermite-Gauss modes

$$\Phi_n(x) = \frac{1}{(2\pi)^{\frac{1}{4}} 2^{\frac{n}{2}} \sqrt{n!}} H_n[(x - x_0^{\text{opt}})/\sqrt{2}] \times \exp[-\frac{1}{4}(x - x_0^{\text{opt}})^2], \quad (25)$$

where $H_n(x)$ are the Hermite polynomials. In this case, we have then $G_{22} = 1/8$.

Figure 2 shows the resulting precision H_s as a function of s on a log-log scale for different intensities q and different measurements of the family Eq. (24). Direct numerical evaluation of the Fisher information Eq. (17) was done using a computational basis $\{\Phi_n\}$ of dimension 30 and no further approximation. With $s \rightarrow 0$ all precisions quickly converge towards the quantum limit and all the measurements Eq. (24) become optimal. Notice however that performances over a wider range of separations are sensitive to measurement

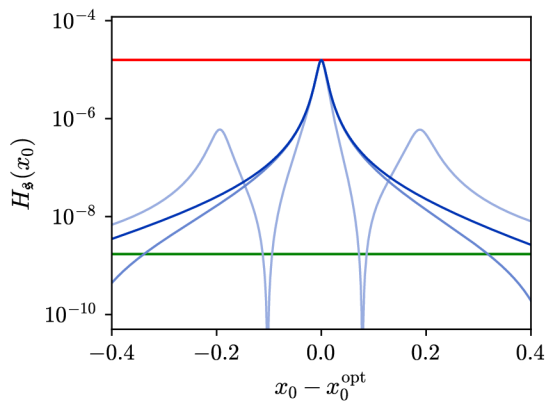


FIG. 3. The precision H_s for misaligned measurements Eq. (24) (solid blue lines) compared to the corresponding quantum (top red) and direct imaging (bottom green) limits. The parameters of the measurements are $\phi = \pi/20$ (light blue), $\pi/2$ (light blue), and $9\pi/20$ (dark blue). Observe the log vertical scale. A Gaussian PSF of a unit width $\sigma = 1$ is assumed; and the sources satisfy $s = 0.03$ and $q = 0.1$.

parameter ϕ and values close to $\phi = \pi/2$ provide the best overall performance.

Having potential applications of our proposed detection scheme in mind we realize that achieving the quantum limits requires knowing the true values of the measured parameters. In particular, the measurement must be optimally displaced to reach the quantum limits and this displacement, through Eq. (22), depends on all the unknown signal parameters. Consequently different displacements should be used for different signals.

Can one hope to saturate the quantum limits for all signals with a fixed measurement? Unfortunately, the answer is negative. Let us consider the estimation of a signal with strongly overlapping components $s \ll 1$ of highly unequal intensities $q \rightarrow 0$ (the same analysis can be carried out for $q \rightarrow 1$), so that the weak component is outshined. To gain significant information about the weak component, the bright one must be almost completely suppressed in one of the measurement outputs. This is ensured by projecting the signal on a state that is nearly orthogonal to the bright component. That crucial projection, though, depends on both the signal centroid and separation.

Our optimal measurement also behaves in this way. Let us look at the value of x_0^{opt} in the limit $q \rightarrow 0$; i.e., when $|\Psi_{-}\rangle$ is the bright component. In this case, $x_0^{\text{opt}} \rightarrow s_0 - s$ coincides with the center of the bright component. But, this means that $|\Phi_0\rangle = |\Psi_{-}\rangle$ and the two outputs described by $|\pi_0\rangle$ and $|\pi_1\rangle$ project on subspaces orthogonal to the bright component, as anticipated.

In practice, the performance will be compromised by any misalignment with respect to x_0^{opt} . This effect is examined in Fig. 3, where the quantum limit and the direct intensity imaging are compared with different misaligned measurements Eq. (24). Being about two orders of magnitude below the Rayleigh limit, such imperfections cause a loss of precision. Even then, the advantage with respect to direct imaging persists over a wide range of displacements x_0 , demonstrating the robustness of our detection scheme. Again, setting $\phi \approx \pi/2$ seems to be the best option. For this particular example, the

measurement can be misaligned by as much as 0.4σ from x_0^{opt} and still beat the direct imaging limits in measuring separations two orders of magnitude below the Rayleigh limit.

Such an inherent robustness of optimal detection schemes hints at using adaptive strategies to achieve the quantum limits. One plausible way would be to spend a portion of the photon pool to obtain a first estimate of the optimal displacement $\hat{x}_0^{\text{opt}} = \hat{s}_0 - \hat{s}(1 - 2\hat{q})/2$. Since this quantity is closely related to the weighted centroid, direct imaging can be used in this step. Then, the estimated \hat{x}_0^{opt} can be used with the optimal measurement Eq. (24) in the next step to refine the estimates of the signal parameters and so forth.

Having considered the fundamental aspects of the problem, how does one implement the optimal measurement in practice for one particular setting of the displacement? This amounts to performing simultaneous projections on three mutually orthogonal states. There exists a unitary transformation taking this triplet into another set of orthogonal vectors, where the latter set is experimentally feasible. For example, the optimal projections can be mapped on three different pixels of a CCD camera. Unitary transformations of this kind can be always realized with a set of nonabsorbing masks. Alternatively, giving up some performance, the implementation can be facilitated by splitting the signal beam and measuring the three projections separately. This leads to a photon loss and a threefold decrease of the precisions H_α , which can be tolerated for sufficiently small separations.

V. CONCLUSIONS

We have examined the ultimate limits for the simultaneous estimation of centroid, separation, and relative intensities of two incoherent point sources. Our results indicate that the optimal sub-Rayleigh resolution limit can be achieved for any real-valued amplitude PSF provided the system output is projected onto a suitable complete set of modes. Particularly useful modes can be generated from the derivatives of the system PSF, which in the limit of small separations can access all available information with a few projections.

For equally bright sources, our proposed projection is optimal whereas, for unbalanced signals, its performance deteriorates with the parameter q . While some of our findings were illustrated explicitly for Gaussian PSFs, our framework is general and can be applied to other relevant cases.

All in all, this constitutes an important application of multiparameter quantum estimation theory to a more realistic imaging setting. Our analysis provides a toolbox for achieving optimal resolution and paves the way for further experimental demonstrations and innovative solutions in scientific, industrial, and biomedical domains.

ACKNOWLEDGMENTS

We acknowledge financial support from the Grant Agency of the Czech Republic (Grant No. 18-04291S), the Palacký University (Grant No. IGA_PrF_2018_003), the European Space Agency's ARIADNA scheme (Contract No. 4000120780/17/NL/LF/as), and the Spanish MINECO (Grant No. FIS2015-67963-P).

- [1] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
- [2] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*, 2nd ed. (North Holland, Amsterdam, 2003).
- [3] V. Giovannetti, S. Lloyd, and L. Maccone, Advances in quantum metrology, *Nat. Photon.* **5**, 222 (2011).
- [4] R. Demkowicz-Dobrzanski, M. Jarzyna, and J. Kolodynski, Quantum limits in optical interferometry, *Prog. Opt.* **60**, 345 (2015).
- [5] S. L. Braunstein and C. M. Caves, Statistical Distance and the Geometry of Quantum States, *Phys. Rev. Lett.* **72**, 3439 (1994).
- [6] H. Yuen and M. Lax, Multiple parameter quantum estimation and measurement of nonselfadjoint observables, *IEEE Trans. Inf. Theory* **19**, 740 (1973).
- [7] V. P. Belavkin, Generalized uncertainty relations and efficient measurements in quantum systems, *Theor. Math. Phys.* **26**, 213 (1976).
- [8] M. Szczykulska, T. Baumgratz, and A. Datta, Multi-parameter quantum metrology, *Adv. Phys. X* **1**, 621 (2016).
- [9] C. W. Helstrom, Minimum mean-squared error of estimates in quantum statistics, *Phys. Lett. A* **25**, 101 (1967).
- [10] K. Matsumoto, A new approach to the Cramér-Rao-type bound of the pure-state model, *J. Phys. A: Math. Gen.* **35**, 3111 (2002).
- [11] R. D. Gill and M. I. Guta, in *On Asymptotic Quantum Statistical Inference* (Institute of Mathematical Statistics, Beachwood, OH, 2013), p. 105.
- [12] L. Pezzè, M. A. Ciampini, N. Spagnolo, P. C. Humphreys, A. Datta, I. A. Walmsley, M. Barbieri, F. Sciarrino, and A. Smerzi, Optimal Measurements for Simultaneous Quantum Estimation of Multiple Phases, *Phys. Rev. Lett.* **119**, 130504 (2017).
- [13] J.-D. Yue, Y.-R. Zhang, and H. Fan, Quantum-enhanced metrology for multiple phase estimation with noise, *Sci. Rep.* **4**, 5933 (2014).
- [14] M. Zuppardo, J. P. Santos, G. De Chiara, M. Paternostro, F. L. Semião, and G. M. Palma, Cavity-aided quantum parameter estimation in a bosonic double-well Josephson junction, *Phys. Rev. A* **91**, 033631 (2015).
- [15] T. J. Proctor, P. A. Knott, and J. A. Dunningham, Multiparameter Estimation in Networked Quantum Sensors, *Phys. Rev. Lett.* **120**, 080501 (2018).
- [16] A. J. den Dekker and A. van den Bos, Resolution: A survey, *J. Opt. Soc. Am. A* **14**, 547 (1997).
- [17] P. R. Hemmer and T. Zapata, The universal scaling laws that determine the achievable resolution in different schemes for super-resolution imaging, *J. Opt.* **14**, 083002 (2012).
- [18] C. Cremer and R. B. Masters, Resolution enhancement techniques in microscopy, *Eur. Phys. J. H* **38**, 281 (2013).
- [19] Lord Rayleigh, Investigations in optics, with special reference to the spectroscope, *Philos. Mag. J. Sci.* **8**, 261 (1879); **8**, 403 (1879); **8**, 477 (1879).
- [20] L. Schermelleh, R. Heintzmann, and H. Leonhardt, A guide to super-resolution fluorescence microscopy, *J. Cell Biol.* **190**, 165 (2010).
- [21] M. Tsang, R. Nair, and X.-M. Lu, Quantum Theory of Superresolution for Two Incoherent Optical Point Sources, *Phys. Rev. X* **6**, 031033 (2016).
- [22] R. Nair and M. Tsang, Far-Field Superresolution of Thermal Electromagnetic Sources at the Quantum Limit, *Phys. Rev. Lett.* **117**, 190801 (2016).
- [23] S. Z. Ang, R. Nair, and M. Tsang, Quantum limit for two-dimensional resolution of two incoherent optical point sources, *Phys. Rev. A* **95**, 063847 (2017).
- [24] C. Lupo and S. Pirandola, Ultimate Precision Bound of Quantum and Subwavelength Imaging, *Phys. Rev. Lett.* **117**, 190802 (2016).
- [25] J. Rehacek, M. Paúr, B. Stoklasa, Z. Hradil, and L. L. Sánchez-Soto, Optimal measurements for resolution beyond the Rayleigh limit, *Opt. Lett.* **42**, 231 (2017).
- [26] R. Kerviche, S. Guha, and A. Ashok, Fundamental limit of resolving two point sources limited by an arbitrary point spread function, in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)* (IEEE, Aachen, 2017), pp. 441–445.
- [27] M. Tsang, Conservative classical and quantum resolution limits for incoherent imaging, *J. Mod. Opt.* **65**, 1385 (2018).
- [28] S. Zhou and L. Jiang, A modern description of Rayleigh's criterion, [arXiv:1801.02917](https://arxiv.org/abs/1801.02917).
- [29] M. Paur, B. Stoklasa, Z. Hradil, L. L. Sanchez-Soto, and J. Rehacek, Achieving the ultimate optical resolution, *Optica* **3**, 1144 (2016).
- [30] F. Yang, A. Taschilina, E. S. Moiseev, C. Simon, and A. I. Lvovsky, Far-field linear optical superresolution via heterodyne detection in a higher-order local oscillator mode, *Optica* **3**, 1148 (2016).
- [31] W. K. Tham, H. Ferretti, and A. M. Steinberg, Beating Rayleigh's Curse by Imaging Using Phase Information, *Phys. Rev. Lett.* **118**, 070801 (2017).
- [32] F. Yang, R. Nair, M. Tsang, C. Simon, and A. I. Lvovsky, Fisher information for far-field linear optical superresolution via homodyne or heterodyne detection in a higher-order local oscillator mode, *Phys. Rev. A* **96**, 063829 (2017).
- [33] A. Chrostowski, R. Demkowicz-Dobrzański, M. Jarzyna, and K. Banaszek, On superresolution imaging as a multiparameter estimation problem, *Int. J. Quantum Inf.* **15**, 1740005 (2017).
- [34] J. Rehacek, Z. Hradil, B. Stoklasa, M. Paúr, J. Grover, A. Krzic, and L. L. Sánchez-Soto, Multiparameter quantum metrology of incoherent point sources: Towards realistic superresolution, *Phys. Rev. A* **96**, 062107 (2017).
- [35] D. Petz and C. Ghinea, Introduction to quantum Fisher information, in *Quantum Probability and Related Topics* (World Scientific, Singapore, 2011), Vol. 27, pp. 261–281.
- [36] J. Suzuki, Explicit formula for the Holevo bound for two-parameter qubit-state estimation problem, *J. Math. Phys.* **57**, 042201 (2016).
- [37] C. Vaneph, T. Tufarelli, and M. G. Genoni, Quantum estimation of a two-phase spin rotation, *Quantum Meas. Quantum Metrol.* **1**, 12 (2013).
- [38] P. J. D. Crowley, A. Datta, M. Barbieri, and I. A. Walmsley, Tradeoff in simultaneous quantum-limited phase and loss estimation in interferometry, *Phys. Rev. A* **89**, 023845 (2014).
- [39] S. Ragy, M. Jarzyna, and R. Demkowicz-Dobrzański, Compatibility in multiparameter quantum metrology, *Phys. Rev. A* **94**, 052108 (2016).
- [40] J. M. Bernardo and A. F. M. Smith, *Bayesian Theory* (Wiley, Sussex, 2000).

Chapter 5

Axial superlocalization with vortex beams

In this work, I carried out theoretical calculations and numerical simulations presented in the paper. I was also involved in the preparation of the manuscript.

This chapter presents a follow-up on the previous work [26], where we studied the superresolution of the axial distance considering Gaussian beams. Here, we introduced the higher-order beams, namely Laguerre-Gaussian vortexes, and analyzed their ability to recover information about the axial distance [2].

Achieving high precision in axial resolution is crucial for three-dimensional (3D) optical imaging systems. This work investigates the ultimate precision in axial localization using vortex beams. It focuses on Laguerre-Gauss (LG) beams and the potential improvements in microscopy methods by replacing traditional intensity sensors with advanced mode-sorting techniques.

The complex amplitude of the LG beam reads

$$\begin{aligned} \Psi(r, \phi, z) = \langle r, \phi, z | p, l \rangle = & \sqrt{\frac{2p!}{\pi(p+|l|)!}} \frac{1}{w(z)} \left(\frac{\sqrt{2}r}{w(z)} \right)^{|l|} \\ & \times L_p^{|l|} \left(\frac{2r^2}{w(z)^2} \right) \exp \left(-r^2 \left[\frac{1}{w(z)^2} - i \frac{k}{2R(z)} - il\phi - i\psi_{pl}(z) \right] \right), \end{aligned} \quad (5.1)$$

where (r, ϕ, z) are cylindrical coordinates, $L_p^{|l|}(\cdot)$ is the generalized Laguerre polynomial and (l, p) are azimuthal and radial mode indexes, respectively. We showed that QFI for pure LG modes reads

$$Q = \frac{2p(p+|l|) + 2p + |l| + 1}{z_R^2}. \quad (5.2)$$

We should note that in the case of the Gaussian mode LG_{00} , the QFI reduces to $\frac{1}{z_R^2}$, which coincides with our previous result on axial resolution from propagating Gaussian beam. The QFI (5.2) is also linear in $|l|$, which means we can improve axial resolution using an LG beam with a higher azimuthal index.

Interesting rotating behavior arises from superposing two LG beams (5.1) with different azimuthal indexes l and l' ,

$$|\Psi_{ll'}\rangle = \frac{1}{\sqrt{2}}(|\text{LG}_{l0}\rangle + |\text{LG}_{0l'}\rangle). \quad (5.3)$$

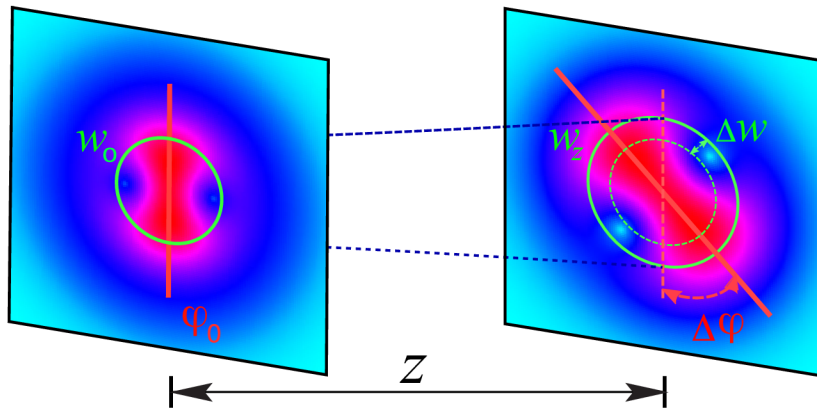


Figure 5.1: Sketch of the evolution of a superposition of two LG modes with azimuthal indexes $l, l' = 0, 2$. During propagation, the intensity pattern rotates and diverges.

The intensity pattern is sketched in the Fig. 5.1. QFI for such a superposed state reads

$$\mathcal{Q} = \frac{1}{z_R^2} [4 + 2(|l| + |l'|) + (|l| - |l'|)^2]. \quad (5.4)$$

If the second component is left in the fundamental mode $l' = 0$, the QFI is maximal for the maximum available azimuthal index l , and the bound becomes quadratic. Even though the theory suggests that we should use higher-order superpositions for axial displacement estimation, as we will see, we cannot access most information about the propagation distance just from simple intensity detection. Considering single parameter estimation, one can always find a von Neumann measurement projecting the signal onto eigenstates of the symmetric logarithmic derivative (5.1). However, such measurements can be impractical for implementation. Therefore, we analyzed the information we can gain about axial distance from simple intensity measurement. Since the detection can be considered as a random process, in consequence, normalized beam intensity can be viewed as a probability distribution

$$p(r, \phi|z) = |\Psi(r, \phi|z)|^2. \quad (5.5)$$

As the detection event follows Poissonian statistics, the classical Fisher information (3.9) reads

$$\mathcal{F} = \int_0^\infty \int_0^{2\pi} \frac{[\partial_z p(r, \phi|z)]^2}{p(r, \phi|z)} r dr d\phi. \quad (5.6)$$

Evaluation of the integral (5.6) for pure LG beams shows that we can indeed saturate the QFI (5.2),

$$\mathcal{F}_{\text{opt}} = Q, \quad (5.7)$$

and that the optimal detection plane is located at $z_{\text{opt}} = \pm z_R$. However, numerical analysis suggests that this result does not hold for superpositions of LG beams. As we can see from Fig. 5.2, where we plotted the ratio of QFI and numerically calculated FI for different superpositions, the optimal detection plane is no longer at $z = z_R$ and that the larger the azimuthal index l is, the smaller portion of the total QFI can be extracted from intensity measurement.

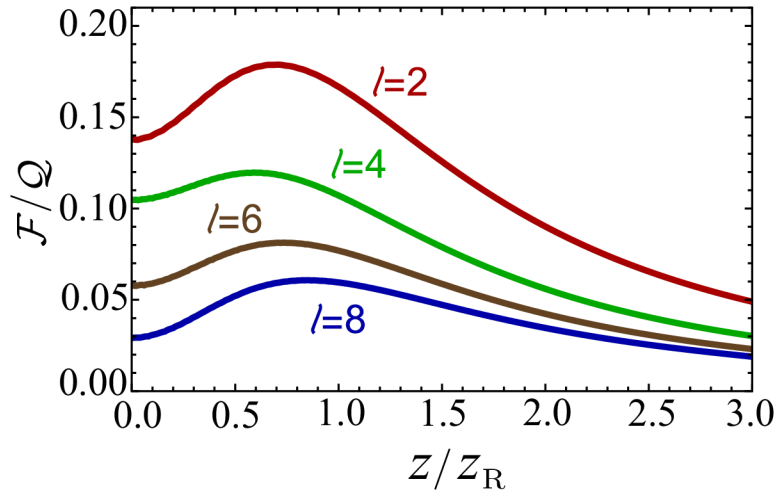


Figure 5.2: Fisher information normalized by a quantum Fisher information for different superposition of LG modes as a function of position z of the detection plane. Maximum information that we can extract from one intensity plane is no longer at the distance $z = z_R$.

In conclusion, this work establishes the quantum limits for axial localization using vortex beams, showing that pure LG beams can achieve these limits with optimal intensity detection. For superpositions of LG beams, the benefits of rotating intensity profiles suggest that microscopy methods could significantly improve by adopting advanced mode-sorting techniques. This work underscores the inadequacy of simple intensity detection in capturing the full information about axial displacement. Instead, it advocates for using spatial-mode projections derived from quantum detection theory, unlocking the potential for axial superlocalization in 3D imaging.

Quantum Science and Technology



PAPER

Axial superlocalization with vortex beams

OPEN ACCESS

RECEIVED
19 November 2020

REVISED
6 February 2021

ACCEPTED FOR PUBLICATION
22 February 2021

PUBLISHED
31 March 2021

Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



D Koutný¹, Z Hradil¹, J Řeháček¹ and L L Sánchez-Soto^{2,3,*}

¹ Department of Optics, Palacký University, 17. listopadu 12, 77146 Olomouc, Czech Republic

² Departamento de Óptica, Facultad de Física, Universidad Complutense, 28040 Madrid, Spain

³ Max-Planck-Institut für die Physik des Lichts, Staudtstraße 2, 91258 Erlangen, Germany

* Author to whom any correspondence should be addressed.

E-mail: lsanchez@fis.ucm.es

Keywords: quantum resolution limits, quantum Fisher information, orbital angular momentum

Abstract

Improving axial resolution is paramount importance for three-dimensional optical imaging systems. Here, we investigate the ultimate precision in axial localization using vortex beams. For Laguerre–Gauss (LG) beams, this limit can be achieved with just an intensity scan. The same is not true for superpositions of LG beams, in particular for those with intensity profiles that rotate on defocusing. Microscopy methods based on rotating vortex beams may thus benefit from replacing traditional intensity sensors with advanced mode-sorting techniques.

1. Introduction

Discerning the nanometer-scale details of living cells, tissues, and materials is of utmost importance for many modern research efforts. A trail toward to this holy grail was opened with the appearance of a set of methods, dubbed with the generic term of superresolution microscopy [1, 2], capable of bypassing the diffraction limit [3–5]: a barrier that was traditionally thought to be impenetrable.

A number of these techniques can also reveal three-dimensional (3D) structural details: relevant examples include stimulated-emission-depletion microscopy [6], PSF engineering [8–12], photoactivated-localization microscopy [7], and multiplane detection [13–15], to cite only but a few. All of them rely on a very accurate localization of point sources; they differ in how point objects are excited and how the corresponding emitted photons are collected.

For 3D imaging, the emitter is fluorescently labeled and determining its axial position is an indispensable ingredient. This problem has been thoroughly examined and some impressive results have been demonstrated so far [16]. However, the fundamental depth precision achievable by any such engineering method has been considered only lately [17–19]. The rationale behind is to make a systematic use of the quantum Fisher information (QFI) [20] and the associated quantum Cramér–Rao bound (QCRB) to get a measurement-independent limit [21, 22]. This is much along the lines of the work of Tsang and co-workers to quantify the transverse two-point resolution [23–27], which has led to the dispelling of the Rayleigh curse [28–31].

In a recent work [32], the ultimate precision in axial localization using Gaussian beams has been established. This limit can be attained with just one intensity scan, as long as the detection plane is placed at one optimal position.

In this paper, we generalize those results and derive quantum limits for axial localization with Laguerre–Gauss (LG) beams, which carry quantized orbital angular momentum [33]. Here, the beam waist acts as a realization of the light emitted by a point source after e.g. mode conversion. Another relevant situation is the reflection of the beam from a surface in surface topology measurements, etc. By linearly superposing different LG modes, one can realize beams with amplitude, phase, and intensity patterns that simply rotate, under free space propagation, maintaining the transverse shape. These rotating structures lie at the core of a variety of sensing techniques [34–37].

We demonstrate that a meager part of the full (quantum) information is available in intensity scans and only a small fraction of this can be attributed to the rotation. This clearly confirms the potential of modal

expansions inspired by quantum information protocols [38], which allow for reaching the QCRB and thus can be considered as the optimal measurement. Our results make 3D superresolution imaging more feasible and potentially useful for improving the resolution of optical microscopes.

2. Theoretical model

The problem we address here is to estimate the distance traveled by a vortex beam from the beam waist to an arbitrary detection plane. We thus consider the beam waist as an object whose axial distance is to be determined. In what follows, we shall represent the fields using Dirac notation, for it makes it straightforward to expand the theory to other types of light states.

We take the beam to be represented by the pure state $|\Psi(0)\rangle$, where $z = 0$ denotes the position of the object plane. The axial displacement is thus characterized by a unitary operation

$$|\Psi(z)\rangle = e^{iGz}|\Psi(0)\rangle, \quad (2.1)$$

where the Hermitian operator G is the generator. To pinpoint the specific form of G , it is appropriate to use the transverse-position representation $\Psi(x, y; z) = \langle x, y | \Psi(z) \rangle$. It then follows directly from equation (2.1) that

$$i\partial_z \Psi(x, y; z) = -G\Psi(x, y; z). \quad (2.2)$$

On the other hand, vortex beams are solutions of the paraxial Helmholtz equation [39, 40]

$$2ik\partial_z \Psi(x, y, z) = \nabla_T^2 \Psi(x, y, z), \quad (2.3)$$

where k is the wavenumber and $\nabla_T^2 = \partial_{xx} + \partial_{yy}$ is the transverse Laplacian. A direct comparison leads us to

$$G \mapsto -\frac{1}{2k} \nabla_T^2. \quad (2.4)$$

The detection plane is placed at z , wherein we perform an arbitrary measurement. Given the formal analogies between spatial modes in wave optics and pure states in quantum theory, and also the mathematical similarities in describing evolution and detection of such objects, the amount of information about the axial distance z carried by the measured signal is quantified by the QFI. For pure states, as it is our case, the QFI reduces to [21]

$$\mathcal{Q}(z) = 4 \text{Var}(G). \quad (2.5)$$

Except for the factor 4, the QFI is the variance of the generator G computed in the initial state $|\Psi(0)\rangle$. According to the time-honored QCRB, the variance of any unbiased estimator \hat{z} of the axial distance z satisfies

$$\text{Var}(\hat{z}) \geq \frac{1}{\mathcal{Q}(z)}, \quad (2.6)$$

whose saturation provides the ultimate precision in axial distance estimation.

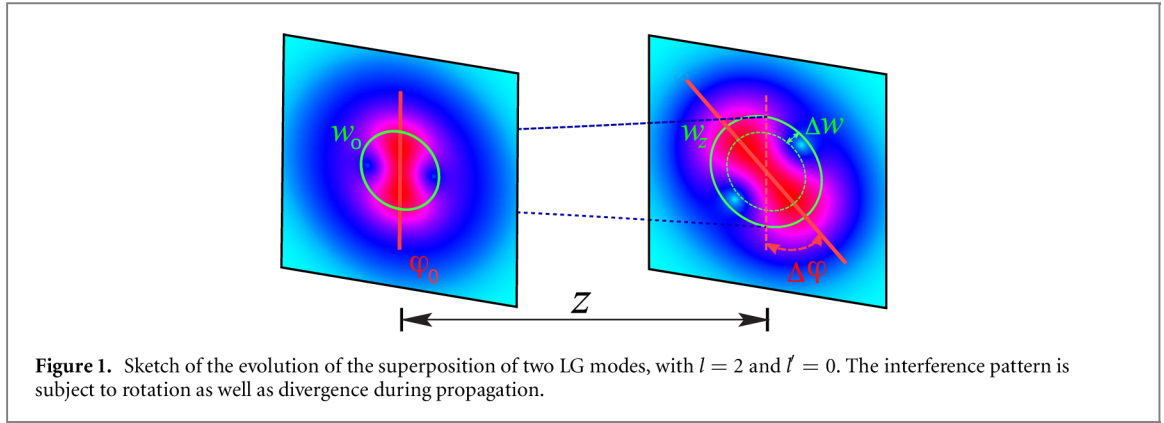
To proceed further, we take the structure of the transverse field to correspond with LG modes

$$\begin{aligned} \text{LG}_{pl}(r, \phi, z) = \langle r, \phi, z | p, l \rangle &= \sqrt{\frac{2p!}{\pi(p+|l|)!}} \frac{1}{w(z)} \left[\frac{\sqrt{2}r}{w(z)} \right]^{|l|} \\ &\times L_p^{|l|} \left(\frac{2r^2}{w(z)^2} \right) e^{-\frac{r^2}{w(z)^2}} \exp \left(i \left[\frac{kr^2}{2R(z)} - l\phi - \psi_{pl}(z) \right] \right), \end{aligned} \quad (2.7)$$

where (r, ϕ, z) are cylindrical coordinates, $L_p^{|l|}(\cdot)$ is the generalized Laguerre polynomial, $l \in \{0, \pm 1, \pm 2, \dots\}$ is the azimuthal mode index and $p \in \{0, 1, 2, \dots\}$ is the radial index. The parameters $R(z)$, $w(z)$, and $\psi_{pl}(z)$ are

$$\begin{aligned} R(z) &= z \left[1 + (z/z_R)^2 \right], \\ w^2(z) &= w_0^2 \left[1 + (z/z_R)^2 \right], \\ \psi_{pl}(z) &= (2p + |l| + 1) \arctan(z/z_R), \end{aligned} \quad (2.8)$$

and represent the radius of curvature of the wave front, the beam radius, and the Gouy phase [41], respectively, at an axial distance z from the beam waist located at $z = 0$. Here, $z_R = kw_0^2/2$ is the Rayleigh length and w_0 the beam waist radius [42].



3. Quantum limit for axial localization with vortex beams

To facilitate the derivation of the QFI corresponding to axial displacements of vortex beams it is advantageous to use an established correspondence between eigenstates of a two-dimensional harmonic oscillator and paraxial beams [43]. For the case of LG beams, one defines

$$a_{\pm} = \frac{1}{\sqrt{2}}(a_{\xi} \mp ia_{\eta}), \quad (3.1)$$

where a_{ξ} and a_{η} are dimensionless bosonic operators for each independent amplitude of the oscillator and obey the standard commutation relations $[a_s, a_{s'}^{\dagger}] = \delta_{ss'}$ ($s, s' \in \{\xi, \eta\}$). Similar relations are obeyed by a_{\pm} .

The eigenstates of the harmonic oscillator $|n_+, n_-\rangle$, generated by the action of a_+^{\dagger} and a_-^{\dagger} on the vacuum, are nothing but LG modes with azimuthal and radial indices given by

$$l = n_+ - n_-, \quad p = \min(n_+, n_-). \quad (3.2)$$

Applying the usual definition of the momentum operator $p_s = \frac{1}{\sqrt{2i}}(a_s - a_s^{\dagger})$, we have

$$\tilde{\nabla}_T^2 = p_{\xi}^2 + p_{\eta}^2 = (p_{\xi} + ip_{\eta})(p_{\xi} - ip_{\eta}) = (a_+ - a_-^{\dagger})(a_+^{\dagger} - a_-). \quad (3.3)$$

If we recall (2.4), and take into account that, in the proper units, $(x, y) \mapsto (\sqrt{2}\xi/w_0, \sqrt{2}\eta/w_0)$, we get $\nabla_T^2 \mapsto 2\tilde{\nabla}_T^2/w_0^2$. In this way, the QFI of a pure LG mode $|n_+, n_-\rangle$ reads:

$$\mathcal{Q}(z) = \frac{1}{z_R^2}(2n_+n_- + n_+ + n_- + 1) = \frac{1}{z_R^2}[2p(p + |l|) + 2p + |l| + 1]. \quad (3.4)$$

For the particular case of the Gaussian mode LG_{00} we get

$$\mathcal{Q}(z) = \frac{1}{z_R^2}, \quad (3.5)$$

that is, the quantum bound (per single detection) is precisely the Rayleigh range [44]. Note that the QFI is *linear* in $|l|$, which means that axial localization can be improved by using LG beams with large OAM.

Other sets of Gaussian transverse modes can be characterized using the so-called Hermite–Laguerre sphere [45, 46]. These modes are represented by a point, of spherical coordinates (θ, ϕ) , on that sphere, and they are generated by the rotated operators

$$\begin{aligned} a_1(\theta, \phi) &= a_+ e^{-i\phi/2} \cos\left(\frac{\theta}{2}\right) + a_- e^{i\phi/2} \sin\left(\frac{\theta}{2}\right), \\ a_2(\theta, \phi) &= -a_+ e^{-i\phi/2} \sin\left(\frac{\theta}{2}\right) + a_- e^{i\phi/2} \cos\left(\frac{\theta}{2}\right). \end{aligned} \quad (3.6)$$

In particular, $\theta = \pi/2$ gives rise to Hermite–Gauss (HG) modes, and $\theta = 0, \pi$ to LG modes analyzed above. Combining equations (3.3) and (3.6), we get a direct generalization of (3.4); viz,

$$\mathcal{Q} = \frac{1}{z_R^2}[4 + n_1 + n_2(3 + n_2) + n_1(3 + 4n_2) + (n_1 - n_1^2 + n_2 + 4n_1n_2 - n_2^2) \cos(2\theta)], \quad (3.7)$$

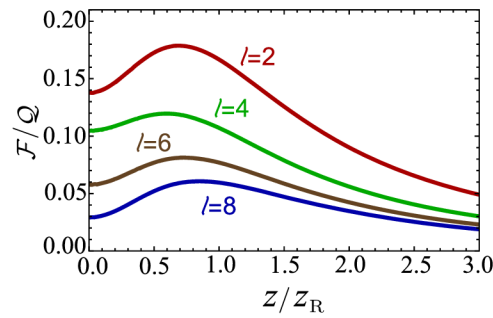


Figure 2. Fisher information about axial distance (relative to the QFI) for different superpositions $(|LG_{0l}\rangle + |LG_{00}\rangle)/\sqrt{2}$ as a function of the position z (relative to the beam waist) of the detection plane.

where n_1 and n_2 denote the eigenvalues of the corresponding number operators. Notice that the QFI (3.7) is independent of ϕ and is optimized by (3.4). This proves that LG modes are better than their HG counterparts, as axial localization is concerned.

Apart from a trivial divergence, transversal intensity profiles of pure vortex beams do not change on propagation. More complex intensity transformations can be realized by superposing two or more vortex beams. In particular, rotating structures are of interest in microscopy [34–37]. Rotation of highly symmetric spots is easily detected and the corresponding defocusing parameter can be estimated from the measured rotation angle. This is roughly sketched in figure 1.

To simplify the details as much as possible, we take the simple example of the superposition of two LG modes with different azimuthal numbers $l \neq l'$, and $p = p' = 0$

$$|\Psi_{ll'}\rangle = \frac{1}{\sqrt{2}}(|LG_{0l}\rangle + |LG_{0l'}\rangle) = \frac{1}{\sqrt{2}}(|n_+, 0\rangle + |n'_+, 0\rangle). \quad (3.8)$$

In this case, the QFI about the axial position of the source reads

$$\mathcal{Q} = \frac{1}{z_R^2} [4 + 2(|l| + |l'|) + (|l| - |l'|)^2]. \quad (3.9)$$

Considering OAM as a resource for axial localization and using the maximum available OAM for one of the components of the superposition, $l = \pm |l_{\max}|$, the QFI is maximized when the second component is in the fundamental mode $l' = 0$, whereupon $\mathcal{Q}_{\max} = [4 + 2|l_{\max}| + |l_{\max}|^2]/z_R^2$, and the quantum bound becomes *quadratic* in $|l|$. At first sight, the better performance of vortex superpositions (3.9) over pure vortex beams (3.4) is somehow related to beam rotation. This might suggest a proxy for the measured axial distance traveled from the waist to the detection plane. However, as we show in the next section, things are not that simple.

4. Intensity detection

In single-parameter estimation, the QFI can always be accessed and the corresponding QCRB saturated with a von Neumann measurement projecting the measured signal on the eigenstates of the symmetric logarithmic derivative of the density matrix [21]. In our context, the practical implementation of such measurements requires a spatial mode demultiplexer/sorter that performs simultaneous projection of the signal onto a complete orthonormal set of modes [47, 48]. Mode separation is usually achieved with a sequence of spatial light modulators implementing a suitable unitary transformation. Unavoidable systematic errors and losses introduced by such complicated experimental setups may ruin any theoretical advantage offered by optimal strategies.

In consequence, we consider the performance of the possibly inferior, but much more robust intensity detection, because it is the simplest method at hand for the experimentalist. As the information (3.4) and (3.7) about the axial distance is carried by both intensity and phase of the measured beam, we might aptly ask how much information is sacrificed by completely ignoring the phase.

As usual, due to the noise, the detection can be considered as a random process. In consequence, the (normalized) beam intensity,

$$p(r, \phi|z) = |\Psi(r, \phi; z)|^2, \quad (4.1)$$

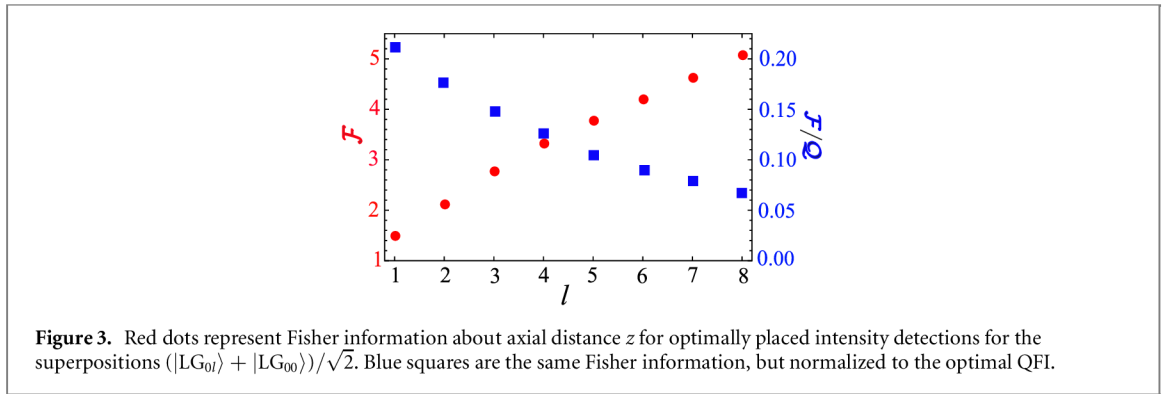


Figure 3. Red dots represent Fisher information about axial distance z for optimally placed intensity detections for the superpositions $(|LG_{0l}\rangle + |LG_{00}\rangle)/\sqrt{2}$. Blue squares are the same Fisher information, but normalized to the optimal QFI.

can be seen as the probability density of a detection conditional on the axial distance z . We take the detection as dominated by shot noise, which obeys a Poisson distribution [49]: although this neglects nonclassical effects, it is still a suitable model for realistic microscopy.

The classical Fisher information about z , per single detection, thus reads [50]

$$\mathcal{F} = \int_0^\infty \int_0^{2\pi} \frac{[\partial_z p(r, \phi|z)]^2}{p(r, \phi|z)} r dr d\phi, \quad (4.2)$$

and it is a suitable tool to quantify the information content about axial displacements accessible from the detected transversal intensity profile. For simplicity, we take the pixel size negligibly small, so that any sampling effect can be ignored. We also define the radial and azimuthal Fisher informations

$$\begin{aligned} \mathcal{F}_r &= \int_0^\infty \frac{\left[\partial_z \int_0^{2\pi} p(r, \phi|z) d\phi \right]^2}{\int_0^{2\pi} p(r, \phi|z) d\phi} r dr, \\ \mathcal{F}_\phi &= \int_0^{2\pi} \frac{\left[\partial_z \int_0^\infty p(r, \phi|z) r dr \right]^2}{\int_0^\infty p(r, \phi|z) r dr} d\phi, \end{aligned} \quad (4.3)$$

respectively. They quantify the information of the radial and the azimuthal intensity components due to the z -dependence. For example, beam rotation only contributes to \mathcal{F}_ϕ , whereas beam divergence contributes to \mathcal{F}_r .

For pure $|LG_{pl}\rangle$ modes, $\Psi(r, \phi; z) = LG_{pl}(r, \phi, z)$, the classical Fisher information can be obtained in a closed form. First notice the integrand of (4.2) is independent of ϕ . Carrying out z -derivatives of Laguerre polynomials using the relation $\partial_z L_p^l(z) = -L_{p-1}^{l+1}(z)$, and changing the integration variable r yields

$$\mathcal{F} = \frac{4p!}{(|l|+p)!} \left[\frac{\partial_z w(z)}{w(z)} \right]^2 \int_0^\infty e^{-t} t^{|l|} \left[2t L_{p-1}^{|l|+1}(t) + (t - |l| - 1) L_p^{|l|}(t) \right]^2 dt. \quad (4.4)$$

We are thus left with evaluating six different integrals involving products of Laguerre polynomials. Compact expressions for each case can be obtained from the generalized orthogonality relation [51]

$$\int_0^\infty e^{-t} t^\mu L_p^l(t) L_{p'}^{l'}(t) dt = (-1)^{p+p'} \Gamma(\mu+1) \sum_{k=0}^{\min(p,p')} \binom{\mu-l}{p-k} \binom{\mu-l'}{p'-k} \binom{\mu+k}{k}. \quad (4.5)$$

Straightforward simplifications leave us with the final result

$$\mathcal{F} = 4[2p(p+|l|) + 2p+|l|+1] \left[\frac{\partial_z w(z)}{w(z)} \right]^2 = \frac{2p(p+|l|) + 2p+|l|+1}{R(z)^2/4}. \quad (4.6)$$

For the planes of maximal wavefront curvature $z_{\text{opt}} = \pm z_R$, where $R(z_R)^2 = 4z_R^2$, the quantum limit is saturated with intensity sensitive detection

$$\mathcal{F}_{\text{opt}} = \mathcal{Q}. \quad (4.7)$$

Thus, for any pure LG beam and any beam waist location, two detection planes can be found, where complete information about the axial distance can be extracted with intensity-only detection. Hence, full potential of high-order vortex beams for axial metrology can be exploited with direct detection techniques. This generalizes the results obtained for Gaussian beams [32].

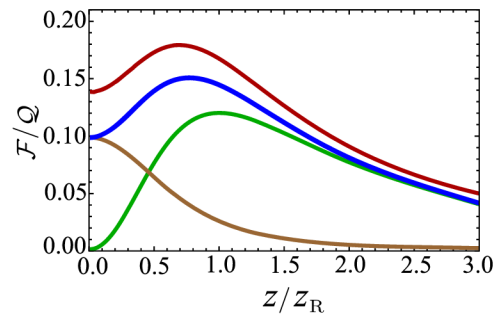


Figure 4. Radial (green), azimuthal (brown) and total (red) Fisher information \mathcal{F} about axial displacement as a function of detector axial position z (in units of z_R) for the vortex superposition $(|LG_{02}\rangle + |LG_{00}\rangle)/\sqrt{2}$. The sum of radial and azimuthal components is also shown (blue). All quantities are normalized to the corresponding optimal QFI. Notice that for small (large) propagation distances, most FI comes from azimuthal (radial) intensity distribution. Some information is stored in correlations between those two marginal distributions, as can be seen from the gap between the blue and the red curves.

However, numerical analysis suggests that this result does not hold for superpositions of LG modes: a single transversal intensity scan is no longer optimal for any detector position z . In figure 2 we plot the classical Fisher information (normalized to the optimal QFI) for different vortex superpositions. First, notice that the optimal detector position is no longer at $z = \pm z_R$, as it was for pure LG beams. Second, the larger the angular momentum l carried by the signal, the smaller portion of the total QFI can be extracted with intensity measurements.

This is further illustrated in figure 3. Locating intensity detectors at the optimal detection planes for every value of l , the classical Fisher information grows sublinearly with l , so that progressively smaller portion of quadratic QFI is available from intensity data and the gap between the quantum bound and the performance of the best estimator from intensity data worsens with l .

For example, for $l = 2$, only about 17% of the QFI is available from intensity scan at the optimal detector position, as we can see in figure 4. Of this, only a small fraction is due to changes in the azimuthal intensity profile, as arises, e.g., in rotations, except for very small propagation distances, where azimuthal profile accounts for up to 10% of QFI. This ratio quickly approaches zero with increasing l .

The information content of rotating beam intensity distributions is surprisingly low. As bad as it sounds, this is not a negative result, but encouraging news for quantum metrology. It highlights inadequacy of simple detection techniques in this particular metrological scenario, where intensity detection fails to reveal all potentially accessible information about the parameter of interest. The hidden Information has to be accessed with advanced detection techniques. Here a generic tool is projecting/sorting the signal into optimal set of spatial modes derivable from the quantum detection theory. Hence, we uncover a huge potential of axial superlocalization based on spatial-mode projections applied to higher order vortex beam superpositions [44].

5. Concluding remarks

We have established the ultimate quantum limits for axial localization using vortex beams. For pure LG beams, this limit is attained with an intensity scan with the detector located at one of two optimal planes. For superpositions of LG beams, in particular of those with intensity profiles rotating on defocusing, this is no longer true. This means that microscopy methods based on rotating vortex beams may benefit from replacing traditional intensity scans with advanced mode-sorting techniques.

Data availability statement

No new data were created or analysed in this study.

Acknowledgments

We thank Robert W Boyd and Aaron Z Goldberg for helpful discussions. We acknowledge financial support of the project ApresSF, under the QuantERA programme, which has received funding from the European Union's Horizon 2020 research and innovation programme, the program H2020 (project StormyTune),

Grant Agency of the Czech Republic (Grant No. 18-04291S), the Palacký University (Grant No. IGA_PrF_2020_004), and the Spanish Ministerio de Ciencia e Innovación (Grant PGC2018-099183-B-I00).

ORCID iDs

D Koutný  <https://orcid.org/0000-0003-1558-833X>

Z Hradil  <https://orcid.org/0000-0001-8732-4638>

J Řeháček  <https://orcid.org/0000-0002-4969-2760>

L L Sánchez-Soto  <https://orcid.org/0000-0002-7441-8632>

References

- [1] 2009 Focus issue: super-resolution imaging *Nat. Photonics* **3** 361–420
- [2] Schermelleh L, Ferrand A, Huser T, Eggeling C, Sauer M, Biehlaier O and Drummen G P C 2019 Super-resolution microscopy demystified *Nat. Cell Biol.* **21** 72–84
- [3] Abbe E 1873 Ueber einen neuen Beleuchtungsapparat am Mikroskop *Arch. Mikrosk. Anat.* **9** 469–80
- [4] Rayleigh L 1879 XXXI. Investigations in optics, with special reference to the spectroscope *London, Edinburgh Dublin Phil. Mag. J. Sci.* **8** 261–74
- [5] Rayleigh L 1879 XLVI. Investigations in optics, with special reference to the spectroscope *London, Edinburgh Dublin Phil. Mag. J. Sci.* **8** 403–11
- [6] Rayleigh L 1879 LVI. Investigations in optics, with special reference to the spectroscope *London, Edinburgh Dublin Phil. Mag. J. Sci.* **8** 477–86
- [7] Ram S, Ward E S and Ober R J 2006 Beyond Rayleigh's criterion: a resolution measure with application to single-molecule microscopy *Proc. Natl Acad. Sci.* **103** 4457–62
- [8] Hell S W and Wichmann J 1994 Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy *Opt. Lett.* **19** 780–2
- [9] Betzig E, Patterson G H, Sougrat R, Lindwasser O W, Olenych S, Bonifacino J S, Davidson M W, Lippincott-Schwartz J and Hess H F 2006 Imaging intracellular fluorescent proteins at nanometer resolution *Science* **313** 1642
- [10] Huang B, Wang W, Bates M and Zhuang X 2008 Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy *Science* **319** 810
- [11] Pavani S R P, Thompson M A, Biteen J S, Lord S J, Liu N, Twieg R J, Piestun R and Moerner W E 2009 Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function *Proc. Natl Acad. Sci.* **106** 2995
- [12] Jia S, Vaughan J C and Zhuang X 2014 Isotropic three-dimensional super-resolution imaging with a self-bending point spread function *Nat. Photon.* **8** 302
- [13] Tamburini F, Anzolin G, Umbriaco G, Bianchini A and Barbieri C 2006 Overcoming the Rayleigh criterion limit with optical vortices *Phys. Rev. Lett.* **97** 163903
- [14] Paúr M, Stoklasa B, Grover J, Krzic A, Sánchez-Soto L L, Hradil Z and Řeháček J 2018 Tempering Rayleigh's curse with PSF shaping *Optica* **5** 1177–80
- [15] Juette M F, Gould T J, Lessard M D, Mlodzianoski M J, Nagpure B S, Bennett B T, Hess S T and Bewersdorff J 2008 Three-dimensional sub-100 nm resolution fluorescence microscopy of thick samples *Nat. Methods* **5** 527
- [16] Dalgarno P A, Dalgarno H I C, Putoud A, Lambert R, Paterson L, Logan D C, Towers D P, Warburton R J and Greenaway A H 2010 Multiplane imaging and three dimensional nanoscale particle tracking in biological microscopy *Opt. Express* **18** 877–84
- [17] Abrahamsson S et al 2012 Fast multicolor 3D imaging using aberration-corrected multifocus microscopy *Nat. Methods* **10** 60–3
- [18] von Diezmann A, Shechtman Y and Moerner W E 2017 Three-dimensional localization of single molecules for super-resolution imaging and single-particle tracking *Chem. Rev.* **117** 7244–75
- [19] Tsang M 2015 Quantum limits to optical point-source localization *Optica* **2** 646–53
- [20] Nair R and Tsang M 2016 Interferometric superlocalization of two incoherent optical point sources *Opt. Express* **24** 3684–701
- [21] Backlund M P, Shechtman Y and Walsworth R L 2018 Fundamental precision bounds for three-dimensional optical localization microscopy with Poisson statistics *Phys. Rev. Lett.* **121** 023904
- [22] Petz D and Ghinea C 2011 *Introduction to Quantum Fisher Information* vol 27 (Singapore: World Scientific) pp 261–81
- [23] Helstrom C W 1976 *Quantum Detection and Estimation Theory* (New York: Academic)
- [24] Holevo A S 2003 *Probabilistic and Statistical Aspects of Quantum Theory* 2nd edn (Amsterdam: North-Holland)
- [25] Tsang M, Nair R and Lu X-M 2016 Quantum theory of superresolution for two incoherent optical point sources *Phys. Rev. X* **6** 031033
- [26] Lupo C and Pirandola S 2016 Ultimate precision bound of quantum and subwavelength imaging *Phys. Rev. Lett.* **117** 190802
- [27] Nair R and Tsang M 2016 Far-field superresolution of thermal electromagnetic sources at the quantum limit *Phys. Rev. Lett.* **117** 190801
- [28] Ang S Z, Nair R and Tsang M 2016 Quantum limit for two-dimensional resolution of two incoherent optical point sources *Phys. Rev. A* **95** 063847
- [29] Tsang M 2017 Subdiffraction incoherent optical imaging via spatial-mode demultiplexing *New J. Phys.* **19** 023054
- [30] Paúr M, Stoklasa B, Hradil Z, Sánchez-Soto L L and Řeháček J 2016 Achieving the ultimate optical resolution *Optica* **3** 1144–7
- [31] Yang F, Tashchilina A, Moiseev E S, Simon C and Lvovsky A I 2016 Far-field linear optical superresolution via heterodyne detection in a higher-order local oscillator mode *Optica* **3** 1148–52
- [32] Tham W K, Ferretti H and Steinberg A M 2016 Beating Rayleigh's curse by imaging using phase information *Phys. Rev. Lett.* **118** 070801
- [33] Yang F, Nair R, Tsang M, Simon C and Lvovsky A I 2017 Fisher information for far-field linear optical superresolution via homodyne or heterodyne detection in a higher-order local oscillator mode *Phys. Rev. A* **96** 063829
- [34] Řeháček J, Paúr M, Stoklasa B, Koutný D, Hradil Z and Sánchez-Soto L L 2019 Intensity-based axial localization at the quantum limit *Phys. Rev. Lett.* **123** 193601

- [33] Allen L, Beijersbergen M W, Spreeuw R J C and Woerdman J P 1992 Orbital angular momentum of light and the transformation of Laguerre–Gaussian laser modes *Phys. Rev. A* **45** 8185–9
- [34] Piestun R, Schechner Y Y and Shamir J 2000 Propagation-invariant wave fields with finite energy *J. Opt. Soc. Am. A* **17** 294–303
- [35] Greengard A, Schechner Y Y and Piestun R 2006 Depth from diffracted rotation *Opt. Lett.* **31** 181–3
- [36] Prasad S 2013 Rotating point spread function via pupil-phase engineering *Opt. Lett.* **38** 585–7
- [37] Baránek M, Bouchal P, Šiler M and Bouchal Z 2015 Aberration resistant axial localization using a self-imaging of vortices *Opt. Express* **23** 15316–31
- [38] Fabre C and Treps N 2020 Modes and states in quantum optics *Rev. Mod. Phys.* **92** 035005
- [39] Franke-Arnold S, Allen L and Padgett M 2008 Advances in optical angular momentum *Laser Photonics Rev.* **2** 299–313
- [40] Barnett S M, Babiker M and Padgett M J 2017 Optical orbital angular momentum *Phil. Trans. R. Soc. A* **375** 20150444
- [41] Feng S and Winful H G 2001 Physical origin of the Gouy phase shift *Opt. Lett.* **26** 485–7
- [42] Siegman A E 1986 *Lasers* (Oxford: Oxford University Press)
- [43] Nienhuis G 2017 Analogies between optical and quantum mechanical angular momentum *Phil. Trans. R. Soc. A* **375** 20150443
- [44] Zhou Y, Yang J, Hassett J D, Rafsanjani S M H, Mirhosseini M, Vamivakas A N, Jordan A N, Shi Z and Boyd R W 2019 Quantum-limited estimation of the axial separation of two incoherent point sources *Optica* **6** 534–41
- [45] Simon R and Agarwal G S 2000 Wigner representation of Laguerre–Gaussian beams *Opt. Lett.* **25** 1313–5
- [46] Visser J and Nienhuis G 2004 Orbital angular momentum of general astigmatic modes *Phys. Rev. A* **70** 013809
- [47] Rehacek J, Paúr M, Stoklasa B, Hradil Z and Sánchez-Soto L L 2017 Optimal measurements for resolution beyond the Rayleigh limit *Opt. Lett.* **42** 231–4
- [48] Rehacek J, Hradil Z, Stoklasa B, Paúr M, Grover J, Krzic A and Sánchez-Soto L L 2017 Multiparameter quantum metrology of incoherent point sources: towards realistic superresolution *Phys. Rev. A* **96** 062107
- [49] Fuhrmann D R, Preza C, O’Sullivan J A, Snyder D L and Smith W H 2004 Spectrum estimation from quantum-limited interferograms *IEEE Trans. Signal Process.* **52** 950–61
- [50] Fisher R A 1925 Theory of statistical estimation *Math. Proc. Cambridge Philos. Soc.* **22** 700–25
- [51] Rassias T M and Srivastava H M 1992 The orthogonality property of the classical Laguerre polynomials *Appl. Math. Comput.* **50** 167–73

Chapter 6

Neural network quantum state tomography

In this work, I was responsible for the structure of deep neural networks, data preparation for the training of neural networks, and numerical analysis of the results and their comparison to other tomographical techniques. I was also involved in the preparation of the manuscript.

This chapter presents a neural network-based method of quantum state tomography [3]. Quantum state tomography is crucial for certifying and verifying quantum states in modern quantum technologies. The objective of quantum state tomography is to estimate an unknown quantum state through measurements performed on a finite set of identical copies of the system. Traditional methods, such as linear inversion and maximum likelihood estimation, have limitations regarding accuracy and computational efficiency, especially for large-scale quantum systems.

Recent advancements in machine learning, particularly neural networks (NNs), have shown promise in addressing data-driven problems in quantum information. This study explores the application of NNs to quantum state tomography, confirming that the positivity constraint can be effectively implemented with trained networks, thus enabling state-of-the-art deep learning methods for quantum state reconstruction under various types of noise.

QST is an inverse problem to estimate a quantum system's density matrix ρ from measurement outcomes. This work focuses on a d -dimensional quantum system described by a $d \times d$ density matrix ρ . The measurement outcomes are represented by positive operator-valued measures (POVMs), and the probability of detecting a specific outcome is given by Born's rule (3.11), $p_i = \text{Tr}(\Pi_i \rho)$.

The inversion of the linear equations given by Born's rule provides an intuitive but non-optimal solution, often resulting in non-physical states. Alternative methods like semidefinite programming (SDP) and maximum likelihood estimation (MLE) ensure positivity but are computationally intensive. This work introduces a NN-based approach to overcome these limitations, leveraging the efficiency and scalability of deep learning.

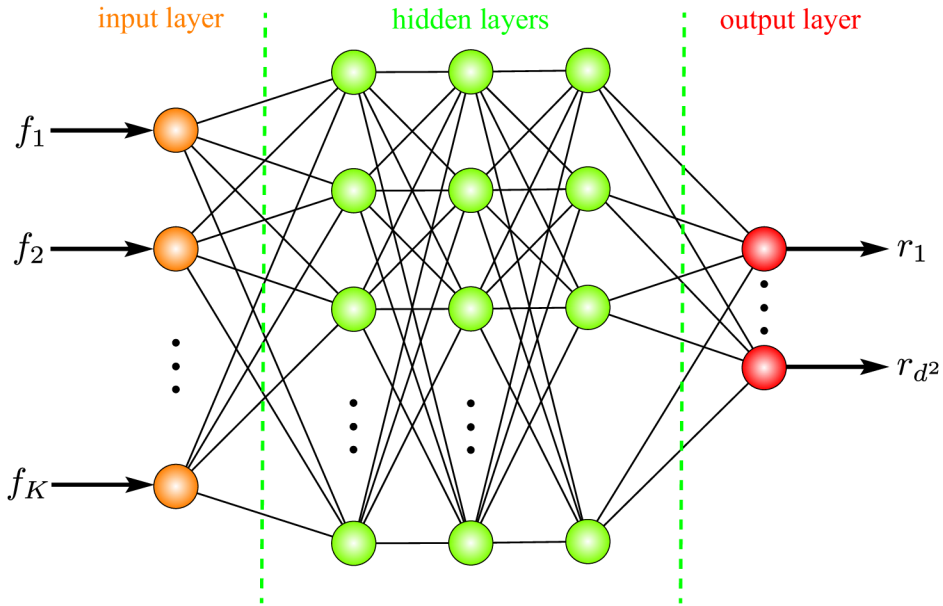


Figure 6.1: Sketch of a NN for estimating the Bloch parameters or the elements of the Cholesky matrix from sampled statistics \mathbf{f} . The blue circles form deep layers of a neural network, representing about 2×10^5 trainable parameters. To fulfill the bounds of the estimated parameters, the output layer has a hyperbolic tangent as an activation function.

The proposed NN architecture, depicted in Fig. 6.1, for QST, consists of a feed-forward network with multiple layers. The input to the NN is the observed frequencies of measurement outcomes, and the output is either the Bloch vector or the elements of the Cholesky matrix representing the quantum state. The network uses the rectified linear unit (ReLU) activation function in hidden layers and the hyperbolic tangent function in the output layer to ensure the outputs are within the appropriate bounds. The NN is trained using stochastic gradient descent with the Nadam optimizer, incorporating Nesterov-accelerated adaptive momentum estimation. The training dataset includes theoretical probabilities and sampled statistics to improve the network's handling of undersampled regimes and noisy data.

The performance of the NN-based approach is benchmarked against traditional methods like linear inversion, SDP, and iterative MLE. Results are shown in Fig. 6.2. We evaluated the average Hilbert-Schmidt distance between the true and estimated quantum states across different dimensions (3, 5, 7, and 9). NNs are about three orders of magnitude faster than SDP and about four orders faster than MLE. The average errors for both NN ap-

proaches are similar to those of SDP and MLE in all dimensions considered. We also showed that NNs can efficiently learn the positivity constraint, resulting in a higher percentage of positive semidefinite states than linear inversion, particularly in the undersampled regime.

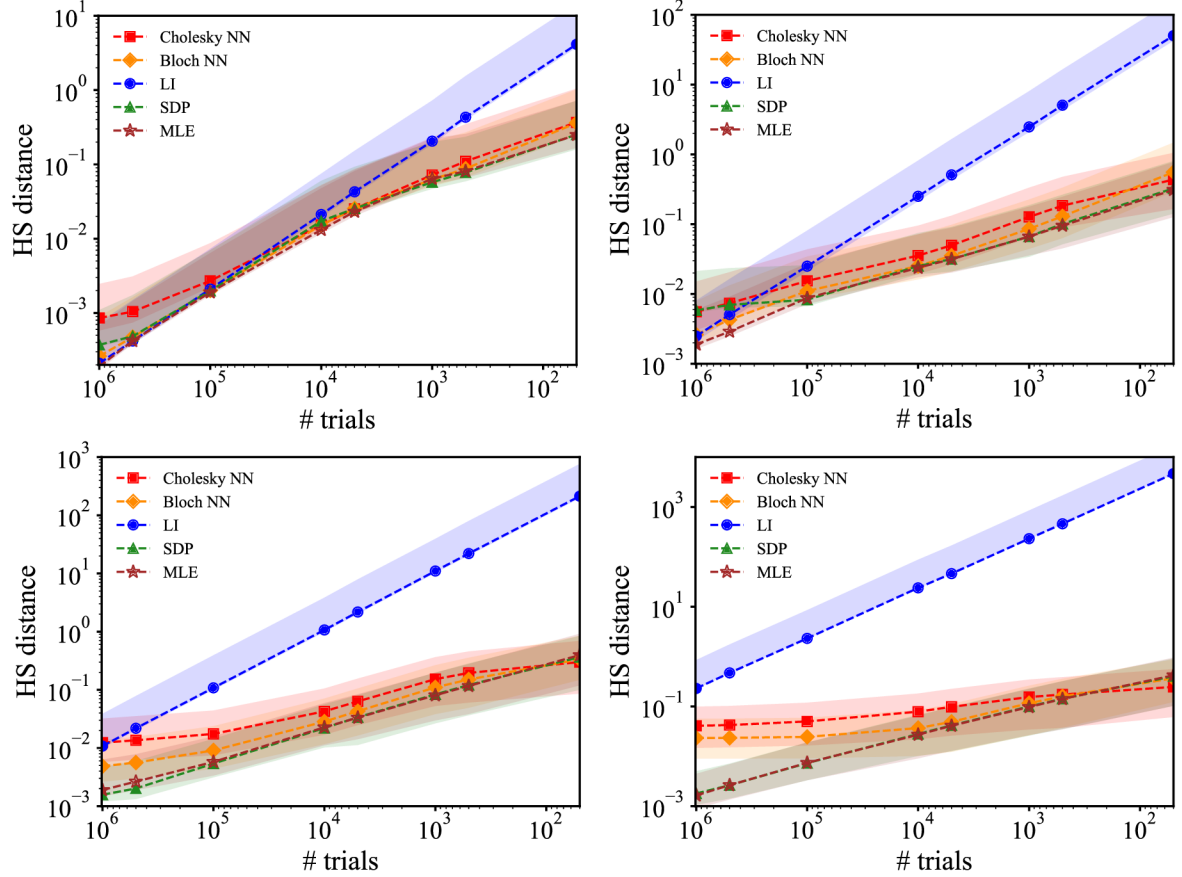




Figure 6.2: Average error in the means of the Hilbert-Schmidt distance for different estimating strategies on the number of trials is shown. The upper left panel depicts errors in estimating quantum states in 3 dimensions. The upper right panel depicts results for dimension 5, and the lower left and right panels depict results for reconstruction errors in dimensions 7 and 9, respectively. The average errors for both NNs in the undersampled regime are in the same order as for SDP and MLE. Confidence intervals of 80% are depicted in respective colors.

We confirmed that NNs can significantly enhance the efficiency and accuracy of QST. By mapping input experimental data to valid density matrices, NNs provide a powerful tool for quantum state reconstruction, offering substantial speed advantages over traditional methods. This work opens up new possibilities for applying deep learning techniques to various quantum information tasks, potentially transforming how quantum states are analyzed and verified. Future research could extend the NN-based approach to larger quantum systems and explore its application in real-time quantum state estimation. Additionally, integrating advanced neural network architectures, such as variational autoencoders and generative

adversarial networks, could further improve the performance and scalability of quantum state tomography.

Neural-network quantum state tomographyDominik Koutný, Libor Motka, Zdeněk Hradil, and Jaroslav Řeháček *Department of Optics, Palacky University, 17. listopadu 12, 77146 Olomouc, Czech Republic*Luis L. Sánchez-Soto *Departamento de Óptica, Facultad de Física, Universidad Complutense, 28040 Madrid, Spain
and Max-Planck-Institut für die Physik des Lichts, 91058 Erlangen, Germany*

(Received 3 December 2021; revised 1 May 2022; accepted 16 June 2022; published 6 July 2022)

We revisit the application of neural networks to quantum state tomography. We confirm that the positivity constraint can be successfully implemented with trained networks that convert outputs from standard feed-forward neural networks to valid descriptions of quantum states. Any standard neural-network architecture can be adapted with our method. Our results open possibilities to use state-of-the-art deep-learning methods for quantum state reconstruction under various types of noise.

DOI: [10.1103/PhysRevA.106.012409](https://doi.org/10.1103/PhysRevA.106.012409)**I. INTRODUCTION**

Modern quantum technologies exploit distinctive features of quantum systems to achieve performances unattainable by classical strategies. This potential advantage hinges on the capability to create, manipulate, and measure quantum states. Any experimental procedure in this area requires a reliable certification of these steps: this is precisely the province of quantum state tomography (QST) [1].

The goal of QST is to estimate the unknown quantum state through measurements performed on a finite set of identical copies of the system. If the state is described by the density matrix ρ , living in a d -dimensional Hilbert space, $O(d/\varepsilon)$ copies are required to obtain an estimate of ρ with an error (understood as total variation distance) less than ε [2]. This clearly illustrates the resource requirements of QST for large-scale systems.

In a broad sense, QST is an inverse problem [3–5]. As such, the linear inversion [6] is probably the most intuitive approach to the topic. Yet it has some cons too: it might report a nonphysical state and the mean squared error bound of the estimate cannot be determined analytically. To bypass these drawbacks a variety of useful QST methods, such as Bayesian tomography [7,8], compressed sensing [9,10], or matrix-product states [11,12], are at hand, although the maximum-likelihood estimation (MLE) is still the most common approach [13,14].

From a modern perspective, QST is fundamentally a data processing problem, trying to extract information from as

few noisy measurements as possible. Therefore, the estimation algorithms used in QST can be easily translated into tasks in machine learning (ML) [15–17]. Actually, neural networks (NNs) have been used to address data-driven problems: examples in quantum information include identifying phase transitions [18], detecting nonclassical features [19–21], quantum error correction [22–25], calibrating quantum devices [26,27], speeding up quantum optimal control [28], and designing quantum experiments [29–31], to cite only a few.

Recently, ML has been applied to QST with very promising results [32–40]. In particular, generative models [41,42], usually restricted Boltzmann machines, have been used to treat the measurement outcomes on a quantum state [43]. These are NNs containing two layers, visible and hidden, with all-to-all connections between the neurons in different layers and none inside each layer. This technique, although powerful, suffers from difficulties with sampling and a lack of straightforward training for larger models.

The use of feed-forward architectures, including recurrent NNs, has been recently advocated [44,45] because these architectures are easier to train without any need for sampling steps, using gradient-based optimization with backpropagation. However, generative tasks in ML often use variational autoencoders [46] and generative adversarial NNs [47]. These are now being actively explored for learning quantum states [48–51].

Our motivation in this paper is to address the benefits of NN-based reconstruction over standard techniques. To fairly benchmark the performance we pick three representative estimators, namely, linear inversion, its positivized version, and MLE, and compare them with a typical NN estimator, obtained with a feed-forward architecture. As measurement, we choose the so-called square root measurement, which was introduced as a “pretty good measurement” [52] for distinguishing possibly nonorthogonal states. Using the Hilbert-Schmidt distance between the true and the reported states as our main indicator, our results suggest that NNs

predict unknown quantum states about three orders of magnitude faster compared to linear and MLE estimates. Interestingly, the average errors are similar for all the estimators considered in all dimensions. This confirms the power of deep-learning-based tools for the quantum realm.

This paper is organized as follows. In Sec. II, we briefly discuss the basic tools of QST we need for our purposes. In Sec. III we describe the details of our NN architecture and training methods. Then, we present the performance of the different estimators in Sec. IV, while our conclusions are summarized in Sec. V.

II. BACKGROUND

We first set the stage for our model. We shall be considering a d -dimensional quantum system, described by a $d \times d$ density matrix ϱ , which requires $n \equiv d^2 - 1$ independent real numbers for its specification.

The goal of QST is to estimate ϱ from measurements performed on identically prepared copies of the system. These measurements are, in general, represented by positive operator-valued measures (POVMs) [53]: they are a set of positive Hermitian operators $\{\Pi_\ell\}$, with the properties

$$\Pi_\ell \geq 0, \quad \Pi_\ell^\dagger = \Pi_\ell, \quad \sum_\ell \Pi_\ell = \mathbb{1}. \quad (2.1)$$

Each POVM element represents a single output of the measuring apparatus. We take every measurement as yielding m distinct outcomes (which we assume to be discrete). According to Born's rule, the probability of detecting the ℓ th output is given by

$$p_\ell = \text{Tr}(\varrho \Pi_\ell). \quad (2.2)$$

To invert this equation, it is convenient to map both ϱ and $\{\Pi_\ell\}$ into a suitable vector form. To this end, we use a traceless Hermitian operator basis $\{\Gamma_k\}$ ($k = 0, \dots, n$) and $\Gamma_0 = \mathbb{1}$, satisfying $\text{Tr}(\Gamma_j) = 0$ and $\text{Tr}(\Gamma_j \Gamma_k) = \delta_{jk}$. In this way, we get the parametrization

$$\varrho = r_0 \Gamma_0 + \sum_{k=1}^n r_k \Gamma_k, \quad \Pi_\ell = w_\ell + \sum_{k=1}^n \mathbf{C}_{\ell k} \Gamma_k. \quad (2.3)$$

Although the condition $\text{Tr} \varrho = 1$ sets $r_0 = 1/d$, we leave r_0 as a parameter to keep the same number of unknowns as for the approach using Cholesky decomposition to be described later. The important point is that the state is characterized by the Bloch vector [54–57] $r_k = \text{Tr}(\varrho \Gamma_k)$, whereas $\mathbf{C}_{\ell k} = \text{Tr}(\Pi_\ell \Gamma_k)$ is a $m \times n$ real matrix describing the explicit relation between the theoretical probabilities \mathbf{p} and the state parameters \mathbf{r} .

In consequence, the inverse problem we have to solve turns out to be the linear system

$$\mathbf{p} = \mathbf{C} \mathbf{r}, \quad (2.4)$$

where we have omitted an unessential constant term that can be incorporated into the following discussion in a straightforward way.

In the presence of noise and with a finite number of copies, the collected data, we will denote by \mathbf{f} , deviates from the expected values \mathbf{p} . The ultimate goal of QST is to infer the

signal parameters \mathbf{r} from the measured noisy data \mathbf{f} . A naive solution is to use the estimator

$$\hat{\mathbf{r}}_{\text{LI}} = \mathbf{C}^{-} \mathbf{f}, \quad (2.5)$$

where the \mathbf{C}^{-} stands for pseudoinverse [58–60] and the subscript LI reminds us that this is a linear inversion approach.

This $\hat{\mathbf{r}}_{\text{LI}}$ is also known as the ordinary least-squares estimator [61]. As heralded before, the resulting $\hat{\mathbf{r}}_{\text{LI}}$ is no longer guaranteed to represent a positive semidefinite operator. One might ensure positivity by setting the negative eigenvalues to zero, which has been called the “quick and dirty” approach [62], although this performs poorly.

Another alternative is to use instead the generalized least-square estimator [63], defined as $\hat{\mathbf{r}}_{\text{GLS}} = [(\mathbf{X}^{-1} \mathbf{C})^+ \mathbf{X}^{-1}] \mathbf{f}$, where \mathbf{X} is such that $\mathbf{C} \mathbf{C}^\dagger$ is the data covariance matrix. Under the Gauss-Markov assumptions [64] it is the best linear unbiased estimator (usually known as BLUE) [65]. However, for small and medium sized datasets, a reliable estimation of the data covariances is not possible, and then $\hat{\mathbf{r}}_{\text{LI}}$ turns out to be a handy estimator.

To circumvent these obstructions we might follow yet another route, introducing instead a semidefinite program that solves (2.4), together with the positivity constraint. The resulting estimator, denoted by $\hat{\mathbf{r}}_{\text{SDP}}$, is thus a solution of

$$\begin{aligned} & \text{minimize} \quad \|\mathbf{f} - \mathbf{C} \mathbf{r}\| \\ & \text{subject to} \quad \varrho \geq 0 \quad \text{and} \quad \text{Tr} \varrho = 1. \end{aligned} \quad (2.6)$$

Finally, to make our analysis complete and consider a nonlinear estimator, we also incorporate the MLE, which guarantees positivity of the resulting quantum state. Although there is vast literature on the subject, the MLE estimate $\hat{\varrho}_{\text{MLE}}$ can be seen as the fixed point of the iterative map [66]

$$\varrho_{k+1} \leftarrow \lambda_k R \varrho_k R, \quad (2.7)$$

with

$$R = \sum_j \frac{f_j}{p_j} \Pi_j \quad (2.8)$$

and λ_k is a normalization constant. The resulting Bloch-vector $\hat{\mathbf{r}}_{\text{MLE}}$ estimate is asymptotically unbiased as $f_j \rightarrow p_j$. Usually, a few thousands of iterations are needed to observe the stationary point of the map (2.7).

From a numerical point of view, an efficient way to deal with the positivity constraint is to directly decompose the density operator using the famous Cholesky factorization [67]

$$\varrho = \frac{\mathbf{A} \mathbf{A}^\dagger}{\text{Tr}(\mathbf{A} \mathbf{A}^\dagger)}, \quad (2.9)$$

where \mathbf{A} is a complex lower triangular matrix and \mathbf{A}^\dagger its Hermitian conjugate. Born's rule then turns to a set of nonlinear equations, which are rather complicated to solve. For this purpose, we adopt ML techniques.

III. NEURAL-NETWORK ESTIMATORS

Our goal is to build a NN that links the input observed frequencies \mathbf{f} to either an output true Bloch vector $\hat{\mathbf{r}}_{\text{NN}}$ or

a Cholesky matrix \mathbf{A}_{NN} . The sampled frequencies \mathbf{f} serve as input to the NN, which transforms them into an output Bloch vector or a Cholesky matrix via a series of linear transformations, each followed by evaluation of some nonlinear function. The structure of such transformations is represented by neurons ordered into deep layers. More precisely, the values $\mathbf{z}^{(k)}$ of the neurons in k th layer are

$$\begin{aligned} \mathbf{z}^{(k)} &= f(\mathbf{y}^{(k)}), \\ \mathbf{y}^{(k)} &= \mathbf{W}^{(k-1 \rightarrow k)} \mathbf{z}^{(k-1)} + \mathbf{b}^{(k-1)}, \end{aligned} \quad (3.1)$$

where \mathbf{W} is a matrix of weights connecting neighboring layers that together with the vector of biases \mathbf{b} forms a set of trainable parameters. The nonlinear activation scalar function f can be chosen arbitrarily, depending on the problem. In our case, the rectified linear unit function $f_{\text{ReLU}}(x) := \max(0, x)$ is used in every deep layer except at the output layer, where we take $f(x) = \tanh(x)$. The hyperbolic tangent function maps real numbers into the $(-1, 1)$ interval which coincides with restriction on elements of both Bloch vector and Cholesky matrix. Symbolically, we can express the Bloch vector (and similarly the elements of the Cholesky matrix) as

$$\mathbf{r}_{\text{NN}} = f_{\text{tanh}} \circ \tilde{\mathbf{W}}^{(n-1 \rightarrow \text{out})} \circ \dots \circ f_{\text{ReLU}} \circ \tilde{\mathbf{W}}^{(\text{in} \rightarrow 1)} \circ \mathbf{f}, \quad (3.2)$$

where $\tilde{\mathbf{W}}$ is a shortcut for $\{\mathbf{W}, \mathbf{b}\}$.

The NN learns by minimizing the loss function. We chose to work with mean squared error, which takes the form

$$\mathcal{L} = \left\langle \sum_{k=0}^n |r_k - z_k^{\text{out}}(\tilde{\mathbf{W}}, f)|^2 \right\rangle, \quad (3.3)$$

where $\langle \cdot \rangle$ denotes the average value in the state ρ . Optimization in the NN is done by backpropagating the error. This is arguably the workhorse of most ML algorithms and definitely the standard approach in most situations, which is working with batches of data. The term and its general use in NNs was coined in [68] and a modern overview is given in a textbook [69].

We minimize the value of the loss function \mathcal{L} over all components of a given dataset to update weights and biases $\tilde{\mathbf{W}}$, using a stochastic gradient-based optimization, which is of core practical importance in many fields [70]. A widely accepted algorithm is Adam [71], which is straightforward to implement, computationally efficient, and has little memory requirements. We use an improved version that incorporates Nesterov-accelerated adaptive momentum estimation (Nadam) [72], since recent results indicate that it has better performance [73]. At the step t , the Nadam procedure updates parameters in the form

$$\tilde{\mathbf{W}}_t \leftarrow \tilde{\mathbf{W}}_{t-1} - \eta \frac{\tilde{\mathbf{m}}_t}{\sqrt{\tilde{\mathbf{h}}_t} + \epsilon}, \quad (3.4)$$

with

$$\begin{aligned} \mathbf{g}_t &\leftarrow \nabla_{\tilde{\mathbf{W}}_{t-1}} \mathcal{L}(\tilde{\mathbf{W}}_{t-1}), \\ \hat{\mathbf{g}} &\leftarrow \frac{\mathbf{g}_t}{1 - \prod_{j=1}^t \mu_j}, \\ \mathbf{m}_t &\leftarrow \mu \mathbf{m}_{t-1} + (1 - \mu) \mathbf{g}_t, \end{aligned}$$

$$\begin{aligned} \hat{\mathbf{m}}_t &\leftarrow \frac{\mathbf{m}_t}{1 - \prod_{j=1}^{t+1} \mu_j}, \\ \mathbf{n}_t &\leftarrow \nu \mathbf{n}_{t-1} + (1 - \nu) \mathbf{g}_t^2, \\ \hat{\mathbf{n}}_t &\leftarrow \frac{\mathbf{n}_t}{1 - \nu^t}, \\ \tilde{\mathbf{m}}_t &\leftarrow (1 - \mu_t) \hat{\mathbf{g}}_t + \mu_{t+1} \hat{\mathbf{m}}_t. \end{aligned} \quad (3.5)$$

Here, η represents the learning rate, μ the exponential decay rate for the first moment estimates $\hat{\mathbf{m}}$, ν 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. We set the numerical values $\{\eta, \mu, \nu, \epsilon\}$ to $\{0.001, 0.9, 0.999, 10^{-7}\}$.

All the above is implemented in the Keras [74] and Tensorflow [75] libraries for PYTHON. The corresponding code can be found in Ref. [76]. In every epoch, the training dataset is divided into 100 batches. The number of epochs needed to find a global minimum of the loss function varies across different deep NNs. In general, the training stopped after 400–2000 epochs, depending on the dimension in which we estimate quantum states. We defined an early stopping after not finding the better minimum of the loss function in the 200 consecutive epochs. We stress that both $\hat{\mathbf{r}}_{\text{NN}}$ and $\hat{\mathbf{r}}_{\text{LI}}$ estimates and the ensuing quantum states are Hermitian matrices but do not incorporate the positivity constraint, in contradistinction to $\hat{\mathbf{r}}_{\text{SDP}}$ and $\hat{\mathbf{r}}_{\text{MLE}}$.

IV. RESULTS

Our deep NN is built as follows: The input layer is fed by observed frequencies \mathbf{f} for different quantum states, followed by eight layers consisting of $(200, 180^{\otimes 2}, 160^{\otimes 4}, 100)$ neurons with the ReLU activation function. The output layer, with

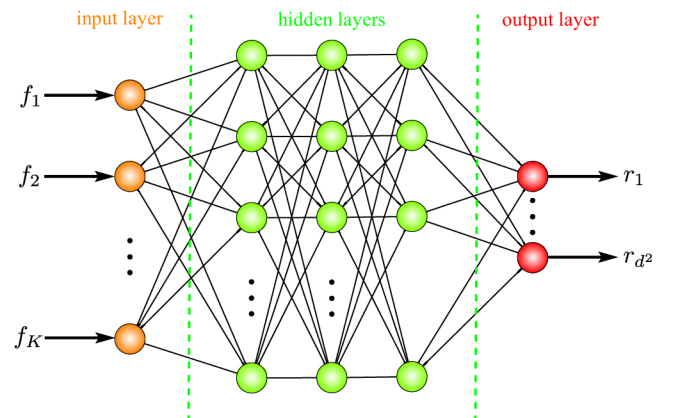


FIG. 1. Sketch of a NN for estimating the Bloch parameters or the elements of the Cholesky matrix from the sampled statistics \mathbf{f} . The filled blue ovals represent neurons in deep layers, representing about 2×10^5 trainable parameters. The output layer has a hyperbolic tangent as an activation function since it has the same bounds as the Bloch parameters and the elements of the Cholesky matrix. In the hidden layers, the rectified linear unit is used as an activation function. The structure of the NN is the same, independently of the dimension or specific parametric representation of quantum states.

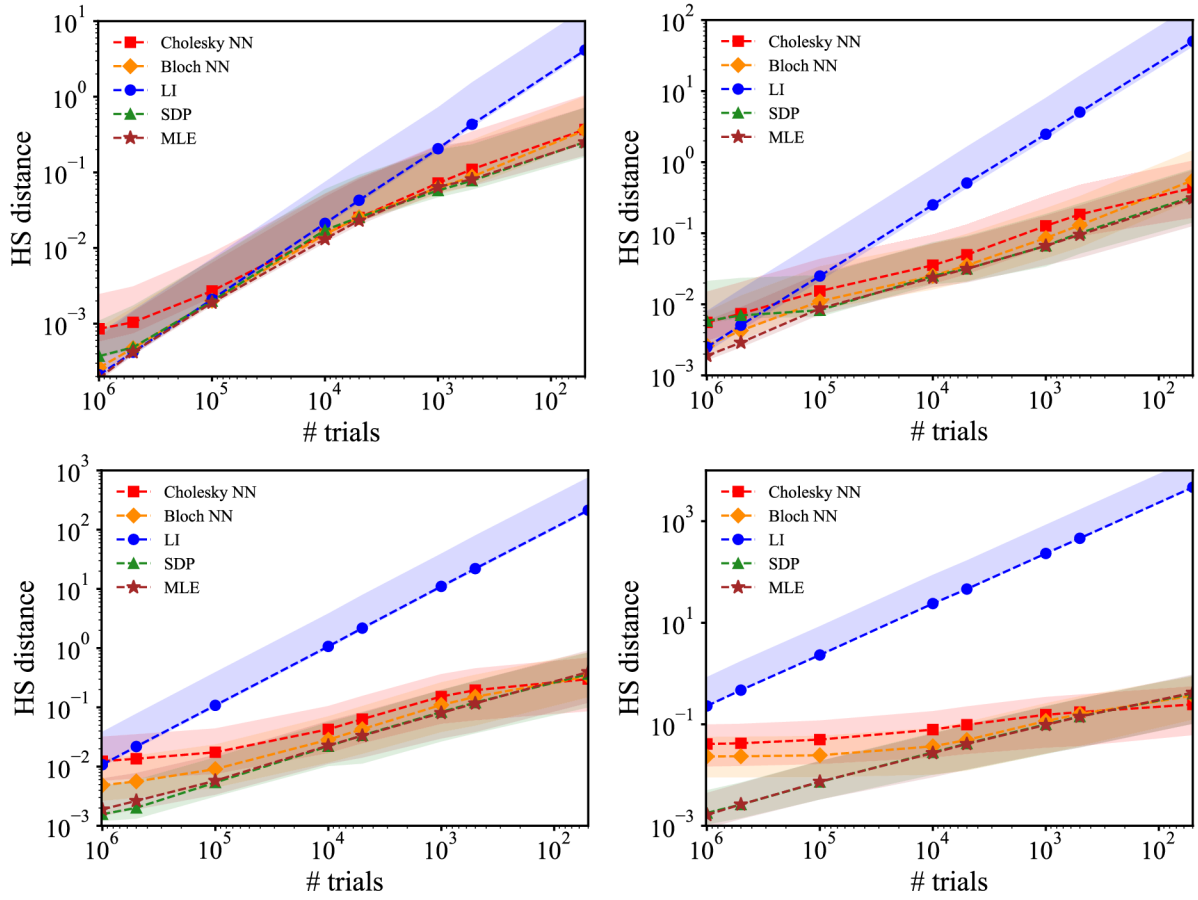


FIG. 2. Average Hilbert-Schmidt distance for different estimating strategies on the number of trials with which we sampled the true probability distribution \mathbf{p} in dimensions 3 (upper left panel), 5 (upper right panel), 7 (lower left panel), and 9 (upper right panel). The insets indicate the corresponding estimators. The average errors for both NNs in the undersampled regime are of the same order for SDP and MLE in all dimensions. When using highly sampled statistics, number of trials $> 10^5$, MLE starts to outperform both NN approaches. Confidence intervals of 80% are depicted in respective colors.

a hyperbolic tangent activation function, serves as an estimate of Bloch vector or the elements of a Cholesky matrix.

The structure of the NN was adjusted heuristically, after having tried multiple settings with differing number of free parameters and deep layers. In terms of the distance between estimated and true quantum states, we got on average the same accuracy for a NN with two layers. However, it was observed that NNs are more likely to return parameters corresponding to positive semidefinite matrices, compared to, e.g., the LI method.

The NN sketch is presented in Fig. 1. We trained in total eight NNs, each with the same structure, for the inference of quantum states in dimensions $d = 3, 5, 7$, and 9. As our target states, we use random density matrices ϱ distributed according to [77]

$$\varrho = \frac{XX^\dagger}{\text{Tr}(XX^\dagger)}, \quad (4.1)$$

with X pertaining to the Ginibre ensemble [78], that is, with real and imaginary parts of each matrix entry being independent normal random variables. These are implemented in PYTHON using QUTIP [79].

As heralded in the Introduction, as our measurement scheme, we choose the square root measurement, defined by the rank-one POVM

$$\Pi_\ell = G^{-1/2} |\phi_\ell\rangle \langle \phi_\ell| G^{-1/2}, \quad G = \sum_\ell |\phi_\ell\rangle \langle \phi_\ell|, \quad (4.2)$$

where $|\phi_\ell\rangle$ are randomly generated Haar-distributed pure states [80] ($\ell = 0, \dots, n$). This POVM is known to be optimal, in the sense that the measurement vectors are the closest in the squared norm to the given states [81,82].

For each dimension, the training (validation) dataset contains 8×10^5 (2×10^5) points. One-quarter of the data in the training dataset are probabilities $p_\ell = \text{Tr}(\varrho \Pi_\ell)$ sampled with a random number of trials, ranging from d^2 up to 10^5 ; the rest are theoretical probabilities. Each input vector, containing either theoretical or sampled statistics, corresponds to different randomly generated quantum states (4.1), whereas the measurement (4.2) is fixed for all states in each dimension.

Ideally, all data points in the training dataset should consist of only theoretical probabilities, so that the NN can extract the appropriate transformation. However, in the undersampled regime, it turns out to be beneficial for the NN to see examples of the sampled statistics. In this case, with a training dataset

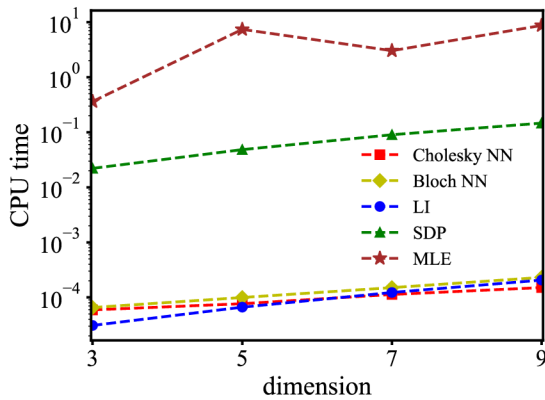


FIG. 3. Average time per state estimation for different tomographical methods, as indicated in the inset. The NN approach is about three orders faster than SDP and about four orders faster than MLE.

containing only theoretical probabilities, the NN would lose the ability to predict positive matrices, while with only sampled probabilities in the training dataset, the NN would have a

hard time in correctly learning the mapping from probabilities to the Bloch vector or the Cholesky matrix.

The output layer consists of true values of the Bloch vector or elements of the Cholesky matrix. All NNs are trained for at most 2000 epochs, which takes about 12 h for every NN when estimating density matrices in dimension 9. After the training procedure, we compared estimates of quantum states based on NNs with standard methods, namely, LI, SPD, and MLE.

Results are shown in Fig. 2. For each dimension d , we generated 10^3 random density matrices and, using the same measurement scheme as for the training, we obtained a set of 10^3 probability distributions, often called a test set. After getting the test set, we sampled each probability distribution with a number of trials depicted on the horizontal axis. Then, we used trained NNs, LI, MLE, and the SDP to reconstruct density matrices from the sampled statistics. We plot the average Hilbert-Schmidt distance between true and inferred quantum states as a function of the number of trials. As we can see, NNs outperform the $\hat{\rho}_{LI}$ estimator and are better or give errors in the same order of magnitude as $\hat{\rho}_{SPD}$.

For a fixed number of trials, the NN-based Bloch vector has a better average error compared to the Cholesky one, but tends to report Hermitian matrices with nonpositive least

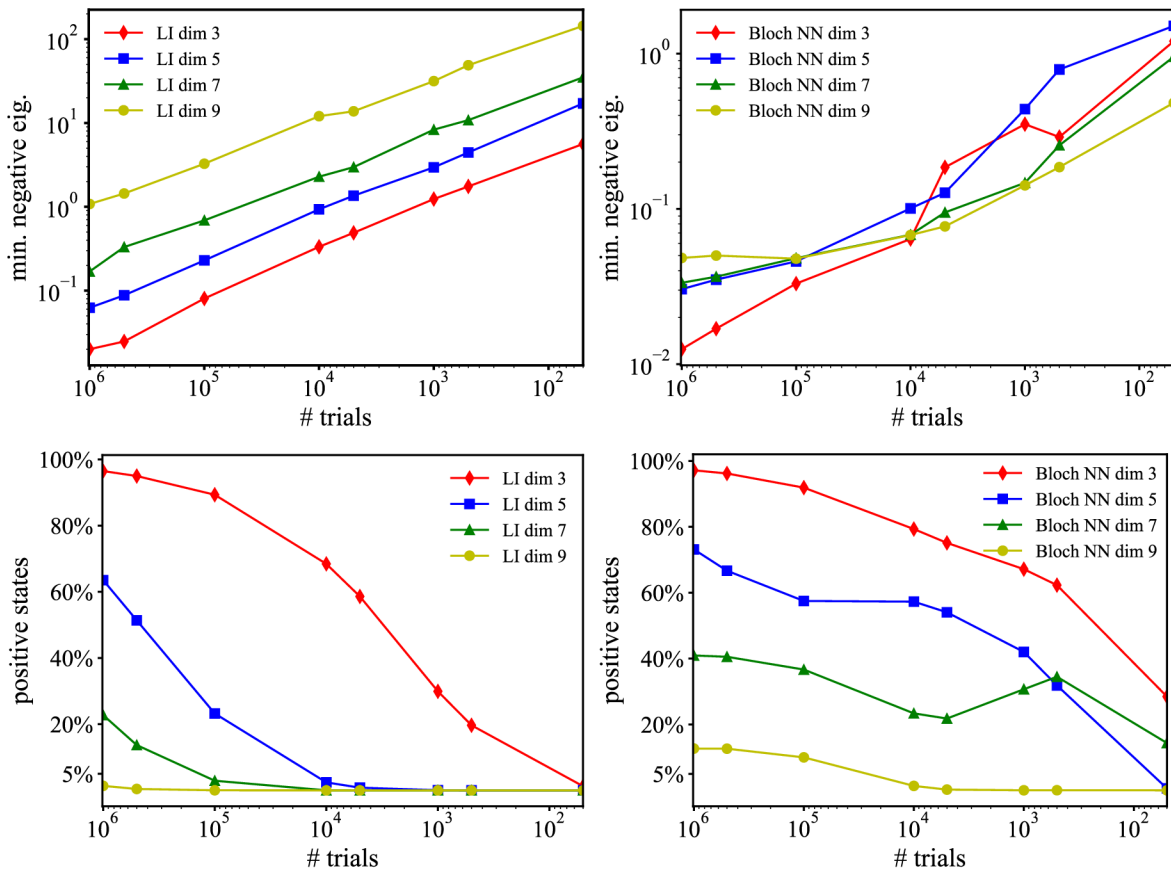


FIG. 4. Quality of estimators based on linear inversion and NN. In all the panels, red diamonds, blue squares, green triangles, and yellow circles represent the results for estimators in dimensions 3, 5, 7, and 9, respectively. The upper plots show the dependence of the most negative eigenvalue on the number of trials for those dimensions and for the LI estimator (left) and the Bloch NN estimator (right). The bottom panels depict the percentage of positive semidefinite states amongst all the reconstructed states with LI (left) and Bloch NN (right). The results are obtained from datasets containing 10^4 statistics in each dimension.

eigenvalues. Moreover, NNs work relatively well in the undersampled regime. Interestingly, NNs show the ability to extrapolate beyond the number of samples on which they were trained. This can be appreciated in the lower Hilbert-Schmidt distance for 10^6 trials, when only up to 10^5 trials were used to sample the true statistics.

The combination of theoretical and sampled probabilities in the training set was balanced in such a way that the NNs work well in the undersampled regime, but also provide accurate estimates when the number of trials is high. Of course, when the number of trials goes to infinity; i.e., when working with theoretical probabilities, the $\hat{\mathbf{r}}_{\text{LI}}$ and $\hat{\mathbf{r}}_{\text{MLE}}$ estimators converge to the true state of the system.

Figure 3 shows an analysis of the performance of different estimators. We depict the average time per single evaluation of $\hat{\mathbf{r}}_{\text{LI}}$, $\hat{\mathbf{r}}_{\text{SDP}}$, $\hat{\mathbf{r}}_{\text{MLE}}$, and both NNs. For the NNs, the times shown are only those associated with the prediction phase, not the training (which takes a lot longer). The semidefinite program infers the quantum state from sampling statistics at around 10^{-1} s. MLE turns out to be the most time-consuming procedure. Compared to linear inversion and both NN approaches, MLE predicts quantum states about 10^4 slower. Akin to linear inversion, NN predicts unknown quantum states from the data about three orders faster compared to SPD estimates and about four orders faster compared to the MLE estimates.

Figure 4 summarizes our performance analysis. We compare the quality of $\hat{\mathbf{r}}_{\text{LI}}$ and $\hat{\mathbf{r}}_{\text{NN}}$ in terms of the *quantumness* of the inferred states. We show the mean of the largest negative eigenvalues and ratio of positive semidefinite quantum states among the set of estimated Hermitian matrices on the measured statistics sampled with a given number of trials. Note that we have excluded results from MLE and SDP, for those estimators always reconstruct a positive matrix.

As one can see in Fig. 4, NNs can learn the positivity constraint. For example, in dimension 9, considering $\hat{\mathbf{r}}_{\text{LI}}$, only 1% of all Hermitian matrices are positive semidefinite, compared to 17% using NN, estimated from statistics sampled

with 10^6 trials. In the undersampled regime, where the number of trials is in the order of the number of projectors, NNs in each dimension predict the higher number of positive quantum states compared to the $\hat{\mathbf{r}}_{\text{LI}}$ estimator.

The issue with predicting quantum states with negative eigenvalues is that it is more prevalent for states that are singular (i.e., with vanishing determinant), as these are the states that sit on the boundary in the generalized Bloch representation. As such, the results in Fig. 4 depend on the purity of the states to be reported.

V. CONCLUDING REMARKS

In summary, we have shown how NN can assist in the reconstruction of quantum states. The NN maps the input experimental data to a valid density matrix up to three orders of magnitude faster than the standard QST. This presents a significant advantage for data postprocessing during tomography. The NN learns to represent the state in a way that is well suited for the problem.

Our results confirm how some of the latest ideas from deep learning can be quite easily adapted and applied to quantum information tasks with just a few tweaks to incorporate the rules of quantum physics. This opens up a wealth of possible applications, which are the object of intense investigation.

ACKNOWLEDGMENTS

The authors thank Miroslav Ježek for useful discussions and two anonymous reviewers for their constructive and detailed comments. This work was supported by the European Union's Horizon 2020 Research and Innovation Programme under the QuantERA Programme through the project ApresSF and from the EU Grant No. 899587 (Project Stormytune), the Palacký University Grant No. IGA_PrF_2021_002, and the Spanish Ministerio de Ciencia e Innovación Grant No. PGC2018-099183-B-I00.

-
- [1] *Quantum State Estimation*, Lecture Notes in Physics, Vol. 649, edited by M. G. A. Paris and J. Řeháček (Springer, Berlin, 2004)
 - [2] R. O'Donnell and J. Wright, Efficient quantum tomography, in *Proceedings of the 48th Annual ACM Symposium on Theory of Computing* (Association for Computing Machinery, Cambridge, MA, 2016), pp. 899–912.
 - [3] L. M. Artiles, R. D. Gill, and M. I. Gutá, An invitation to quantum tomography, *J. R. Stat. Soc., Ser. B* **67**, 109 (2005).
 - [4] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation* (SIAM, Philadelphia, 2005), p. 348.
 - [5] R. Aster, B. Borchers, and C. Thurber, *Parameter Estimation and Inverse Problems*, 2nd ed. (Academic, Boston, 2013).
 - [6] B. Qi, Z. Hou, L. Li, D. Dong, G. Xiang, and G. Guo, Quantum state tomography via linear regression estimation, *Sci. Rep.* **3**, 3496 (2013).
 - [7] R. Blume-Kohout, Optimal, reliable estimation of quantum states, *New J. Phys.* **12**, 043034 (2010).
 - [8] C. Granade, J. Combes, and D. G. Cory, Practical Bayesian tomography, *New J. Phys.* **18**, 033024 (2016).
 - [9] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, Quantum State Tomography via Compressed Sensing, *Phys. Rev. Lett.* **105**, 150401 (2010).
 - [10] D. Ahn, Y. S. Teo, H. Jeong, F. Bouchard, F. Hufnagel, E. Karimi, D. Koutný, J. Řeháček, Z. Hradil, G. Leuchs, and L. L. Sánchez-Soto, Adaptive Compressive Tomography with No *a priori* Information, *Phys. Rev. Lett.* **122**, 100404 (2019).
 - [11] M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, and Y. K. Liu, Efficient quantum state tomography, *Nat. Commun.* **1**, 149 (2010).
 - [12] 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, and C. F. Roos, Efficient tomography of a quantum many-body system, *Nat. Phys.* **13**, 1158 (2017).
 - [13] Z. Hradil, Quantum-state estimation, *Phys. Rev. A* **55**, R1561 (1997).

- [14] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, Measurement of qubits, *Phys. Rev. A* **64**, 052312 (2001).
- [15] K. Murphy, *Machine Learning: A Probabilistic Perspective* (The MIT Press, Cambridge, 2012).
- [16] M. A. Nielsen, *Neural Networks and Deep Learning* (Determination Press, 2015).
- [17] P. Mehta, M. Bukov, C.-H. Wang, A. G. R. Day, C. Richardson, C. K. Fisher, and D. J. Schwab, A high-bias, low-variance introduction to machine learning for physicists, *Phys. Rep.* **810**, 1 (2019).
- [18] B. S. Rem, N. Käming, M. Tarnowski, L. Asteria, N. Fläschner, C. Becker, K. Sengstock, and C. Weitenberg, Identifying quantum phase transitions using artificial neural networks on experimental data, *Nat. Phys.* **15**, 917 (2019).
- [19] V. Gebhart and M. Bohmann, Neural-network approach for identifying nonclassicality from click-counting data, *Phys. Rev. Research* **2**, 023150 (2020).
- [20] C. You, M. A. Quiroz-Juárez, A. Lambert, N. Bhusal, C. Dong, A. Perez-Leija, A. Javaid, R. J. León-Montiel, and O. S. Magaña-Loaiza, Identification of light sources using machine learning, *Appl. Phys. Rev.* **7**, 021404 (2020).
- [21] C. Harney, S. Pirandola, A. Ferraro, and M. Paternostro, Entanglement classification via neural network quantum states, *New J. Phys.* **22**, 045001 (2020).
- [22] G. Torlai and R. G. Melko, Neural Decoder for Topological Codes, *Phys. Rev. Lett.* **119**, 030501 (2017).
- [23] P. Baireuther, T. E. O'Brien, B. Tarasinski, and C. W. J. Beenakker, Machine-learning-assisted correction of correlated qubit errors in a topological code, *Quantum* **2**, 48 (2018).
- [24] S. Krastanov and L. Jiang, Deep neural network probabilistic decoder for stabilizer codes, *Sci. Rep.* **7**, 11003 (2017).
- [25] D. Fitzek, M. Eliasson, A. F. Kockum, and M. Granath, Deep q-learning decoder for depolarizing noise on the toric code, *Phys. Rev. Research* **2**, 023230 (2020).
- [26] E. Flurin, L. S. Martin, S. Hacoen-Gourgy, and I. Siddiqi, Using a Recurrent Neural Network to Reconstruct Quantum Dynamics of a Superconducting Qubit from Physical Observations, *Phys. Rev. X* **10**, 011006 (2020).
- [27] 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. Applied* **15**, 034080 (2021).
- [28] N. Leung, M. Abdelhafez, J. Koch, and D. Schuster, Speedup for quantum optimal control from automatic differentiation based on graphics processing units, *Phys. Rev. A* **95**, 042318 (2017).
- [29] M. Krenn, M. Malik, R. Fickler, R. Lapkiewicz, and A. Zeilinger, Automated Search for New Quantum Experiments, *Phys. Rev. Lett.* **116**, 090405 (2016).
- [30] A. A. Melnikov, H. Poulsen Nautrup, M. Krenn, V. Dunjko, M. Tiersch, A. Zeilinger, and H. J. Briegel, Active learning machine learns to create new quantum experiments, *Proc. Natl. Acad. Sci. USA* **115**, 1221 (2018).
- [31] L. O'Driscoll, R. Nichols, and P. A. Knott, A hybrid machine learning algorithm for designing quantum experiments, *Quantum Mach. Intell.* **1**, 5 (2019).
- [32] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, Neural-network quantum state tomography, *Nat. Phys.* **14**, 447 (2018).
- [33] G. Carleo, Y. Nomura, and M. Imada, Constructing exact representations of quantum many-body systems with deep neural networks, *Nat. Commun.* **9**, 5322 (2018).
- [34] T. Xin, S. Lu, N. Cao, G. Anikeeva, D. Lu, J. Li, G. Long, and B. Zeng, Local-measurement-based quantum state tomography via neural networks, *npj Quantum Inf.* **5**, 109 (2019).
- [35] A. M. Palmieri, E. Kovlakov, F. Bianchi, D. Yudin, S. Straupe, J. D. Biamonte, and S. Kulik, Experimental neural network enhanced quantum tomography, *npj Quantum Inf.* **6**, 20 (2020).
- [36] A. Melkani, C. Gneiting, and F. Nori, Eigenstate extraction with neural-network tomography, *Phys. Rev. A* **102**, 022412 (2020).
- [37] S. Lohani, B. T. Kirby, M. Brodsky, O. Danaci, and R. T. Glasser, Machine learning assisted quantum state estimation, *Mach. Learn.: Sci. Technol.* **1**, 035007 (2020).
- [38] Y. Liu, D. Wang, S. Xue, A. Huang, X. Fu, X. Qiang, P. Xu, H.-L. Huang, M. Deng, C. Guo, X. Yang, and J. Wu, Variational quantum circuits for quantum state tomography, *Phys. Rev. A* **101**, 052316 (2020).
- [39] Y. Quek, S. Fort, and H. K. Ng, Adaptive quantum state tomography with neural networks, *npj Quantum Inf.* **7**, 105 (2021).
- [40] J. Carrasquilla and G. Torlai, How to use neural networks to investigate quantum many-body physics, *PRX Quantum* **2**, 040201 (2021).
- [41] J. Carrasquilla, G. Torlai, R. G. Melko, and L. Aolita, Reconstructing quantum states with generative models, *Nat. Mach. Intell.* **1**, 155 (2019).
- [42] S. Lloyd and C. Weedbrook, Quantum Generative Adversarial Learning, *Phys. Rev. Lett.* **121**, 040502 (2018).
- [43] E. S. Tiunov, V. V. Tiunova (Vyborova), A. E. Ulanov, A. I. Lvovsky, and A. K. Fedorov, Experimental quantum homodyne tomography via machine learning, *Optica* **7**, 448 (2020).
- [44] Z. Cai and J. Liu, Approximating quantum many-body wave functions using artificial neural networks, *Phys. Rev. B* **97**, 035116 (2018).
- [45] P. Cha, P. Ginsparg, F. Wu, J. Carrasquilla, P. L. McMahon, and E.-A. Kim, Attention-based quantum tomography, *Mach. Learn.: Sci. Technol.* **3**, 01LT01 (2022).
- [46] D. Kingma and M. Welling, An introduction to variational autoencoders, *Found. Trends Mach. Learn.* **12**, 307 (2019).
- [47] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, Generative adversarial nets, in *Advances in Neural Information Processing Systems* (Curran Associates, Montreal, 2014), Vol. 3, pp. 2672–2680.
- [48] A. Rocchetto, E. Grant, S. Strelchuk, G. Carleo, and S. Severini, Learning hard quantum distributions with variational autoencoders, *npj Quantum Inf.* **4**, 28 (2018).
- [49] C. Zoufal, A. Lucchi, and S. Woerner, Quantum generative adversarial networks for learning and loading random distributions, *npj Quantum Inf.* **5**, 103 (2019).
- [50] S. Ahmed, C. Sánchez Muñoz, F. Nori, and A. F. Kockum, Quantum State Tomography with Conditional Generative Adversarial Networks, *Phys. Rev. Lett.* **127**, 140502 (2021).
- [51] S. Ahmed, C. Sánchez Muñoz, F. Nori, and A. F. Kockum, Classification and reconstruction of optical quantum states with deep neural networks, *Phys. Rev. Research* **3**, 033278 (2021).
- [52] P. Hausladen and W. K. Wootters, A 'pretty good' measurement for distinguishing quantum states, *J. Mod. Opt.* **41**, 2385 (1994).
- [53] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).

- [54] G. Kimura, The Bloch vector for n -level systems, *Phys. Lett. A* **314**, 339 (2003).
- [55] R. A. Bertlmann and P. Krammer, Bloch vectors for qudits, *J. Phys. A: Math. Theor.* **41**, 235303 (2008).
- [56] I. P. Mendaš, Classification and time evolution of density matrices for a n -state system, *J. Math. Phys.* **49**, 092102 (2008).
- [57] E. Brüning, H. Mäkelä, A. Messina, and F. Petruccione, Parametrizations of density matrices, *J. Mod. Opt.* **59**, 1 (2012).
- [58] R. Penrose, A generalized inverse for matrices, *Math. Proc. Cambridge Philos. Soc.* **51**, 406 (1955).
- [59] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications* (Wiley, New York, 1977).
- [60] S. L. Campbell and C. D. J. Meyer, *Generalized Inverses of Linear Transformations* (Dover, New York, 1991).
- [61] C. Lawson and R. Hanson, *Solving Least Squares Problems* (Prentice-Hall, Englewood Cliffs, NJ, 1974).
- [62] M. S. Kaznady and D. F. V. James, Numerical strategies for quantum tomography: Alternatives to full optimization, *Phys. Rev. A* **79**, 022109 (2009).
- [63] T. Opatrný, D. G. Welsch, and W. Vogel, Least-squares inversion for density-matrix reconstruction, *Phys. Rev. A* **56**, 1788 (1997).
- [64] M. Hallin, Gauss–Markov theorem in statistics, in *Encyclopedia of Environmetrics* (Wiley, New York, 2006).
- [65] S. M. Kay, *Fundamentals of Statistical Signal Processing* (Prentice-Hall, Upper Saddle River, NJ, 1993), Vol. 1.
- [66] J. Řeháček, Z. Hradil, E. Knill, and A. I. Lvovsky, Diluted maximum-likelihood algorithm for quantum tomography, *Phys. Rev. A* **75**, 042108 (2007).
- [67] D. Watkins, *Fundamentals of Matrix Computations* (Wiley, New York, 1991).
- [68] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, Learning representations by back-propagating errors, *Nature (London)* **323**, 533 (1986).
- [69] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT Press, Cambridge, MA, 2016).
- [70] S. Ruder, An overview of gradient descent optimization algorithms, [arXiv:1609.04747](https://arxiv.org/abs/1609.04747).
- [71] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, [arXiv:1412.6980](https://arxiv.org/abs/1412.6980).
- [72] T. Dozat, in *Proceedings of the 4th International Conference on Learning Representations* (San Juan, Puerto Rico, 2016).
- [73] E. M. Dogo, O. J. Afolabi, N. I. Nwulu, B. Twala, and C. O. Aigbavboa, A comparative analysis of gradient descent-based optimization algorithms on convolutional neural networks, in *2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS) Belgaum, India* (IEEE, New York, 2018), pp. 92–99.
- [74] F. Chollet *et al.*, Keras (2015), <https://github.com/fchollet/keras>.
- [75] M. A. *et al.*, TensorFlow: Large-scale machine learning on heterogeneous systems, software available from tensorflow.org.
- [76] <https://github.com/dkoutny/ReconstructionOfQuantumStatesViaNNs>.
- [77] V. A. Osipov, H. Sommers, and K. Życzkowski, Random Bures mixed states and the distribution of their purity, *J. Phys. A: Math. Theor.* **43**, 055302 (2010).
- [78] J. Ginibre, Statistical ensembles of complex, quaternion, and real matrices, *J. Math. Phys.* **6**, 440 (1965).
- [79] J. R. Johansson, P. D. Nation, and F. Nori, QuTiP: An open-source Python framework for the dynamics of open quantum systems, *Comput. Phys. Commun.* **183**, 1760 (2012).
- [80] K. Życzkowski, K. A. Penson, I. Nechita, and B. Collins, Generating random density matrices, *J. Math. Phys.* **52**, 062201 (2011).
- [81] Y. C. Eldar and G. D. Forney, On quantum detection and the square-root measurement, *IEEE Trans. Inf. Theory* **47**, 858 (2001).
- [82] N. Dalla Pozza and G. Pierobon, Optimality of square-root measurements in quantum state discrimination, *Phys. Rev. A* **91**, 042334 (2015).

Chapter 7

Deep learning of quantum entanglement

In this work, I was responsible for implementing neural networks, theoretical analysis, numerical simulations, and experimental data analysis. I was also involved in the manuscript preparation process.

This chapter presents a novel method of quantifying entanglement from incomplete measurements [4]. We showed that with the help of deep learning methods, we achieved a quantification error of up to an order of magnitude lower than the state-of-the-art quantum tomography. We aimed at the quantification of entanglement in two-qubit systems. However, we showed that we could recover the information about entanglement measures from incomplete data in the systems of up to five qubits.

Quantifying quantum entanglement is critical in fundamental quantum research and practical quantum applications. Entanglement is a key resource for various quantum information processing tasks, including quantum communication, computation, and cryptography. Traditional methods for entanglement quantification, such as quantum state tomography, require either detailed prior knowledge of the system or extensive experimental procedures involving complete sets of measurements. Due to their high computational and experimental demands, these methods are often impractical for complex or large-scale quantum systems.

This work introduces an innovative approach leveraging deep learning techniques to quantify quantum entanglement from incomplete measurement data. Our method utilizes neural networks to directly estimate the degree of entanglement without necessitating the full reconstruction of the quantum state. This approach is particularly advantageous as it circumvents the need for full-state tomography, significantly reducing the experimental

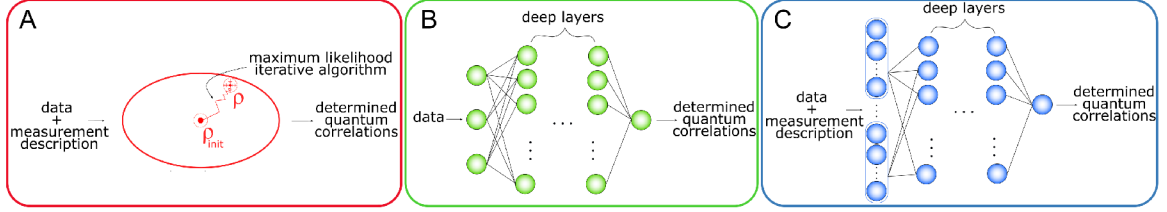


Figure 7.1: Schematics of three different approaches to quantifying quantum entanglement. The maximum likelihood tomographical approach (A) infers the most likely state from data and initial guess ρ_{init} . The green deep neural network represents a fully connected neural network that infers quantum entanglement directly from specific measurements. In contrast, the blue convolutional deep neural network infers quantum entanglement from both measured data as the measurement description.

and computational burden.

We use three approaches to determine the concurrence and the mutual information from an incomplete data set. We show them schematically in Fig. 7.1. We use the maximum likelihood algorithm (MaxLik) (Fig. 7.1A, red), measurement-specific deep neural networks (Fig. 7.1B, green), and a measurement-independent deep neural network (Fig. 7.1C, blue). The MaxLik algorithm iteratively finds the quantum state from which we can calculate the entanglement measure. On the other hand, measurement-specific networks take data from a given measurement and directly predict entanglement. The measurement-independent network works with an arbitrary measurement and directly infers the value of the quantum entanglement measure.

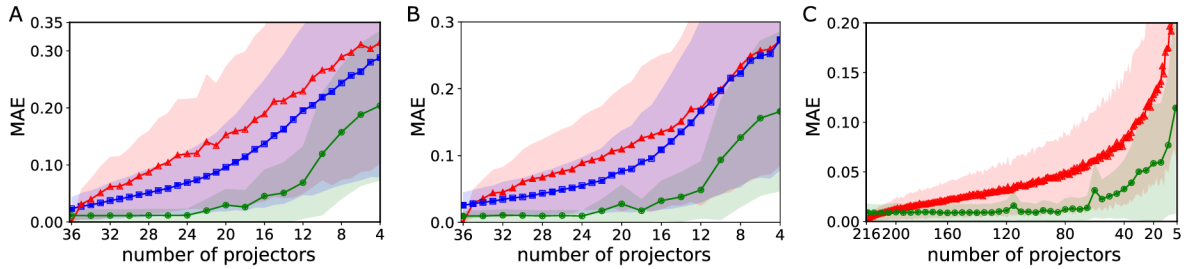


Figure 7.2: The mean absolute error (MAE) versus the number of measurement projectors for (A) two-qubit concurrence, (B) two-qubit mutual information, and (C) three-qubit mutual information matrix. Red triangles depict MAE for the MaxLik; blue squares stand for the values of MAE computed from measurement-independent deep neural network, and, lastly, green circles represent the values of MAE computed from measurement-specific deep neural networks. The DNNs outperform the MaxLik approach regarding entanglement quantification accuracy and consistency, given by smaller errors and uncertainty intervals, even for incomplete measurements.

We compared the three approaches based on how accurately they can infer entanglement measures, namely the concurrence and the quantum mutual information from in-

complete datasets. Results are shown in Fig. 7.2, where we plotted the mean average error (MAE) between the true value of the entanglement measure and the predicted one. In all panels, the red curve stands for MAE for MaxLik, the green curve for measurement-specific DNNs, and the blue curve for measurement-independent DNN. The MAE depends on the number of measurement projectors. The lower amount of projectors leads to higher MAE. As we can see, the measurement-independent DNNs outperform the tomographical approach in estimating the concurrence (Fig. 7.2 A), the two-qubit quantum mutual information (Fig. 7.2 B) and three-qubit quantum mutual information (Fig. 7.2 C). We also generalized the estimation of a quantum mutual information matrix for the systems of up to five qubits.

Moreover, we demonstrated the performance of our DNNs-based method on experimental data. We collected data from two entangled photon sources: continuously pumped spontaneous downconversion and resonantly pumped semiconductor quantum dot. In both cases, measurement-specific DNNs outperformed the MaxLik approach and inferred quantum correlations with higher accuracy.

In conclusion, by developing novel methods of neural networks and deep learning, we outperformed the traditional techniques for quantifying quantum correlations, such as state tomography. The best-performing approach is the measurement-specific DNNs, trained to predict concurrence or quantum mutual information from a fixed set of projectors. Furthermore, we generalized our approach to the systems of up to five qubits, where we showed that measurement-specific DNNs represent a more accurate method to quantify entanglement measures. Furthermore, we demonstrated the ability of our approaches to predict quantum measures from noisy experimental data at higher accuracy than standard methods.



PHYSICS

Deep learning of quantum entanglement from incomplete measurements

Dominik Koutný^{1*}, Laia Ginés², Magdalena Moczala-Dusanowska³, Sven Höfling⁴, Christian Schneider⁵, Ana Predojević², 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 long-standing 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

¹Department of Optics, Faculty of Science, Palacký University, 17. listopadu 12, 77146 Olomouc, Czechia. ²Department of Physics, Stockholm University, 10691 Stockholm, Sweden. ³Princeton Institute of Materials, Princeton University, Princeton, NJ 08544, USA. ⁴Technische Physik, Physikalisches Institut and Würzburg-Dresden Cluster of Excellence ct.qmat, Universität Würzburg, Am Hubland, D-97074 Würzburg, Germany. ⁵Institute 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.

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

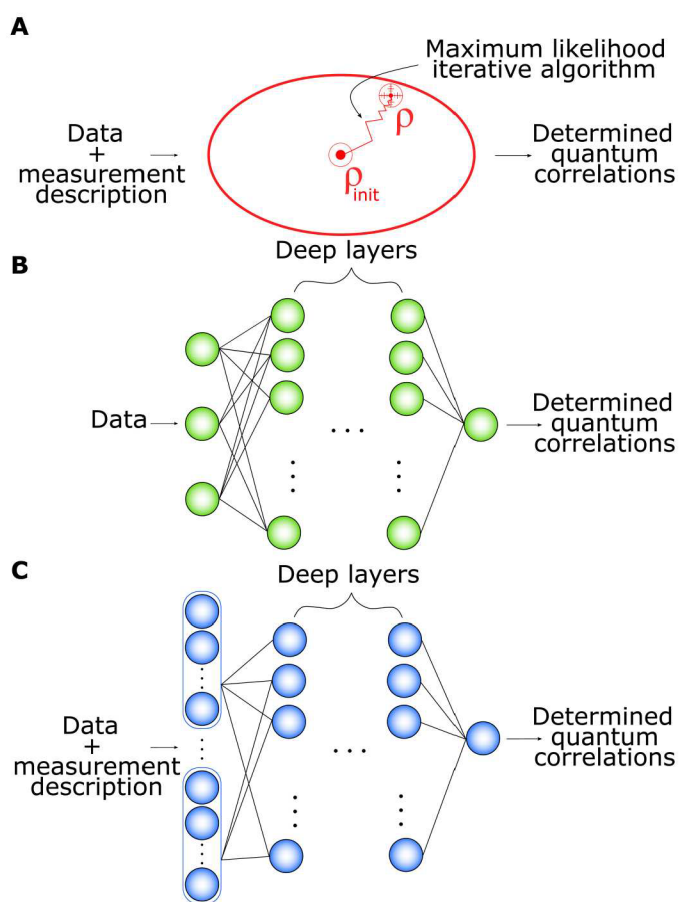


Fig. 1. Schematics of the three methods that we used to infer the quantum correlations. (A) The maximum likelihood algorithm (MaxLik) finds the most likely quantum state ρ based on the measured data and an initial guess ρ_{init} . (B) Green DNN represents a fully connected neural network that infers 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.

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 $\text{Tr}\{\rho M_m\}$ 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 resort 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

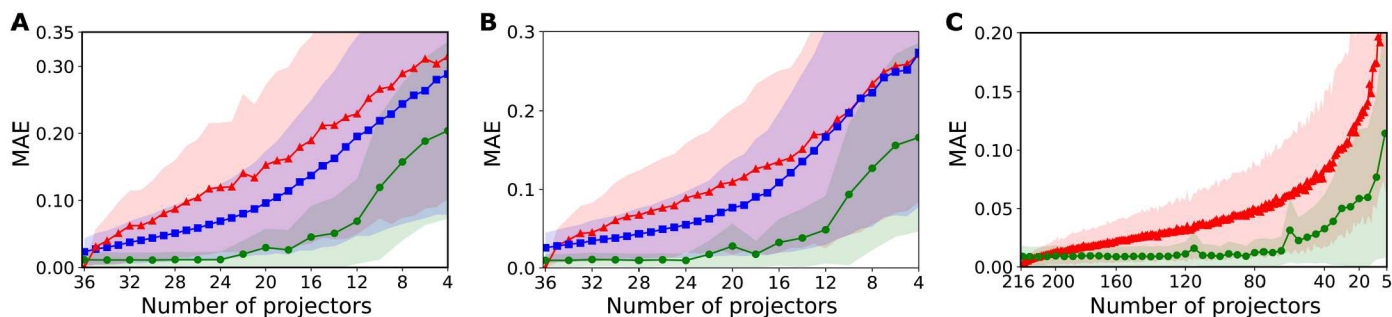


Fig. 2. Entanglement quantification error for the two- and three-qubit systems. The mean absolute error (MAE) versus the number of measurement projectors for (A) two-qubit concurrence, (B) two-qubit mutual information, and (C) three-qubit mutual information matrix. Red triangles depict MAE for the MaxLik, 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.

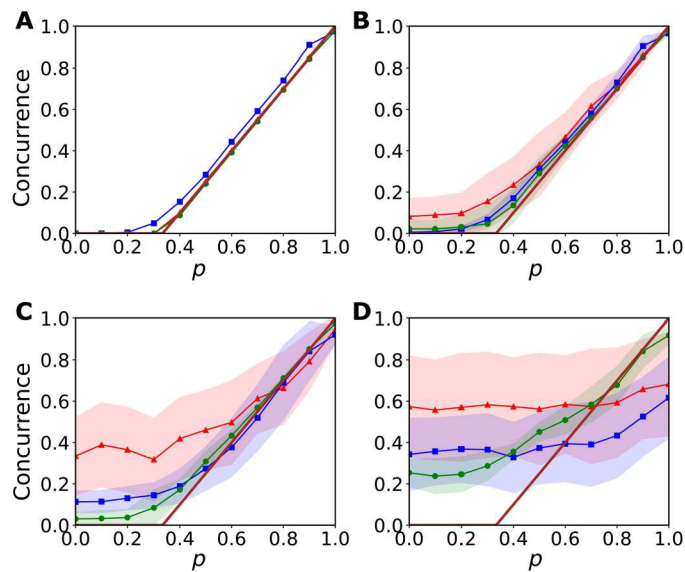


Fig. 3. Entanglement quantification error for the Werner state. The dependence of values of concurrence for the two-qubit Werner state $\rho_W(p)$ characterized by the parameter $p \in [0,1]$. Values of the concurrence determined from (A) 36, (B) 28, (C) 18, and (D) 8 measurement settings. In each panel, 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 $\mathcal{F} \equiv \{\mathcal{F}_{AB}, \mathcal{F}_{AC}, \mathcal{F}_{BC}\}$, with subscripts referring to the three different ways of partition. To determine all three numbers \mathcal{F} 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{F} . 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{F} 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.

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 non-ideal 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)

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

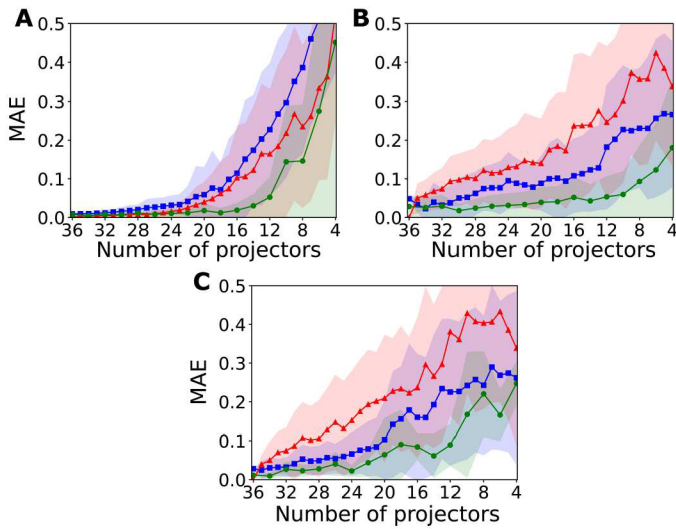


Fig. 4. Performance of MaxLik and DNN-based approaches for an experimental datasets. We show the dependence of the MAE on the number of projectors. (A and B) Spontaneous parametric downconversion sources and (C) semiconductor quantum dot source. The concurrence of experimentally prepared quantum states 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

$$\mathcal{C}(\rho) \equiv \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\} \tag{1}$$

with $\lambda_1, \dots, \lambda_4$ 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 ρ_{sup} / \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

$$\mathcal{I}_{ij} = \frac{1}{2} [\mathcal{S}(\rho_i) + \mathcal{S}(\rho_j) - \mathcal{S}(\rho_{ij})] \tag{2}$$

and is constructed from the one and two point von Neumann entropies (76), $\mathcal{S}(\rho_i) = -\text{Tr}\{\rho_i \log_d \rho_i\}$ and $\mathcal{S}(\rho_{ij}) = -\text{Tr}\{\rho_{ij} \log_d \rho_{ij}\}$, with ρ_i and ρ_{ij} standing for reduced density matrices, $\rho_i = \text{Tr}_{k \neq i}\{\rho\}$ and $\rho_{ij} = \text{Tr}_{k \neq ij}\{\rho\}$ 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 $\{f_j\}$ 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 $\mathcal{L}(\rho) \propto \sum_{j=1}^N f_j \log p_j(\rho)$, 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 $p_i = \text{Tr}\{\rho M_i\}$. 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

$$\rho^{(k+1)} \leftarrow \mu_k G^{-1/2} R \rho^{(k)} R G^{-1/2} \quad (3)$$

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)

$$\rho = \frac{(1 + U^\dagger) G G^\dagger (1 + U)}{\text{Tr}(1 + U^\dagger) G G^\dagger (1 + U)} \quad (4)$$

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

$$\begin{aligned} \mathcal{D}^{\text{input}} &= \{\text{Tr}\{\rho M_1\}, \dots, \text{Tr}\{\rho M_{36}\}\} \\ \mathcal{D}^{\text{output}} &= \{\mathcal{Q}(\rho)\} \end{aligned} \quad (5)$$

where the length of the input vector $\mathcal{D}^{\text{input}}$ is different for various measurement-specific neural networks, ranging from full 36 projectors down to 4. The output $\mathcal{Q}(\rho)$ 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 $\mathcal{Q}_{\text{true}}$ and the values $\mathcal{Q}_{\text{predicted}}^\theta$ predicted by the networks. The loss function thus takes a form

$$\mathcal{L} = |\mathcal{Q}_{\text{true}} - \mathcal{Q}_{\text{predicted}}^\theta| \quad (6)$$

and the minimum is found by minimizing the \mathcal{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

$$\theta_t \leftarrow \theta_{t-1} - \eta \frac{\bar{\mathbf{m}}_t}{\sqrt{\hat{\mathbf{n}}_t} + \epsilon} \quad (7)$$

with

$$\begin{aligned} \mathbf{g}_t &\leftarrow \nabla_{\theta_{t-1}} \mathcal{L}(\theta_{t-1}), \\ \hat{\mathbf{g}}_t &\leftarrow \frac{\mathbf{g}_t}{t}, \\ &1 - \prod_{j=1}^t \mu_j \\ \mathbf{m}_t &\leftarrow \mu \mathbf{m}_{t-1} + (1 - \mu) \mathbf{g}_t, \\ \hat{\mathbf{m}}_t &\leftarrow \frac{\mathbf{m}_t}{t+1}, \\ &1 - \prod_{j=1}^t \mu_j \\ \mathbf{n}_t &\leftarrow \nu \mathbf{n}_{t-1} + (1 - \nu) \mathbf{g}_t^2, \\ \hat{\mathbf{n}}_t &\leftarrow \frac{\mathbf{n}_t}{1-\nu^t}, \\ \bar{\mathbf{m}}_t &\leftarrow (1 - \mu_t) \hat{\mathbf{g}}_t + \mu_{t+1} \hat{\mathbf{m}}_t \end{aligned} \quad (8)$$

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 $\{\eta, \mu, \nu, \epsilon\}$ 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 restructuring of the input layer that inputs not only the vector of probabilities p but also the description of the measurement $\{M_i\}$ itself

$$\mathcal{D}^{\text{input}} = \{M_1, \text{Tr}\{\rho M_1\}, \dots, M_{36}, \text{Tr}\{\rho M_{36}\}\} \quad (9)$$

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 than 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. Gärtner, Time-dependent variational principle for open quantum systems with artificial neural networks. *Phys. Rev. Lett.* **127**, 230501 (2021).
43. A. Szabó, C. Castelnuovo, 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 Ek, C. W. J. Beenakker, Measuring $\text{Tr} \rho^n$ 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. Gühne, 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. Klobus, 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. Černoč, 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. Miřuda, R. Stárek, J. Provazník, O. Leskovařanová, L. Miřta, 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. Řeháček, J. Fiurášek, M. Jeřek, 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. Jeřek, 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. Řeháček, 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. Miřuda, M. Duřek, M. Jeřek, 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

Deep learning of quantum entanglement from incomplete measurements

Dominik Koutný, Laia Ginés, Magdalena Mocza#a-Dusanowska, Sven Höfling, Christian Schneider, Ana Predojevi#, and Miroslav Ježek

Sci. Adv., **9** (29), eadd7131.
DOI: 10.1126/sciadv.add7131

View the article online

<https://www.science.org/doi/10.1126/sciadv.add7131>

Permissions

<https://www.science.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of service](#)

Science Advances (ISSN) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.
Copyright © 2023 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).

Definitions of MLME and SDP tomographic methods and their comparison with MaxLik and DNNs:

In the main text we compare the direct DNN approach to indirect quantum state reconstruction. We use the MaxLik for the quantum state reconstruction and we describe it in the Methods. Here, we review other approaches for quantum state reconstruction, namely maximum-likelihood maximum-entropy (MLME) and semidefinite programming (SDP). Furthermore, we discuss the convergence properties in the case of incomplete measurements and compare the performance of all methods for ideal probabilities and finite sampled data.

The MLME (81) is based on maximization of the log-likelihood functional together with the von Neumann entropy $\mathcal{S}(\rho) = -\text{Tr}\{\rho \log \rho\}$. The total functional reads $\mathcal{F}(\rho, \lambda) \propto \mathcal{L}(\rho) + \lambda \mathcal{S}(\rho)$. In a similar way as for the log-likelihood functional, one can derive from \mathcal{F} an iterative map $\rho^{(k+1)} \leftarrow \mu_k T \rho^{(k)} T$. The operator T takes the form $R - \lambda(\log \rho_k - \text{Tr}\{\rho_k \log \rho_k\})$, with R being the same as in Eq. (3) of the main text, $R = \sum_i f_i/p_i M_i$, and μ_k is the normalization constant ensuring the unit trace of the estimator. The MaxLik and the MLME are both constrained to positive estimates, $\rho \geq 0$.

In the SDP (30,82) the estimator is a solution to the constrained optimization problem, $\rho_{\text{est}} = \min_{\rho} \|\mathcal{A}(\rho) - \mathbf{p}\|$, such that $\rho \geq 0$. The operator \mathcal{A} represents the action of projectors M_i on the quantum state according to the Born's rule $\text{Tr}\{\rho M_i\}$ and \mathbf{p} is the vector of probabilities. We considered two widely used norms $\|\cdot\|$ as a cost function, namely the l_1 norm defined on elements of a vector space \mathbf{x} as $\|\mathbf{x}\|_{l_1} = \sum_i |x_i|$ and the l_2 norm that acts on vectors as $\|\mathbf{x}\|_{l_2} = \sqrt{\sum_i |x_i|^2}$.

Let us comment on how the introduced quantum state reconstruction methods cope with incomplete data. The MaxLik converges to an estimate within a plateau, which is given by input data and an initial iteration. In our case, we start from a maximally mixed state, which spans the whole Hilbert space and allows the iteration process to go “anywhere”. It was shown that the MaxLik method represents an unbiased estimator except for very small amount of measurement data and states close to pure states (76,77). The MLME converges to a state with maximum entropy within a plateau of states given by the measurement data likelihood. Here, we also start from a maximally mixed state for the same reason as given for the MaxLik. The SDP solves a linear positive problem by the primal-dual interior-point method. Based on what cost function is used, the SDP tends to a more or less sparse solution. The sparse (low entropy) solution is usually connected to the L1 norm. Finally, let us point out that minimizing any cost function subjected to the positivity constraint is a compressed sensing protocol (78).

Here, we compare our novel DNN approach to quantification of entanglement with the four quantum state reconstruction methods we listed above. We compute the MAE between true and inferred values of the concurrence in the same way as in the main text. The results are plotted in Fig. S1 for a two-qubit system and the ideal measurement probabilities (infinite sampled data). The DNNs outperform substantially the other methods. On average, the MaxLik approach quantifies the entanglement with either better or similar accuracy when compared to the MLME and SDP. The only exceptions are measurement with very low number of projectors, where the SDP methods slightly outperform the MaxLik. However, considering their large uncertainty regions, all four quantum-state reconstruction methods perform on par. These results show that (i) the DNN approach is superior, and (ii) the MaxLik can be used as an representative example of quantum-state reconstruction method.

We compare the methods also for noisy data (finite sampled), see Fig. S2. The results remain on par with those obtained in the case of the ideal measurement probabilities (infinite sampled data). Moreover, for a small number of the total detections (e.g. one thousand samples), the measurement-specific DNNs outperform the other methods even for the full 36-projector measurement.

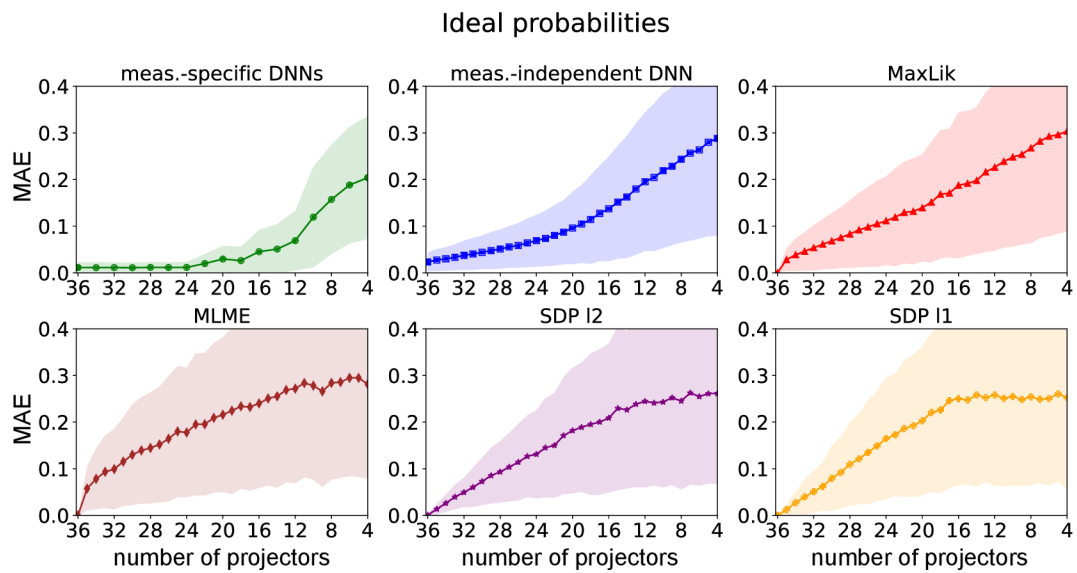


Fig. S1: Comparison of mean absolute error (MAE) of concurrence for different tested methods. The input simulated data are the ideal measurement probabilities. Green circles represent the values of MAE computed from measurement-specific DNNs, MAE for the measurement-independent DNN is depicted in blue squares, Maxlik in red triangles, MLME algorithm in brown diamonds, the SDP with l_2 norm in violet stars and the SDP with l_1 norm in orange crosses.

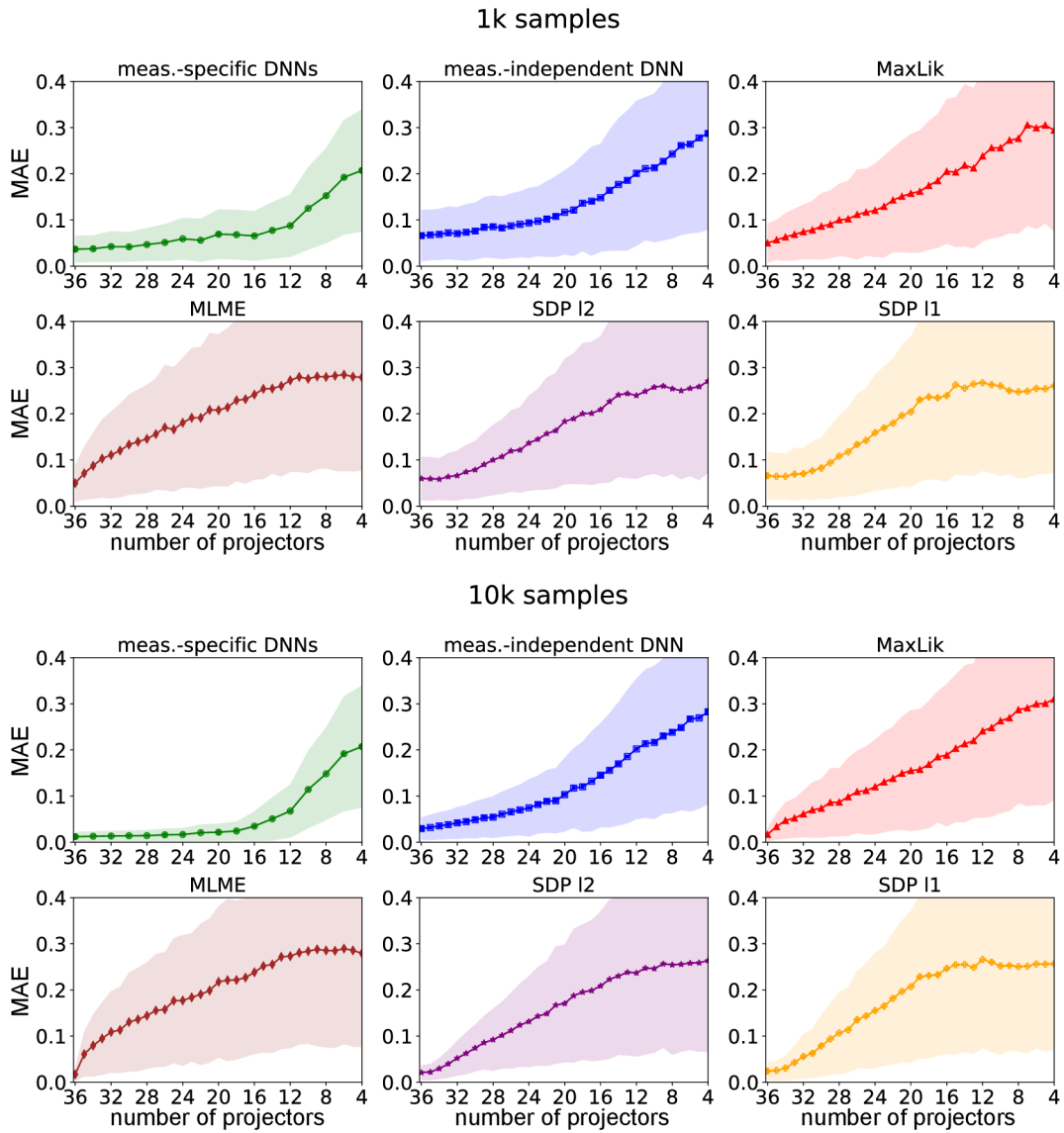


Fig. S2: Comparison of mean absolute error (MAE) of concurrence for different tested methods. The input simulated data are finite sampled; two cases are shown with 10^3 and 10^4 total detections, respectively. Green circles represent the values of MAE computed from measurement-specific DNNs, MAE for the measurement-independent DNN is depicted in blue squares, Maxlik in red triangles, MLME algorithm in brown diamonds, the SDP with l_2 norm in violet stars and the SDP with l_1 norm in orange crosses. Each point is averaged over 1000 randomly generated quantum states and uncertainty regions are depicted in corresponding colors.

Chapter 8

Conclusions

This dissertation focuses on the theoretical analysis of approaches to parameter estimation from data and processing noisy data using neural networks. During my research, we concentrated on theoretical calculations regarding optimal quantum measurements and the determination of quantum and classical limits on the accuracy with which we can estimate parameters from noisy data. Another part of the research was dedicated to the advantages of using deep neural networks for quantum tomography and estimating quantum correlations from incomplete measurements. The work was structured based on four main publications summarizing the most significant results of my doctoral studies.

In this work, we employed a range of advanced methods and tools. Key methods included Fisher information theory and its quantum counterpart, quantum Fisher information, which serve as standard mathematical tools for adequately addressing the problem of parameter estimation from noisy data. Additionally, we applied quantum tomography methods, which facilitate the reconstruction of quantum states from measurements performed on identical copies of the system. Deep neural networks also played a significant role, providing a new approach to estimating quantum entanglement and quantum states, where we achieved increased efficiency and accuracy.

In the article [1], we focused on constructing optimal measurements for two light sources, specifically studying the simultaneous estimation of the centroid, separation, and relative intensities of two incoherent point sources. We calculated the quantum Fisher information matrix analytically for all three parameters. Our results showed that the quantum spatial resolution of two incoherent unbalanced sources could be achieved by projecting the signal onto a suitable complete set of modes. This approach significantly improves measurement accuracy in various applications, such as astronomical observations and microscopy.

In the work [2], we addressed achieving maximum accuracy in axial localization using vortex beams, specifically Laguerre-Gaussian (LG) modes. We demonstrated that the

quantum limit for axial localization could be achieved using an intensity scan if the detector is optimally positioned and the mode is pure. Traditional intensity sensors proved inadequate for estimating the axial distance from superpositions of LG beams, leading to the proposal of using advanced mode sorting methods to fully exploit the potential of vortex beams in microscopy, thereby significantly improving axial resolution.

In the third article [3], we explored the application of neural networks to quantum tomography. We demonstrated that the positivity condition could be effectively implemented into the training set for neural networks, improving predictions using modern deep learning methods for reconstructing quantum states under various types of noise. The neural network-based approach offers significant speed advantages over traditional methods such as semidefinite programming and the maximum likelihood method. However, deep neural networks can also compete with these traditional methods regarding estimation accuracy, even in the presence of noise.

The fourth article [4] presents a new approach to quantifying quantum entanglement using deep neural networks instead of complete quantum state characterization. We showed that even in the case of highly informationally incomplete measurements, neural networks could extract information about quantum entanglement from data compared to standard tomographic techniques. Our neural networks were further tested using experimental data from semiconductor quantum dots and SPDC light sources. For both types of sources, our proposed approach managed to handle the noise present in the experimental data and find an accurate estimate of the quantum entanglement between quantum states.

The results of our research expand the available methods and understanding of fundamental questions in parameter estimation and state reconstruction in quantum physics. The advanced detection and data processing methods we developed represent key building blocks in quantum information processing and quantum communications. Applying deep neural networks in quantum tomography and estimating quantum correlations indicates significant progress in quantum technologies and opens new possibilities for future research. Our work provides both theoretical and practical foundations for future experimental implementations and technological innovations in quantum optics and quantum technologies.

Stručné shrnutí v češtině

Tato disertační práce se zaměřuje na teoretickou analýzu odhadů parametrů z dat a také zpracování zašuměných dat pomocí neuronových sítí. Během mého výzkumu jsme se zaměřili na teoretické výpočty týkající se optimálních kvantových měření a stanovení kvantových a klasických limitů na přesnost, s jakou můžeme odhadovat parametry ze zašuměných dat. Další část výzkumu se věnovala výhodám použití hlubokých neuronových sítí pro kvantovou tomografii a odhad kvantových korelací z neúplných měření. Práce byla strukturována na základě čtyř hlavních publikací, které sumarizují nejvýznamnější výsledky dosažené během mých doktorských studií.

V rámci této práce jsme použili řadu pokročilých metod a nástrojů. Klíčovými metodami byly teorie Fisherovy informace a její kvantový protějšek, kvantová Fisherova informace, které slouží jako standardní matematické nástroje pro adekvátní řešení problému odhadu parametrů ze zašuměných dat. Dále jsme aplikovali metody kvantové tomografie, které umožňují rekonstruovat kvantové stavy z měření provedených na identických kopiích systému. Významnou roli hrály také hluboké neuronové sítě, které poskytly nový přístup k odhadu kvantové provázanosti a kvantových stavů, kde jsme docílili zvýšené efektivity a přesnosti.

V článku [1] jsme se zaměřili na konstrukci optimálního měření pro dva zdroje světla, konkrétně jsme studovali simultánní odhad centroidu, separace a relativních intenzit dvou nekoherentních bodových zdrojů. Pro všechny tři parametry se nám povedlo analyticky napočítat kvantovou Fisherovu informaci. Naše výsledky ukázaly, že kvantová prostorová rozlišovací schopnost dvou nekoherentních nevyvážených zdrojů může být dosažena projektováním signálu na vhodnou kompletní sadu módů. Tento přístup výrazně zlepšuje přesnost měření v různých aplikacích, jako je astronomická pozorování a mikroskopie.

V práci [2] jsme se zabývali dosažením maximální přesnosti v axiální lokalizaci pomocí vírových svazků, konkrétně Laguerreových-Gaussových (LG) módů. Ukázali jsme, že kvantový limit pro axiální lokalizaci může být dosažen pomocí intenzitního skenu, pokud je detektor optimálně umístěn a pokud je jedná o čistý mód. Tradiční intenzitní senzory se ukázaly jako nedostatečné pro odhad axiální vzdálenosti z superpozic LG svazků, což vedlo k návrhu použití pokročilých metod třídění módů pro plné využití potenciálu vírových pa-

prsků v mikroskopii, čímž se výrazně zlepšuje axiální rozlišení.

Ve třetím článku [3] jsme zkoumali aplikaci neuronových sítí na kvantovou tomografii. Demonstrovali jsme, že podmínka positivity může být efektivně implementována do trénovací sady pro neuronové sítě, což zkvalitňuje predikce za pomoci moderních metod hlubokého učení pro rekonstrukci kvantových stavů za různých typů šumu. Přístup založený na neuronových sítích poskytuje nejen výrazné výhody v rychlosti oproti tradičním metodám, jako je semidefinitní programování a metoda maximální věrohodnosti, ale hluboké neuronové sítě také dokážou i za přítomnosti šumu konkurovat těmto tradičním metodám co se týče přesnosti odhadu.

Čtvrtý článek [4] představuje nový přístup ke kvantifikaci kvantového provázání pomocí hlubokých neuronových sítí, namísto úplné charakterizace kvantového stavu. Ukázali jsme, že i v případě silně informačně neúplného měření, neuronové sítě jsou schopny z dat získat informaci o kvantové provázanosti oproti standartním tomografickým technikám. Naše neuronové sítě byly dále testovány na experimentálních datech ze dvou typů zdrojů, polovodičových kvantových teček a SPDC zdrojů světla. Pro oba typy zdrojů se ukázalo, že námi navržený přístup si dokáže poradit i s šumem, kterým jsou experimentální data zatížena a najít přesný odhad kvantové provázanosti mezi kvantovými stavy.

Výsledky našeho výzkumu rozšiřují dostupné metody a porozumění základních otázek na odhady parametrů a rekonstrukci stavů ve kvantové fyzice. Pokročilé metody detekce a zpracování dat, které jsme vyvinuli, představují klíčové stavební bloky nejen v kvantovém zpracování informace, ale i v kvantových komunikacích. Aplikace hlubokých neuronových sítí v kvantové tomografii a odhadu kvantových korelací přináší pokrok v oblasti kvantových technologií a otevírá nové možnosti pro budoucí výzkum. Naše práce poskytuje teoretické i praktické základy pro budoucí experimentální implementace a technologické inovace v kvantové optice a kvantových technologiích.

Publications of the author and a list of citations

The Thesis is based on the following publications:

- J. Řeháček, Z. Hradil, D. Koutný, J. Grover, A. Krzic and L. L. Sánchez-Soto. ‘Optimal measurements for quantum spatial superresolution’. In: *Physical Review A* 98.1 (2018), p. 012103
- D. Koutný, Z. Hradil, J. Řeháček and L. L. Sánchez-Soto. ‘Axial superlocalization with vortex beams’. In: *Quantum Science and Technology* 6.2 (2021), p. 025021
- D. Koutný, L. Motka, Z. Hradil, J. Řeháček and L. L. Sánchez-Soto. ‘Neural-network quantum state tomography’. In: *Physical Review A* 106.1 (2022), p. 012409
- D. Koutný, L. Ginés, M. Moczala-Dusanowska, S. Höfling, Ch. Schneider, A. Predojević and M. Ježek. ‘Deep learning of quantum entanglement from incomplete measurements’. In: *Science Advances* 9.29 (2023)

Other author publications:

- D. Koutný, Y. S. Teo, Z. Hradil and J. Řeháček. ‘Fast universal performance certification of measurement schemes for quantum tomography’. In: *Physical Review A* 94.2 (2016), p. 022113
- M. Mičuda, D. Koutný, M. Miková, I. Straka, M. Ježek and L. Mišta. ‘Experimental demonstration of a fully inseparable quantum state with nonlocalizable entanglement’. In: *Scientific Reports* 7.1 (2017)
- J. Řeháček, Z. Hradil, D. Koutný, J. Grover, A. Krzic and L. L. Sánchez-Soto. ‘Optimal measurements for quantum spatial superresolution’. In: *Physical Review A* 98.1 (2018), p. 012103

- F. Bouchard, D. Koutný, F. Hufnagel, Z. Hradil, J. Řeháček et al. ‘Compressed sensing of twisted photons’. In: *Optics Express* 27.13 (2019), p. 17426
- J. Řeháček, M. Paúr, B. Stoklasa, D. Koutný, Z. Hradil and L. L. Sánchez-Soto. ‘Intensity-Based Axial Localization at the Quantum Limit’. In: *Physical Review Letters* 123.19 (2019), p. 193601
- M. Paúr, B. Stoklasa, D. Koutný, J. Řeháček, Z. Hradil et al. ‘Reading out Fisher information from the zeros of the point spread function’. In: *Optics Letters* 44.12 (2019), p. 3114
- D. Ahn, Y. S. Teo, H. Jeong, D. Koutný, J. Řeháček et al. ‘Adaptive compressive tomography: A numerical study’. In: *Physical Review A* 100.1 (2019), p. 012346
- F. Bouchard, F. Hufnagel, D. Koutný, A. Abbas, A. Sit et al. ‘Quantum process tomography of a high-dimensional quantum communication channel’. In: *Quantum* 3 (2019), p. 138
- Z. Hradil, D. Koutný and J. Řeháček. ‘Exploring the ultimate limits: super-resolution enhanced by partial coherence’. In: *Optics Letters* 46.7 (2021), p. 1728
- M. Peterek, M. Paúr, M. Vitek, D. Koutný, B. Stoklasa et al. ‘Enhancing axial localization with wavefront control’. In: *Optics Express* 31.26 (2023), p. 43722

List of citations (without self-citations):

- J. Řeháček, Z. Hradil, D. Koutný, J. Grover, A. Krzic and L. L. Sánchez-Soto. ‘Optimal measurements for quantum spatial superresolution’. In: *Physical Review A* 98.1 (2018), p. 012103
1. R.B. hammou, A. El Achab and N. Habiballah. ‘Quantum Fisher information matrix of quantum metrology in a Heisenberg XXZ model’. In: *Quantum Studies: Mathematics and Foundations* 11.2 (2024), pp. 263–274
 2. A. Villegas, M.H.M. Passos, S.F. Pereira and J.P. Torres. ‘Optimal parameter estimation of shaped phase objects’. In: *Physical Review A* 109.3 (2024)
 3. W. Krokosz, M. Mazelanik, M. Lipka, M. Jarzyna, W. Wasilewski, K. Banaszek and M. Parniak. ‘Beating the spectroscopic Rayleigh limit via post-processed heterodyne detection’. In: *Optics Letters* 49.4 (2024), pp. 1001–1004

4. A. Sajjad, M.R. Grace and S. Guha. ‘Quantum limits of parameter estimation in long-baseline imaging’. In: *Physical Review Research* 6.1 (2024)
5. J. Miyazaki and K. Matsumoto. ‘Imaginary-free quantum multiparameter estimation’. In: *Quantum* 6 (2022)
6. A.Z. Goldberg, J.L. Romero, Á.S. Sanz and L.L. Sánchez-Soto. ‘Taming singularities of the quantum Fisher information’. In: *International Journal of Quantum Information* 19.8 (2021)
7. S. De, J. Gil-Lopez, B. Brecht, C. Silberhorn, L.L. Sánchez-Soto, Z. Hradil and J. Řeháček. ‘Effects of coherence on temporal resolution’. In: *Physical Review Research* 3.3 (2021)
8. B. Wang, L. Xu, J.-C. Li and L. Zhang. ‘Quantum-limited localization and resolution in three dimensions’. In: *Photonics Research* 9.8 (2021). cited By 7, pp. 1522–1530
9. A. Sajjad, M.R. Grace, Q. Zhuang and S. Guha. ‘Attaining quantum limited precision of localizing an object in passive imaging’. In: *Physical Review A* 104.2 (2021)
10. F. Bao, H. Choi, V. Aggarwal and Z. Jacob. ‘Quantum-accelerated imaging of N stars’. In: *Optics Letters* 46.13 (2021), pp. 3045–3048
11. J. Xin, Y. Li and X.-M. Lu. ‘Optical super-resolution for two unequally bright point sources based on the fractional Hilbert transform’. In: *Physical Review A* 103.5 (2021)
12. C. Oh, S. Zhou, Y. Wong and L. Jiang. ‘Quantum Limits of Superresolution in a Noisy Environment’. In: *Physical Review Letters* 126.12 (2021)
13. J.O. De Almeida, J. Kołodzyński, C. Hirche, M. Lewenstein and M. Skotiniotis. ‘Discrimination and estimation of incoherent sources under misalignment’. In: *Physical Review A* 103.2 (2021)
14. X. Shudong, L. Yanan, X. Jun and L. Xiao-Ming. ‘Incoherent Point Source Resolution Based on Quantum Measurement Optimization’. In: *Laser and Optoelectronics Progress* 58.10 (2021)
15. V. Ansari, B. Brecht, J. Gil-Lopez, J.M. Donohue, J. Řeháček et al. ‘Achieving the Ultimate Quantum Timing Resolution’. In: *PRX Quantum* 2.1 (2021)
16. M. R. Grace, Z. Dutton, A. Ashok and S. Guha. ‘Approaching quantum-limited imaging resolution without prior knowledge of the object location’. In: *Journal of the Optical Society of America B: Optical Physics* 37.8 (2020), pp. 1288–1299

17. I. A. Peshko. ‘Avoiding the “Rayleigh Catastrophe” Using Asymmetric Antennas’. In: *Journal of Applied Spectroscopy* 87.3 (2020), pp. 447–451
18. E. Polino, M. Valeri, N. Spagnolo and F. Sciarrino. ‘Photonic quantum metrology’. In: *AVS Quantum Science* 2.2 (2020)
19. G. Y. Slepyan, S. Vlasenko and D. Mogilevtsev. ‘Quantum Antennas’. In: *Advanced Quantum Technologies* 3.4 (2020)
20. C. Lupo. ‘Subwavelength quantum imaging with noisy detectors’. In: *Physical Review A* 101.2 (2020)
21. Y. L. Len, C. Datta, M. Parniak and K. Banaszek. ‘Resolution limits of spatial mode demultiplexing with noisy detection’. In: *International Journal of Quantum Information* 18.1 (2020)
22. S. Prasad. ‘Quantum limited super-resolution of an unequal-brightness source pair in three dimensions’. In: *Physica Scripta* 95.5 (2020)
23. E. Bisketzi, D. Branford and A. Datta. ‘Quantum limits of localisation microscopy’. In: *New Journal of Physics* 21.12 (2019)
24. M. Tsang. ‘Resolving starlight: a quantum perspective’. In: *Contemporary Physics* 60.4 (2019), pp. 279–298
25. M. Tsang. ‘Semiparametric estimation for incoherent optical imaging’. In: *Physical Review Research* 1.3 (2019)
26. M. Tsang. ‘Quantum limit to subdiffraction incoherent optical imaging’. In: *Physical Review A* 99.1 (2019)
27. I. Peshko, D. Mogilevtsev, I. Karuseichyk, A. Mikhalychev, A. P. Nizovtsev, G. Y. A. Slepyan and A. Boag. ‘Quantum noise radar: Superresolution with quantum antennas by accessing spatiotemporal correlations’. In: *Optics Express* 27.20 (2019), pp. 29217–29231
28. K. A. G. Bonsma-Fisher, W. - K. Tham, H. Ferretti and A. M. Steinberg. ‘Realistic sub-Rayleigh imaging with phase-sensitive measurements’. In: *New Journal of Physics* 21.9 (2019)
29. L. Bakmou, A. Slaoui, M. Daoud and R. Ahl Laamara. ‘Quantum Fisher information matrix in Heisenberg XY model’. In: *Quantum Information Processing* 18.6 (2019)
30. Y. Zhou, J. Yang, J.D. Hassett, S.M.H. Rafsanjani, M. Mirhosseini et al. ‘Quantum-limited estimation of the axial separation of two incoherent point sources’. In: *Optica* 6.5 (2019), pp. 534–541

31. S. Zhou and L. Jiang. ‘Modern description of Rayleigh’s criterion’. In: *Physical Review A* 99.1 (2019)
 32. C. Napoli, S. Piano, R. Leach, G. Adesso and T. Tufarelli. ‘Towards superresolution surface metrology: Quantum estimation of angular and axial separations’. In: *Physical Review Letters* 122.14 (2019)
 33. Y. Wang, Y. Zhang and V.O. Lorenz. ‘Superresolution in interferometric imaging of strong thermal sources’. In: *Physical Review A* 104.2 (2021)
- D. Koutný, Z. Hradil, J. Řeháček and L. L. Sánchez-Soto. ‘Axial superlocalization with vortex beams’. In: *Quantum Science and Technology* 6.2 (2021), p. 025021
1. Y. Wang, Y. Zhang and V.O. Lorenz. ‘Superresolution in interferometric imaging of strong thermal sources’. In: *Physical Review A* 104.2 (2021)
- D. Koutný, L. Motka, Z. Hradil, J. Řeháček and L. L. Sánchez-Soto. ‘Neural-network quantum state tomography’. In: *Physical Review A* 106.1 (2022), p. 012409
1. Z. Zhang and Y.-Z. You. ‘Observing Schrödinger’s cat with artificial intelligence: emergent classicality from information bottleneck’. In: *Machine Learning: Science and Technology* 5.1 (2024)
 2. D. O. Norkin, E. O. Kiktenko and A. K. Fedorov. ‘Reliable confidence regions for quantum tomography using distribution moments’. In: *Physical Review A* 109.3 (2024)
 3. A. Gaikwad, O. Bihani, Arvind and K. Dorai. ‘Neural-network-assisted quantum state and process tomography using limited data sets’. In: *Physical Review A* 109.1 (2024)
 4. J. Wasilewski, T. Paterek and K. Horodecki. ‘Uncertainty of feed forward neural networks recognizing quantum contextuality’. In: *Journal of Physics A: Mathematical and Theoretical* 56.45 (2023)
 5. X. Chai, T. Ma, Q. Guo, Z. Yin, H. Wu and Q. Zhao. ‘Multiqubit State Tomography with only a Few Pauli Measurements’. In: *Physical Review Applied* 20.3 (2023)
 6. J. Wang, S. Zhang, J. Cai, Z. Liao, C. Arenz and R. Betzholz. ‘Robustness of random-control quantum-state tomography’. In: *Physical Review A* 108.2 (2023)
 7. P. M. Vecsei, C. Flindt and J. L. Lado. ‘Lee-Yang theory of quantum phase transitions with neural network quantum states’. In: *Physical Review Research* 5.3 (2023)

8. M. S. Zanetti, D. F. Pinto, M. L. W. Basso and J. Maziero. ‘Simulating noisy quantum channels via quantum state preparation algorithms’. In: *Journal of Physics B: Atomic, Molecular and Optical Physics* 56.11 (2023)
 9. S. Lohani, S. Regmi, J. M. Lukens, R. T. Glasser, T. A. Searles and B. T. Kirby. ‘Dimension-adaptive machine learning-based quantum state reconstruction’. In: *Quantum Machine Intelligence* 5.1 (2023)
 10. I. Gianani and C. Benedetti. ‘Multiparameter estimation of continuous-time quantum walk Hamiltonians through machine learning’. In: *AVS Quantum Science* 5.1 (2023)
 11. X.-D. Wu and S. Cong. ‘Fast and noise-robust quantum state tomography based on ELM’. in: *International Journal of Quantum Information* (2023)
 12. A. Melnikov, M. Kordzanganeh, A. Alodjants and R.-K. Lee. ‘Quantum machine learning: from physics to software engineering’. In: *Advances in Physics: X* 8.1 (2023)
 13. C. Pan and J. Zhang. ‘Deep Learning-Based Quantum State Tomography With Imperfect Measurement’. In: *International Journal of Theoretical Physics* 61.9 (2022)
 14. Z.-T. Li, C.-C. Zheng, F.-X. Meng, H. Zeng, T. Luan, Z.-C. Zhang and X.-T. Yu. ‘Non-Markovian quantum gate set tomography’. In: *Quantum Science and Technology* 9.3 (2024)
 15. Z. Wu, R. Zen, H. P. Casagrande, D. Poletti and S. Bressan. ‘Supervised training of neural-network quantum states for the next-nearest neighbor Ising model’. In: *Computer Physics Communications* 300 (2024)
 16. M.-C. Hsu, E.-J. Kuo, W.-H. Yu, J.-F. Cai and M.-H. Hsieh. ‘Quantum State Tomography via Nonconvex Riemannian Gradient Descent’. In: *Physical Review Letters* 132.24 (2024)
 17. N. Innan, O.I. Siddiqui, S. Arora, T. Ghosh, Y.P. Koçak et al. ‘Quantum state tomography using quantum machine learning’. In: *Quantum Machine Intelligence* 6.1 (2024)
 18. M. Rieger, M. Reh and M. Gärttner. ‘Sample-efficient estimation of entanglement entropy through supervised learning’. In: *Physical Review A* 109.1 (2024)
- D. Koutný, L. Ginés, M. Moczala-Dusanowska, S. Höfling, Ch. Schneider, A. Predojević and M. Ježek. ‘Deep learning of quantum entanglement from incomplete measurements’. In: *Science Advances* 9.29 (2023)

1. M. Rieger, M. Reh and M. Gärttner. ‘Sample-efficient estimation of entanglement entropy through supervised learning’. In: *Physical Review A* 109.1 (2024)
2. J. Denis, F. Damanet and J. Martin. ‘Estimation of the geometric measure of entanglement with Wehrl moments through artificial neural networks’. In: *SciPost Physics* 15.5 (2023)

Bibliography

- [1] J. Řeháček, Z. Hradil, D. Koutný, J. Grover, A. Krzic and L. L. Sánchez-Soto. ‘Optimal measurements for quantum spatial superresolution’. In: *Physical Review A* 98.1 (2018), p. 012103.
- [2] D. Koutný, Z. Hradil, J. Řeháček and L. L. Sánchez-Soto. ‘Axial superlocalization with vortex beams’. In: *Quantum Science and Technology* 6.2 (2021), p. 025021.
- [3] D. Koutný, L. Motka, Z. Hradil, J. Řeháček and L. L. Sánchez-Soto. ‘Neural-network quantum state tomography’. In: *Physical Review A* 106.1 (2022), p. 012409.
- [4] D. Koutný, L. Ginés, M. Moczala-Dusanowska, S. Höfling, Ch. Schneider, A. Predojević and M. Ježek. ‘Deep learning of quantum entanglement from incomplete measurements’. In: *Science Advances* 9.29 (2023).
- [5] G. D. Villiers. *The limits of resolution*. Ed. by Edward R. Pike. Series in optics and optoelectronics. Includes bibliographical references and index. Boca Raton: CRC Press, 2017. 546 pp.
- [6] A. J. den Dekker and A. van den Bos. ‘Resolution: a survey’. In: *Journal of the Optical Society of America A* 14.3 (1997), p. 547.
- [7] B. P. Ramsay, E. L. Cleveland and O. T. Koppius. ‘Criteria and the Intensity-Epoch Slope’. In: *Journal of the Optical Society of America* 31.1 (1941), p. 26.
- [8] Lord Rayleigh. ‘XII. On the manufacture and theory of diffraction-gratings’. In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 47.310 (1874), pp. 81–93.
- [9] Lord Rayleigh. ‘XXXI. Investigations in optics, with special reference to the spectroscope’. In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 8.49 (1879), pp. 261–274.
- [10] E. Abbe. ‘Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung’. In: *Archiv für Mikroskopische Anatomie* 9.1 (1873), pp. 413–468.

- [11] C. M. Sparrow. ‘On Spectroscopic Resolving Power’. In: *The Astrophysical Journal* 44 (1916), p. 76.
- [12] W. V. Houston. ‘A Compound Interferometer for Fine Structure Work’. In: *Physical Review* 29.3 (1927), pp. 478–484.
- [13] W. R. Dawes. ‘Catalogue of Micrometrical Measurements of Double Stars’. In: *Monthly Notices of the Royal Astronomical Society* 27.6 (1867), pp. 217–238.
- [14] H. L. V. Trees. *Detection Estimation and Modulation Theory, Part I. Detection, Estimation, and Filtering Theory*. 1st ed. New York Academy of Sciences Series. Description based on publisher supplied metadata and other sources. Newark: John Wiley Sons, Incorporated, 2013. 11176 pp.
- [15] A. Fischer. ‘Fisher information and Cramér-Rao bound for unknown systematic errors’. In: *Measurement* 113 (2018), pp. 131–136.
- [16] M. Tsang, R. Nair and X.-M. Lu. ‘Quantum Theory of Superresolution for Two Incoherent Optical Point Sources’. In: *Physical Review X* 6.3 (2016), p. 031033.
- [17] C. W. Helstrom, ed. *Quantum detection and estimation theory*. Mathematics in science and engineering v. 123. Master and use copy. Digital master created according to Benchmark for Faithful Digital Reproductions of Monographs and Serials, Version 1. Digital Library Federation, December 2002. New York: Academic Press, 2010. 309 pp.
- [18] M. Hayashi. *Asymptotic Theory of Quantum Statistical Inference: Selected Papers*. WORLD SCIENTIFIC, 2005.
- [19] A. Fujiwara. ‘Strong consistency and asymptotic efficiency for adaptive quantum estimation problems’. In: *Journal of Physics A: Mathematical and General* 39.40 (2006), pp. 12489–12504.
- [20] M. Paúr, B. Stoklasa, Z. Hradil, L. L. Sánchez-Soto and J. Řeháček. ‘Achieving the ultimate optical resolution’. In: *Optica* 3.10 (2016), p. 1144.
- [21] L. Schermelleh, R. Heintzmann and H. Leonhardt. ‘A guide to super-resolution fluorescence microscopy’. In: *Journal of Cell Biology* 190.2 (2010), pp. 165–175.
- [22] S. W. Hell and J. Wichmann. ‘Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy’. In: *Optics Letters* 19.11 (1994), p. 780.
- [23] M. Paúr, B. Stoklasa, D. Koutný, J. Řeháček, Z. Hradil et al. ‘Reading out Fisher information from the zeros of the point spread function’. In: *Optics Letters* 44.12 (2019), p. 3114.

- [24] B. Huang, W. Wang, M. Bates and X. Zhuang. ‘Three-Dimensional Super-Resolution Imaging by Stochastic Optical Reconstruction Microscopy’. In: *Science* 319.5864 (2008), pp. 810–813.
- [25] S. R. P. Pavani, M. A. Thompson, J. S. Biteen, S. J. Lord, N. Liu et al. ‘Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function’. In: *Proceedings of the National Academy of Sciences* 106.9 (2009), pp. 2995–2999.
- [26] J. Řeháček, M. Paúr, B. Stoklasa, D. Koutný, Z. Hradil and L. L. Sánchez-Soto. ‘Intensity-Based Axial Localization at the Quantum Limit’. In: *Physical Review Letters* 123.19 (2019), p. 193601.
- [27] W. S. McCulloch and W. Pitts. ‘A logical calculus of the ideas immanent in nervous activity’. In: *The Bulletin of Mathematical Biophysics* 5.4 (1943), pp. 115–133.
- [28] F. Rosenblatt. ‘The perceptron: A probabilistic model for information storage and organization in the brain.’ In: *Psychological Review* 65.6 (1958), pp. 386–408.
- [29] M Minsky and S. A. Papert. *Perceptrons: An Introduction to Computational Geometry*. The MIT Press, 2017.
- [30] P. Werbos. ‘Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Science. Thesis (Ph. D.). Appl. Math. Harvard University’. PhD thesis. 1974.
- [31] D. E. Rumelhart, E. Hinton G and R. J. Williams. ‘Learning representations by back-propagating errors’. In: *Nature* 323.6088 (1986), pp. 533–536.
- [32] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard and L. D. Jackel. ‘Backpropagation Applied to Handwritten Zip Code Recognition’. In: *Neural Computation* 1.4 (1989), pp. 541–551.
- [33] J. J. Hopfield. ‘Neural networks and physical systems with emergent collective computational abilities.’ In: *Proceedings of the National Academy of Sciences* 79.8 (1982), pp. 2554–2558.
- [34] S. Hochreiter and J. Schmidhuber. ‘Long Short-Term Memory’. In: *Neural Computation* 9.8 (1997), pp. 1735–1780.
- [35] G. Carleo and M. Troyer. ‘Solving the quantum many-body problem with artificial neural networks’. In: *Science* 355.6325 (2017), pp. 602–606.
- [36] G. Torlai and G. G. Melko. ‘Neural Decoder for Topological Codes’. In: *Physical Review Letters* 119.3 (2017), p. 030501.
- [37] S. Krastanov and L. Jiang. ‘Deep Neural Network Probabilistic Decoder for Stabilizer Codes’. In: *Scientific Reports* 7.1 (2017).

- [38] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko and G. Carleo. ‘Neural-network quantum state tomography’. In: *Nature Physics* 14.5 (2018), pp. 447–450.
- [39] X. Gao and L.-M. Duan. ‘Efficient representation of quantum many-body states with deep neural networks’. In: *Nature Communications* 8.1 (2017).
- [40] A. Rocchetto. ‘Stabiliser states are efficiently PAC-learnable’. In: *Quantum Info. Comput.* 18.7–8 (2018), pp. 541–552.
- [41] J. Carrasquilla, D. Luo, F. Pérez, A. Milsted, B. K. Clark, M. Volkovs and L. Aolita. ‘Probabilistic simulation of quantum circuits using a deep-learning architecture’. In: *Physical Review A* 104.3 (2021), p. 032610.
- [42] A. Einstein, B. Podolsky and N. Rosen. ‘Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?’ In: *Physical Review* 47.10 (1935), pp. 777–780.
- [43] E. Schrödinger. ‘Discussion of Probability Relations between Separated Systems’. In: *Mathematical Proceedings of the Cambridge Philosophical Society* 31.4 (1935), pp. 555–563.
- [44] J. S. Bell. ‘On the Einstein Podolsky Rosen paradox’. In: *Physics Physique Fizika* 1.3 (1964), pp. 195–200.
- [45] J. F. Clauser, M. A. Horne, A. Shimony and R. A. Holt. ‘Proposed Experiment to Test Local Hidden-Variable Theories’. In: *Physical Review Letters* 23.15 (1969), pp. 880–884.
- [46] A. Aspect, J. Dalibard and G. Roger. ‘Experimental Test of Bell’s Inequalities Using Time-Varying Analyzers’. In: *Physical Review Letters* 49.25 (1982), pp. 1804–1807.
- [47] Ch. H. Bennett and G. Brassard. ‘Quantum cryptography: Public key distribution and coin tossing’. In: *Theoretical Computer Science* 560 (2014), pp. 7–11.
- [48] P. W. Shor. ‘Algorithms for quantum computation: discrete logarithms and factoring’. In: *Proceedings 35th Annual Symposium on Foundations of Computer Science. SFCS-94*. IEEE Comput. Soc. Press.
- [49] D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter and A. Zeilinger. ‘Experimental quantum teleportation’. In: *Nature* 390.6660 (1997), pp. 575–579.
- [50] J.-W. Pan, Z.-B. Chen, Ch.-Y. Lu, H. Weinfurter, A. Zeilinger and M. Żukowski. ‘Multiphoton entanglement and interferometry’. In: *Reviews of Modern Physics* 84.2 (2012), pp. 777–838.
- [51] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe and J. L. O’Brien. ‘Quantum computers’. In: *Nature* 464.7285 (2010), pp. 45–53.

- [52] B. M. Terhal. ‘Bell inequalities and the separability criterion’. In: *Physics Letters A* 271.5–6 (2000), pp. 319–326.
- [53] J. Roik, K. Bartkiewicz, A. Černoč and K. Lemr. ‘Accuracy of Entanglement Detection via Artificial Neural Networks and Human-Designed Entanglement Witnesses’. In: *Physical Review Applied* 15.5 (2021), p. 054006.
- [54] C. Harney, S. Pirandola, A. Ferraro and M. Paternostro. ‘Entanglement classification via neural network quantum states’. In: *New Journal of Physics* 22.4 (2020), p. 045001.
- [55] J. Gao, L.-F. Qiao, Z.-Q. Jiao, Y.-Ch. Ma, Ch.-Q. Hu et al. ‘Experimental Machine Learning of Quantum States’. In: *Physical Review Letters* 120.24 (2018), p. 240501.
- [56] A. K. Ekert, C. Moura Alves, D. K. L. Oi, M Horodecki, P. Horodecki and L. C. Kwék. ‘Direct Estimations of Linear and Nonlinear Functionals of a Quantum State’. In: *Physical Review Letters* 88.21 (2002), p. 217901.
- [57] K. Audenaert, M. B. Plenio and J. Eisert. ‘Entanglement Cost under Positive-Partial-Transpose-Preserving Operations’. In: *Physical Review Letters* 90.2 (2003), p. 027901.
- [58] A. Peres. ‘Separability Criterion for Density Matrices’. In: *Physical Review Letters* 77.8 (1996), pp. 1413–1415.
- [59] M. Horodecki, P. Horodecki and R. Horodecki. ‘Separability of mixed states: necessary and sufficient conditions’. In: *Physics Letters A* 223.1–2 (1996), pp. 1–8.
- [60] W. K. Wootters. ‘Entanglement of Formation of an Arbitrary State of Two Qubits’. In: *Physical Review Letters* 80.10 (1998), pp. 2245–2248.
- [61] G. Vidal and R. F. Werner. ‘Computable measure of entanglement’. In: *Physical Review A* 65.3 (2002), p. 032314.
- [62] R. Horodecki, P. Horodecki, M. Horodecki and K. Horodecki. ‘Quantum entanglement’. In: *Reviews of Modern Physics* 81.2 (2009), pp. 865–942.
- [63] M. B. Plenio and S. Virmani. ‘An introduction to entanglement measures’. In: *Quantum Information and Computation* 7.1–2 (2007), pp. 1–51.
- [64] L. Amico, R. Fazio, A. Osterloh and V. Vedral. ‘Entanglement in many-body systems’. In: *Reviews of Modern Physics* 80.2 (2008), pp. 517–576.
- [65] O. Gühne and G. Tóth. ‘Entanglement detection’. In: *Physics Reports* 474.1–6 (2009), pp. 1–75.
- [66] J. Eisert, M. Cramer and M. B. Plenio. ‘Colloquium: Area laws for the entanglement entropy’. In: *Reviews of Modern Physics* 82.1 (2010), pp. 277–306.

- [67] R. Barlow. *Statistics. A guide to the use of statistical methods in the physical sciences*. 5. repr. The @Manchester physics series. Enth. Literaturverz. (S. 187-189) u. Index. Chichester [u.a.]: Wiley, 1999. 204 pp.
- [68] Z. Hradil. ‘Quantum-state estimation’. In: *Physical Review A* 55.3 (1997), R1561–R1564.
- [69] D. Koutný, Y. S. Teo, Z. Hradil and J. Řeháček. ‘Fast universal performance certification of measurement schemes for quantum tomography’. In: *Physical Review A* 94.2 (2016), p. 022113.
- [70] M. Mičuda, D. Koutný, M. Miková, I. Straka, M. Ježek and L. Mišta. ‘Experimental demonstration of a fully inseparable quantum state with nonlocalizable entanglement’. In: *Scientific Reports* 7.1 (2017).
- [71] F. Bouchard, D. Koutný, F. Hufnagel, Z. Hradil, J. Řeháček et al. ‘Compressed sensing of twisted photons’. In: *Optics Express* 27.13 (2019), p. 17426.
- [72] D. Ahn, Y. S. Teo, H. Jeong, D. Koutný, J. Řeháček et al. ‘Adaptive compressive tomography: A numerical study’. In: *Physical Review A* 100.1 (2019), p. 012346.
- [73] F. Bouchard, F. Hufnagel, D. Koutný, A. Abbas, A. Sit et al. ‘Quantum process tomography of a high-dimensional quantum communication channel’. In: *Quantum* 3 (2019), p. 138.
- [74] Z. Hradil, D. Koutný and J. Řeháček. ‘Exploring the ultimate limits: super-resolution enhanced by partial coherence’. In: *Optics Letters* 46.7 (2021), p. 1728.
- [75] M. Peterek, M. Paúr, M. Vitek, D. Koutný, B. Stoklasa et al. ‘Enhancing axial localization with wavefront control’. In: *Optics Express* 31.26 (2023), p. 43722.
- [76] R.B. hammou, A. El Achab and N. Habiballah. ‘Quantum Fisher information matrix of quantum metrology in a Heisenberg XXZ model’. In: *Quantum Studies: Mathematics and Foundations* 11.2 (2024), pp. 263–274.
- [77] A. Villegas, M.H.M. Passos, S.F. Pereira and J.P. Torres. ‘Optimal parameter estimation of shaped phase objects’. In: *Physical Review A* 109.3 (2024).
- [78] W. Krokosz, M. Mazelanik, M. Lipka, M. Jarzyna, W. Wasilewski, K. Banaszek and M. Parniak. ‘Beating the spectroscopic Rayleigh limit via post-processed heterodyne detection’. In: *Optics Letters* 49.4 (2024), pp. 1001–1004.
- [79] A. Sajjad, M.R. Grace and S. Guha. ‘Quantum limits of parameter estimation in long-baseline imaging’. In: *Physical Review Research* 6.1 (2024).
- [80] J. Miyazaki and K. Matsumoto. ‘Imaginary-free quantum multiparameter estimation’. In: *Quantum* 6 (2022).

- [81] A.Z. Goldberg, J.L. Romero, Á.S. Sanz and L.L. Sánchez-Soto. ‘Taming singularities of the quantum Fisher information’. In: *International Journal of Quantum Information* 19.8 (2021).
- [82] S. De, J. Gil-Lopez, B. Brecht, C. Silberhorn, L.L. Sánchez-Soto, Z. Hradil and J. Řeháček. ‘Effects of coherence on temporal resolution’. In: *Physical Review Research* 3.3 (2021).
- [83] B. Wang, L. Xu, J.-C. Li and L. Zhang. ‘Quantum-limited localization and resolution in three dimensions’. In: *Photonics Research* 9.8 (2021). cited By 7, pp. 1522–1530.
- [84] A. Sajjad, M.R. Grace, Q. Zhuang and S. Guha. ‘Attaining quantum limited precision of localizing an object in passive imaging’. In: *Physical Review A* 104.2 (2021).
- [85] F. Bao, H. Choi, V. Aggarwal and Z. Jacob. ‘Quantum-accelerated imaging of N stars’. In: *Optics Letters* 46.13 (2021), pp. 3045–3048.
- [86] J. Xin, Y. Li and X.-M. Lu. ‘Optical super-resolution for two unequally bright point sources based on the fractional Hilbert transform’. In: *Physical Review A* 103.5 (2021).
- [87] C. Oh, S. Zhou, Y. Wong and L. Jiang. ‘Quantum Limits of Superresolution in a Noisy Environment’. In: *Physical Review Letters* 126.12 (2021).
- [88] J.O. De Almeida, J. Kołodziej, C. Hirche, M. Lewenstein and M. Skotiniotis. ‘Discrimination and estimation of incoherent sources under misalignment’. In: *Physical Review A* 103.2 (2021).
- [89] X. Shudong, L. Yanan, X. Jun and L. Xiao-Ming. ‘Incoherent Point Source Resolution Based on Quantum Measurement Optimization’. In: *Laser and Optoelectronics Progress* 58.10 (2021).
- [90] V. Ansari, B. Brecht, J. Gil-Lopez, J.M. Donohue, J. Řeháček et al. ‘Achieving the Ultimate Quantum Timing Resolution’. In: *PRX Quantum* 2.1 (2021).
- [91] M. R. Grace, Z. Dutton, A. Ashok and S. Guha. ‘Approaching quantum-limited imaging resolution without prior knowledge of the object location’. In: *Journal of the Optical Society of America B: Optical Physics* 37.8 (2020), pp. 1288–1299.
- [92] I. A. Peshko. ‘Avoiding the “Rayleigh Catastrophe” Using Asymmetric Antennas’. In: *Journal of Applied Spectroscopy* 87.3 (2020), pp. 447–451.
- [93] E. Polino, M. Valeri, N. Spagnolo and F. Sciarrino. ‘Photonic quantum metrology’. In: *AVS Quantum Science* 2.2 (2020).
- [94] G. Y. Slepyan, S. Vlasenko and D. Mogilevtsev. ‘Quantum Antennas’. In: *Advanced Quantum Technologies* 3.4 (2020).

- [95] C. Lupo. ‘Subwavelength quantum imaging with noisy detectors’. In: *Physical Review A* 101.2 (2020).
- [96] Y. L. Len, C. Datta, M. Parniak and K. Banaszek. ‘Resolution limits of spatial mode demultiplexing with noisy detection’. In: *International Journal of Quantum Information* 18.1 (2020).
- [97] S. Prasad. ‘Quantum limited super-resolution of an unequal-brightness source pair in three dimensions’. In: *Physica Scripta* 95.5 (2020).
- [98] E. Bisketzi, D. Branford and A. Datta. ‘Quantum limits of localisation microscopy’. In: *New Journal of Physics* 21.12 (2019).
- [99] M. Tsang. ‘Resolving starlight: a quantum perspective’. In: *Contemporary Physics* 60.4 (2019), pp. 279–298.
- [100] M. Tsang. ‘Semiparametric estimation for incoherent optical imaging’. In: *Physical Review Research* 1.3 (2019).
- [101] M. Tsang. ‘Quantum limit to subdiffraction incoherent optical imaging’. In: *Physical Review A* 99.1 (2019).
- [102] I. Peshko, D. Mogilevtsev, I. Karuseichyk, A. Mikhalychev, A. P. Nizovtsev, G. Y. A. Slepnyan and A. Boag. ‘Quantum noise radar: Superresolution with quantum antennas by accessing spatiotemporal correlations’. In: *Optics Express* 27.20 (2019), pp. 29217–29231.
- [103] K. A. G. Bonsma-Fisher, W. - K. Tham, H. Ferretti and A. M. Steinberg. ‘Realistic sub-Rayleigh imaging with phase-sensitive measurements’. In: *New Journal of Physics* 21.9 (2019).
- [104] L. Bakmou, A. Slaoui, M. Daoud and R. Ahl Laamara. ‘Quantum Fisher information matrix in Heisenberg XY model’. In: *Quantum Information Processing* 18.6 (2019).
- [105] Y. Zhou, J. Yang, J.D. Hassett, S.M.H. Rafsanjani, M. Mirhosseini et al. ‘Quantum-limited estimation of the axial separation of two incoherent point sources’. In: *Optica* 6.5 (2019), pp. 534–541.
- [106] S. Zhou and L. Jiang. ‘Modern description of Rayleigh’s criterion’. In: *Physical Review A* 99.1 (2019).
- [107] C. Napoli, S. Piano, R. Leach, G. Adesso and T. Tufarelli. ‘Towards superresolution surface metrology: Quantum estimation of angular and axial separations’. In: *Physical Review Letters* 122.14 (2019).
- [108] Y. Wang, Y. Zhang and V.O. Lorenz. ‘Superresolution in interferometric imaging of strong thermal sources’. In: *Physical Review A* 104.2 (2021).

- [109] Z. Zhang and Y.-Z. You. ‘Observing Schrödinger’s cat with artificial intelligence: emergent classicality from information bottleneck’. In: *Machine Learning: Science and Technology* 5.1 (2024).
- [110] D. O. Norkin, E. O. Kiktenko and A. K. Fedorov. ‘Reliable confidence regions for quantum tomography using distribution moments’. In: *Physical Review A* 109.3 (2024).
- [111] A. Gaikwad, O. Bihani, Arvind and K. Dorai. ‘Neural-network-assisted quantum state and process tomography using limited data sets’. In: *Physical Review A* 109.1 (2024).
- [112] J. Wasilewski, T. Paterek and K. Horodecki. ‘Uncertainty of feed forward neural networks recognizing quantum contextuality’. In: *Journal of Physics A: Mathematical and Theoretical* 56.45 (2023).
- [113] X. Chai, T. Ma, Q. Guo, Z. Yin, H. Wu and Q. Zhao. ‘Multiqubit State Tomography with only a Few Pauli Measurements’. In: *Physical Review Applied* 20.3 (2023).
- [114] J. Wang, S. Zhang, J. Cai, Z. Liao, C. Arenz and R. Betzholtz. ‘Robustness of random-control quantum-state tomography’. In: *Physical Review A* 108.2 (2023).
- [115] P. M. Vecsei, C. Flindt and J. L. Lado. ‘Lee-Yang theory of quantum phase transitions with neural network quantum states’. In: *Physical Review Research* 5.3 (2023).
- [116] M. S. Zanetti, D. F. Pinto, M. L. W. Basso and J. Maziero. ‘Simulating noisy quantum channels via quantum state preparation algorithms’. In: *Journal of Physics B: Atomic, Molecular and Optical Physics* 56.11 (2023).
- [117] S. Lohani, S. Regmi, J. M. Lukens, R. T. Glasser, T. A. Searles and B. T. Kirby. ‘Dimension-adaptive machine learning-based quantum state reconstruction’. In: *Quantum Machine Intelligence* 5.1 (2023).
- [118] I. Gianani and C. Benedetti. ‘Multiparameter estimation of continuous-time quantum walk Hamiltonians through machine learning’. In: *AVS Quantum Science* 5.1 (2023).
- [119] X.-D. Wu and S. Cong. ‘Fast and noise-robust quantum state tomography based on ELM’. In: *International Journal of Quantum Information* (2023).
- [120] A. Melnikov, M. Kordzanganeh, A. Alodjants and R.-K. Lee. ‘Quantum machine learning: from physics to software engineering’. In: *Advances in Physics: X* 8.1 (2023).
- [121] C. Pan and J. Zhang. ‘Deep Learning-Based Quantum State Tomography With Imperfect Measurement’. In: *International Journal of Theoretical Physics* 61.9 (2022).
- [122] Z.-T. Li, C.-C. Zheng, F.-X. Meng, H. Zeng, T. Luan, Z.-C. Zhang and X.-T. Yu. ‘Non-Markovian quantum gate set tomography’. In: *Quantum Science and Technology* 9.3 (2024).

- [123] Z. Wu, R. Zen, H. P. Casagrande, D. Poletti and S. Bressan. ‘Supervised training of neural-network quantum states for the next-nearest neighbor Ising model’. In: *Computer Physics Communications* 300 (2024).
- [124] M.-C. Hsu, E.-J. Kuo, W.-H. Yu, J.-F. Cai and M.-H. Hsieh. ‘Quantum State Tomography via Nonconvex Riemannian Gradient Descent’. In: *Physical Review Letters* 132.24 (2024).
- [125] N. Innan, O.I. Siddiqui, S. Arora, T. Ghosh, Y.P. Koçak et al. ‘Quantum state tomography using quantum machine learning’. In: *Quantum Machine Intelligence* 6.1 (2024).
- [126] M. Rieger, M. Reh and M. Gärttner. ‘Sample-efficient estimation of entanglement entropy through supervised learning’. In: *Physical Review A* 109.1 (2024).
- [127] J. Denis, F. Damanet and J. Martin. ‘Estimation of the geometric measure of entanglement with Wehrl moments through artificial neural networks’. In: *SciPost Physics* 15.5 (2023).