Multiparameter Method of Moments for Sources Resolving and Characterization

Ilya Karuseichyk

SORBONNE UNIVERSITÉ

DOCTORAL THESIS

Multiparameter Method of Moments for Sources Resolving and Characterization

Author:
Ilya KARUSEICHYK

Supervisor:
Nicolas TREPS

Co-supervisor:
Mattia WALSCHAERS

Members of the jury:

Ivano RUO-BERCHERA	Reviewer
Madalin GUTA	Reviewer
Milena D'ANGELO	Examiner
Damian MARKHAM	President of the jury
Dmitri HOROSHKO	Invited member
Nicolas TREPS	Supervisor

June 2024

Abstract

Quantum metrology presents numerous promising prospects, showing the potential for significant enhancing of the measurement precision across various domains, from imaging to gravitational wave detection. However, assessing whether a given measurement scheme effectively extracts all the available information, as predicted by the quantum Cramer-Rao bound, remains challenging in practical scenarios. Additionally, constructing computationally feasible data-processing algorithms that fully exploit the measured data poses another challenge in multiparameter estimation.

To address these challenges, this thesis adopts the Method of Moments approach to multiparameter estimation — a data-processing technique leveraging the first statistical moments of measurement results. This method provides straightforward estimators with associated sensitivity bounds, facilitating easy computation and relaxing demands on the detection system.

Using this approach, we explore the classical problem of resolving point sources of light and extend its scope to scenarios where bright sources exhibit mutual coherence. Our investigation includes models with diverse statistics and coherence properties, including instances of non-classical statistics or separation-dependent mutual coherence of the sources. By analyzing multiple parameters such as sources' separation, relative and absolute brightness, and phase, we compare the sensitivity of the moment-based spatial mode demultiplexing technique, direct imaging, and the quantum Cramer-Rao bound. Our findings demonstrate a practical estimation approach that often achieves quantum-optimal performance.

Furthermore, we apply the moment-based technique to efficiently characterize Gaussian states using homodyne detection data. We devise an optimal unbiased estimator through algebraic transformations of measured data, providing a simpler alternative to traditional optimization-based methods that are computationally intensive.

Acknowledgments

First and foremost, I wish to express my deepest gratitude to my supervisors, Nicolas Treps and Mattia Walschaers, for their incredible support throughout the three years of my doctoral studies. I felt your moral, psychological, and personal support even before meeting you in person, while I was still in another country. You provided attentive scientific guidance while also granting me the freedom to choose my research topics and methods. I am deeply thankful for the warm and friendly atmosphere within our group and for the sense of freedom and equality. Thank you for the engaging scientific discussions where authority was never used as an argument and everyone remained open to changing their opinions. Thank you for the trust you put in me, for your valuable time, advice, and the resources you provided; for the professional experience I gained under your guidance and for the wonderful moments we shared outside the university; for your help in integrating me into the laboratory team and into French society and culture; and, of course, for your active involvement and assistance in shaping my future career, which I will pursue after leaving the LKB nest. I sincerely hope our professional and personal connection will continue in the future and that our fruitful collaboration will endure.

I would also like to express my gratitude to my colleagues, with whom I spent these wonderful three years, both within the university walls and beyond. Without hesitation, I want to start this long list with Clémentine, who was not only an excellent office mate but also my guide into the world of experimental physics, keeping my theoretical research grounded in reality. It was always a pleasure to communicate with you, and I learned a great deal from our interactions. I would also like to thank Giacomo, who, along with Manuel, was one of the pioneers laying the groundwork for this research. Thank you for the fascinating scientific discussions, your ability to clearly explain complex concepts, your attention to detail, and your readiness to provide quick and helpful advice, even when working in another country. Antonin and Alexander, I am confident that the PESTo project is in good hands with you; it was a pleasure working and not working with you! Nilse, you did an amazing job all these years, bringing humor and life into our daily routines. I was also very happy to work with you on a joint scientific project – it was easy, enjoyable, and productive. Valentina, Claude, David, Carlos, Yann, Leonardo, Bastian, Massimo, Johan, Iris, Francesca, Guilherme, David again, and many other LKB colleagues: each of you has left a mark on my life, and with each of you, I have shared memorable experiences for which I am truly grateful, from you I learned new things about science and life and shared my thoughts and concerns. Thanks to the laboratory management for ensuring smooth operations and shielding us from much of the bureaucracy. A special thanks to Stéphanie and David, who were always efficient and helpful. I am also grateful to Sorbonne University, which prepared my contract in record time and demonstrated a flexible approach to bureaucratic procedures,

adapting them to various non-standard situations. I would also like to acknowledge the PAUSE program, which partially financed my studies and provided invaluable support. My sincere regards go to my colleagues at Thales – Arnaud, Romain, Luc, and of course, Maxime – thank you for the enriching and fruitful experience of scientific collaboration. I would also like to thank my former colleagues from CQOQI – Vlad, Dmitri, Alexander, and others – for their support and the many valuable lessons they provided during the early stages of my research career.

I am deeply thankful to the members of the PhD jury for taking the time to read this manuscript and for attending the defense in person. Ivano, thank you for the stimulating scientific discussions at numerous conferences. Milena, thank you for organizing the excellent summer school in Bari, where I met many remarkable people and gained valuable knowledge. And of course, my special thanks go to Dmitri, whose recommendations enabled me to join this remarkable research group in the first place. Dmitri, your scientific curiosity, enthusiasm, and unwavering ethical standards have always been an example for me to follow.

Undoubtedly, I could not have accomplished this task without the support of my family and friends. My wife, Sonya, has always inspired me to achieve new heights while surrounding me with care and love. I am endlessly grateful for your presence in my life and for everything you do for me. I am also thankful to my parents for their constant support, inexhaustible interest, and patience, for the many pieces of valuable advice, and, of course, for all the hard work they invested in me. My heartfelt thanks go to my friends – Vlad, Marina, Anton, Lyuda, Vova, Misha, Pasha, Nadya, Sergey, and Margo – for being by my side during these years and always being ready to share happy moments or to support me with action or advice in any difficult situation. I deeply appreciate everyone who supported me and stayed in touch; sometimes fate separates us by great distances, but your support remains invaluable.

Contents

List of Abbreviations

(Q)CRB	(Quantum) Cramér-Rao bound
CV	Continuous variable
DI	Direct imaging
EM	Expectation-Maximization
(Q)FI	(Quantum) Fisher information
HG mode	Hermite-Gaussian mode
LO	Local oscillator
LS	List squares
MLE	Maximum likelihood estimation
MoM	Method of moments
MSE	Mean squared error
PSF	Point spread function
SLD	Symmetric logarithmic derivative
SPADE	Spatial-mode demultiplexing

Introduction

Physics, as a science, revolves around measuring phenomena in our surroundings. With advancing technology enabling the exploration of increasingly complex systems, the demand for precise measurements continues to grow. The science and practice of measurement, including all theoretical and practical aspects related to measurement, is called *metrology*. It involves the development of measurement standards, techniques, and methodologies to ensure accuracy, consistency, and reliability in measurements across diverse fields.

Often, the characteristics of interest within a system cannot be directly observed and must be inferred from indirect measurements. Within the realm of optics, a field at the forefront of this thesis, direct measurement is typically limited to light-intensity detection. At the same time, optical metrology proves to be one of the most advanced approaches, offering unparalleled speed and precision of measurements [Gåsvik 2003; Yoshizawa 2017]. In general, indirect inference of unknown parameters from the measurement result is called *parameter estimation* [Kay 1998; Trees 2001].

Classical metrology, analyzing the precision of the parameter estimation, primarily addresses noise arising from experimental imperfections and limitations in measurement devices. The limitation on the estimation precision, coming from the detection noise, is often expressed through statistical bounds like the Cramér–Rao bound (CRB) [Cramér 1946; Rao 1994], based on the Fisher information [Fisher 1925]. Semi-classical approaches also incorporate the shot noise into the detection statistics, acknowledging discrete properties of light, which reflects the quantum nature of light and matter. This intrinsic noise can not be eliminated by improving the measurement devices, the only way to decrease it within the semi-classical model is to obtain higher measurement statistics.

A fully quantum approach to metrology accounts for arbitrary quantum statistics of the measured objects. The statistics of the noise in this case can span beyond the Poissionian model of the shot noise, which opens new possibilities for the optimization of metrological schemes. In the context of optical quantum metrology, an appropriate choice of quantum states of the probe light pulses and of the measurement scheme can significantly increase the sensitivity of the metrological scheme and even achieve better scaling of the sensitivity with the number of photons [Giovannetti 2006; Giovannetti 2011; Demkowicz-Dobrzański 2012; Barbieri 2022; Huang 2024].

However, assessing whether a given measurement scheme effectively extracts all available information from the probe often remains a challenging task in practical scenarios. This challenge stems from the complexity of obtaining full measurement statistics, making the calculation of FI demanding. Additionally, constructing computationally feasible data-processing algorithms that fully exploit the measured data poses another challenge in multiparameter estimation.

The sensitivity of specific metrological schemes, given by the CRB, can be challenging to determine due to difficulties in obtaining full measurement statistics. Constructing an estimator that saturates this bound while remaining computationally feasible poses an even greater challenge [Kay 1998]. To address these issues, this thesis focuses on a specific class of estimators based on the first statistical moments of measured observables. This approach, often referred to as the *method of moments* (MoM), offers several advantages: such estimators are fast and easy to compute, sensitivity bounds for them are easier to determine and saturate, and they relax requirements for the detection system as they don't necessitate high temporal resolution of detection if based on mean measured signals [Kay 1998; Gessner 2019; Gessner 2020b]. We provide a detailed analysis of the properties of this approach in chapter 2.

One of the fundamental examples of the metrological problems, that we briefly consider in section 2.2.4 using the MoM is the estimation of losses in a bosonic channel. Loss estimation proved to be significantly more precise when employing probes with non-classical statistics [Monras 2007; Adesso 2009; Losero 2018; Nair 2018]. In our study, we illustrate that the moment-based approach maintains the benefits of non-classical probe statistics in the majority of scenarios while offering a simple estimator and sensitivity bound.

Optical imaging represents another crucial metrological task, closely linked with the problem of loss estimation. Its objective is to accurately depict the spatial shape of remote or microscopic objects, often being constrained by the diffraction effect within the optical imaging system. Traditionally, improving imaging quality entails making the image (i.e., the measured signal) as similar as possible to the studied object. The initial step in this process involves optimizing the imaging setup itself [Masters 2021]. This optimization includes increasing the numerical aperture (by moving closer to the object, using larger lenses, and employing immersive techniques) and minimizing aberrations, among other strategies. These efforts enable the achievement of resolutions beyond the wavelength of light, with the resolution of the most advanced optical microscopes reaching as low as 200 nm [Villiers 2016].

Another approach to improving optical resolution is often referred to as *active imaging* [Laurenzis 2019]. This set of techniques involves controlling certain properties of light coming from the studied object.

A significant subset of active imaging techniques relies on *fluorescent labeling* of the object [Huang 2009; Leung 2011]. One well-established method in this category is stimulated emission depletion (STED) [Hell 1994; Klar 1999]. STED exploits the non-linear response of fluorescent labels to selectively deactivate some labels, enabling fluorescence from only a tiny spot. This allows for precise localization of emitting spots and estimation of emitter density there, achieving imaging resolutions up to 50 nm. A related technique, called photoactivated localization microscopy (PALM) [Betzig 2006; Hess 2006; Lee 2012], involves active optical switching between fluorescent and non-fluorescent states of labels. This approach enables individual activation and precise localization of emitters. Some imaging techniques rely on stochastic switching between label states, such as stochastic optical reconstruction microscopy (STORM) [Rust 2006; Betzig 2006; Huang 2008; Geissbuehler 2011] and

superresolution optical fluctuation imaging (SOFI) [Dertinger 2009; Dertinger 2013; Vlasenko 2020; Sroda 2020]. In STORM, emitters are individually localized using frames with only one emitter active within the diffraction range. SOFI, on the other hand, enhances overall image sharpness by computing temporal cumulants of a series of images, leveraging independent stochastic switching of sources.

Another example of active imaging relies on controlling the illumination of the object. One well-developed technique in this domain is *structured illumination* imaging [Gustafsson 2000; Gustafsson 2008; Saxena 2015; Ströhl 2016; Heintzmann 2017; Classen 2017; Lim 2024]. In this approach, the studied object is illuminated with spatially structured light in various configurations, allowing for a higher spatial resolution to be achieved by processing the set of obtained images.

Similar approaches are based on correlations of the illumination. These can utilize classical correlations [Valencia 2005; Gatti 2004], as well as quantum spatial and temporal correlations [DAngelo 2005a; Brida 2010; Lopaeva 2013; Meda 2017; Samantaray 2017; Mikhalychev 2019; Berchera 2019; Peshko 2019; Mikhalychev 2021; Defienne 2022] to enhance resolution and contrast while reducing background noise and detection noise. Other correlation-based techniques such as ghost imaging [DAngelo 2005b; Shapiro 2008; Erkmen 2010; Shapiro 2012; Chiuri 2022] and undetected photons imaging [Lemos 2014; Shapiro 2015; Lemos 2022] allow probing the object with lower frequency light while detecting light with higher frequency, thereby enhancing detection efficiency.

There are numerous other optical imaging techniques, such as confocal microscopy [Tata 1998; Price 2011; Gatto Monticone 2014], imaging with the nearfield probing [Guerra 1990; Ohtsu 1995; Hecht 2000]. Each of these techniques offers unique insights into the studied object, along with their own set of advantages and limitations.

Active imaging techniques can achieve remarkable nanometer-scale resolution, but their application range is often limited to microscopy. To utilize these techniques, one typically needs to attach fluorescent labels to the object or control the illumination of the object. This may not always be feasible, especially in remote imaging scenarios. In such cases, one must extract the most information from the light naturally emitted or reflected by the object, which is known as *passive imaging*.

Sometimes, the imaging resolution can be enhanced by processing the image captured by the camera using techniques known as *computational imaging* [Heide 2013; Altmann 2018]. These methods may involve inverting the imaging system transformation with deconvolution algorithms such as Lucy-Richardson [Lucy 1974; Kundur 1996; Sibarita 2005; Khetkeeree 2020], employing neural networks [Xu 2014; Rivenson 2017; Higham 2018; Barbastathis 2019; Belthangady 2019; Hoffman 2021], or other approaches.

The *aperture synthesis* technique is also closely related to computational imaging. This approach is typically used in radio astronomy, where the phase and amplitude of the observed electromagnetic field are measured simultaneously with several remote telescopes [Jennison 1958; Roggemann 1997; Kellermann 2001; Collaboration 2019]. The resolution limit of this approach is defined by the baseline (distance between the telescopes) and significantly exceeds the resolution of individual telescopes. However, realizing this approach in the optical spectral range is challenging due to the difficulty of optical phase synchronization in remote observatories. Thus,

in the optical domain, direct interference of light from each telescope is used by gathering the light in one place with the help of vacuum tubes [Hand 2010]. The baseline of optical astronomical interferometers is limited to several hundred meters, while in the radio domain, it can reach over 100,000 km thanks to satellite-based telescopes [Schuh 2012]. A similar measurement approach, introduced by Hanbury Brown and Twiss, relies on detecting intensity correlations in passive imaging of thermal sources [Brown 1954; Brown 2013; Rosenberg 2022]. This technique allows for determining the angular size of stars with high precision.

Capturing the intensity and direction of light, a technique known as *plenoptic imaging* (or light-field imaging), enables more flexible image processing effects such as refocusing and changing perspective after the image is captured [Wetzstein 2011; DAngelo 2016; Pepe 2017; Di Lena 2018; Massaro 2023; Paniate 2024]. It can also help increase the resolution of imaging in practical scenarios.

Often, the object under study can be fully characterized by a finite set of parameters, such as its moments, Fourier components, local properties in the pixel basis, or positions of point emitters [Helstrom 1967; Kolobov 2000; Beskrovnyy 2005; Motka 2016; Zhou 2019a; Mikhalychev 2019; Tsang 2019c; Pushkina 2021; Fiderer 2021; Köse 2022; Köse 2023]. In such cases, the imaging can be framed as a parameter estimation problem. Consequently, the primary objective of this approach is not to obtain measurement results that resemble the object but to acquire measurement data containing the most information about the parameters of interest. Besides, it enables the utilization of quantum metrology tools to find the ultimate bound for imaging precision and compare it to the sensitivity of the specific measurements employed in practice.

The majority of the previously mentioned techniques rely on detecting light intensity (or correlations of intensities) in a pixel basis using a conventional camera. However, the analysis of the quantum Fisher information shows that this measurement does not extract the full information available in the field. At the same time, the parameter estimation approach opens up the possibility of optimizing an additional degree of freedom that has not been discussed so far — the measurement basis. Our group, specializing in multimode quantum optics, has a rich history of exploring this approach [Fabre 2000; Treps 2003; Treps 2004; Treps 2005; Delaubert 2006a; Delaubert 2006b]. This strategy involves analyzing the informational content of different modes and selecting the most useful mode for measurement, while also strategically allocating quantum resources (like squeezing) in the probe based on this analysis. Initially, this approach was mainly used for the transverse positioning of a beam or single-point source until 2016 when it was generalized to address the problem of resolving incoherent point sources [Tsang 2016]. The developed method was termed *spatial mode demultiplexing* (SPADE). It has demonstrated the potential to significantly increase the information about the separation of two close point sources and saturate the quantum bound in many scenarios. We delve into the details of this approach in section 3.2.3 and provide an overview of a diverse range of recent studies on this topic in section 3.2.4.

Just like with other parameter estimation problems, determining the CRB for the SPADE measurement and constructing a practical estimator that saturates this

bound can be challenging. Most studies on this topic have been conducted under the assumption of faint sources (Poisson limit), which is not always applicable in practical scenarios. In chapter 3, we employ the MoM to study the resolving of bright sources and comprehensively analyze the impact of coherence on resolution, a topic that has been the subject of intense debate in recent years [Larson 2018; Tsang 2019d; Larson 2019; Hradil 2019; Hradil 2021; Liang 2021; Kurdzialek 2022; Liang 2023b; Wang 2023]. Our investigation includes models with diverse statistics and coherence properties, including instances of non-classical statistics or separation-dependent mutual coherence of the sources. By analyzing multiple parameters such as sources' separation, relative and absolute brightness, and phase, we compare the sensitivity of the spatial mode demultiplexing technique, direct imaging, and the quantum Cramer-Rao bound. Our findings demonstrate a practical estimation approach that often achieves quantum-optimal performance.

In addition to efficiently quantifying the sensitivity, the MoM also offers simple estimators that are easy to compute. This characteristic can be valuable in domains where the computational complexity of parameter estimation is critical. One of the examples of such a problem is the characterization of quantum states (or quantum tomography). The setting of this problem and existing approaches to its solution are discussed in section 2.4. In chapter 4, we employ the MoM to construct a computationally efficient estimator that saturates the CRB for characterizing single-mode Gaussian states.

List of publications

This manuscript covers the results, presented in the following papers:

1. [Karuseichyk 2022] Ilya Karuseichyk, Giacomo Sorelli, Mattia Walschaers, Nicolas Treps, and Manuel Gessner. "Resolving mutually-coherent point sources of light with arbitrary statistics." *Physical Review Research* Vol. 4, Iss. 4 (2022): 043010. DOI: 10.1103/PhysRevResearch.4.043010

2. [Karuseichyk 2024] Ilya Karuseichyk, Giacomo Sorelli, Vyacheslav Shatokhin, Mattia Walschaers, and Nicolas Treps. "Exploiting separation-dependent coherence to boost optical resolution." *Physical Review A* Vol. 109, Iss. 4 (2024): 043524. DOI: 10.1103/PhysRevA.109.043524

3. *[In preparation]* Ilya Karuseichyk, Mattia Walschaers, Valentina Parigi, and Nicolas Treps. "Simple and efficient estimator for Gaussian state parameters."

Another related publication, to which I contributed to during my PhD study, is the following experimental study of SPADE:

4. [Rouvière 2024] Clémentine Rouvière, David Barral, Antonin Grateau, Ilya Karuseichyk, Giacomo Sorelli, Mattia Walschaers, and Nicolas Treps. "Ultrasensitive separation estimation of optical sources." *Optica* 11, no. 2 (2024): 166-170. DOI: 10.1364/OPTICA.500039

Its results, however, are not included in the current text.

Chapter 1

Modes, States and Coherence in Quantum Optics

Contents

This chapter provides a concise introduction to fundamental concepts in both classical and quantum optics. Its primary aim is to familiarize readers with key terms and notations that will be utilized in the following chapters, rather than offering an exhaustive tutorial on the subject. The content of this chapter draws heavily from papers [Fabre 2020] and [Glauber 1963], in which interested readers can find more detailed information.

1.1 Modes in classical optics

The behavior of light, as understood through the lens of classical electrodynamics, finds its comprehensive description in Maxwell's equations. These equations describe the interplay between electric and magnetic fields and give us a way to grasp how light travels and evolves. When light propagates through empty space (without

charges), these equations take a specific form, helping us to comprehend the basic rules that govern light without any material medium:

$$\vec{\nabla} \cdot \vec{E}(\vec{r},t) = 0, \qquad \vec{\nabla} \times \vec{E}(\vec{r},t) = -\frac{\partial \vec{B}(\vec{r},t)}{\partial t}, \tag{1.1}$$

$$\vec{\nabla} \cdot \vec{B}(\vec{r},t) = 0, \qquad \vec{\nabla} \times \vec{B}(\vec{r},t) = \mu_0 \varepsilon_0 \frac{\partial \vec{E}(\vec{r},t)}{\partial t}. \tag{1.2}$$

With this system of equations, it is straightforward to demonstrate that the amplitude of the electric (and magnetic) field follows the standard wave equation:

$$\Delta \vec{E}(\vec{r},t) - \frac{1}{c^2}\frac{\partial^2 \vec{E}(\vec{r},t)}{\partial t^2} = 0, \tag{1.3}$$

where $c = (\mu_0 \varepsilon_0)^{-1/2}$. This equation governs the wave-like evolution of light propagation.

Any field configuration $\vec{E}(\vec{r},t)$ as a function of space and time coordinates can be represented by a Fourier integral:

$$\vec{E}(\vec{r},t) = \int_{-\infty}^{\infty} \vec{E}_\omega(\vec{r},\omega) e^{-i\omega t} d\omega. \tag{1.4}$$

We define the positive frequency part of this integral as:

$$\vec{E}^{(+)}(\vec{r},t) = \int_{0}^{\infty} \vec{E}_\omega(\vec{r},\omega) e^{-i\omega t} d\omega, \tag{1.5}$$

and the negative frequency part:

$$\vec{E}^{(-)}(\vec{r},t) = \int_{-\infty}^{0} \vec{E}_\omega(\vec{r},\omega) e^{-i\omega t} d\omega = \int_{0}^{\infty} \vec{E}^*_\omega(\vec{r},\omega) e^{i\omega t} d\omega, \tag{1.6}$$

where we used the property $\vec{E}_\omega(\vec{r},-\omega) = \vec{E}^*_\omega(\vec{r},\omega)$, which follows from the fact that electric field $\vec{E}(\vec{r},t)$ is a real-valued function.

All subsequent analyses will be conducted for the complex positive frequency part $\vec{E}^{(+)}(\vec{r},t)$ of the electric field, bearing in mind that the real field can be expressed as:

$$\vec{E}(\vec{r},t) = \vec{E}^{(+)}(\vec{r},t) + \vec{E}^{(-)}(\vec{r},t), \tag{1.7}$$

and that $\vec{E}^{(-)}(\vec{r},t) = (\vec{E}^{(+)}(\vec{r},t))^*$, as follows from eq. (1.4) and eq. (1.6). By construction of the positive-frequency part of the electric field $\vec{E}^{(+)}(\vec{r},t)$, it is also a solution of Maxwell's equations.

Given the linearity of Maxwell's equations, any linear combination of its solutions is also a solution. If we confine our analysis to cases where the field of interest is enclosed in a spatial box of size V, we can identify a discrete set of orthonormalized solutions $\{\vec{f}_m(\vec{r},t)\}$ that can serve as a basis for decomposing any other solution satisfying the chosen boundary conditions. Each individual solution $\vec{f}_m(\vec{r},t)$ from

this set is called a *mode of the electromagnetic field* and follows the Maxwell equations

$$\left(\Delta - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\vec{f}_m(\vec{r},t) = 0, \quad \vec{\nabla}\cdot\vec{f}_m(\vec{r},t) = 0, \tag{1.8}$$

and orthonormality conditions

$$\frac{1}{V}\int_V d\vec{r}\,\vec{f}_m^*(\vec{r},t)\cdot\vec{f}_n(\vec{r},t) = \delta_{mn}. \tag{1.9}$$

If the set of functions $\{\vec{f}_m(\vec{r},t)\}$ is complete, any solution of Maxwell equations can be decomposed over it:

$$\vec{E}^{(+)}(\vec{r},t) = \sum_m \mathcal{E}_m\vec{f}_m(\vec{r},t). \tag{1.10}$$

The set of complex amplitudes $\mathcal{E}_m$ fully defines the state of the classical electromagnetic field.

1.1.1 Mode transformation

As with any linear space, we are not limited to one specific basis, i.e., a chosen set of modes. If we perform a unitary transformation:

$$\vec{g}_n(\vec{r},t) = \sum_m U_{nm}\vec{f}_m(\vec{r},t), \tag{1.11}$$

with U_{nm} being elements of a unitary matrix, we arrive at another complete orthonormal mode basis $\{\vec{g}_n(\vec{r},t)\}$. Decomposing the field over this new basis we arrive at

$$\vec{E}^{(+)}(\vec{r},t) = \sum_m \mathcal{G}_m\vec{g}_m(\vec{r},t), \tag{1.12}$$

where new amplitudes $\mathcal{G}_m$ can be found as

$$\mathcal{G}_n = \sum_m U_{nm}^*\mathcal{E}_m. \tag{1.13}$$

One can easily show that elements U_{nm} of the unitary matrix can be expressed as

$$U_{nm} = \frac{1}{V}\int_V d\vec{r}\;\vec{g}_n(\vec{r},t)\cdot\vec{f}_m^*(\vec{r},t). \tag{1.14}$$

1.1.2 Transverse modes

For many practical problems, such as imaging for example, a common approach involves considering fields that result from the superposition of plane waves with wave vectors close to a mean value $\vec{k}_0$ (paraxial approximation) and frequencies close to a central frequency $\omega_0 = c|\vec{k}_0|$ (narrow-band approximation). We choose our coordinate system such that $\vec{k}_0$ is parallel to the z-axis, which will coincide with the optical axis of the imaging apparatus. We consider the field with fixed linear

polarization, which can be effectively described using a scalar function

$$E^{(+)}(\vec{r},t) = e^{i(k_0 z-\omega_0 t)} \sum_m \mathcal{V}_m v_m(\vec{r},t), \tag{1.15}$$

where $e^{i(k_0 z-\omega_0 t)}$ represents the carrier plane wave, and $v_m(\vec{r},t)$ are the envelope functions of the different modes, which are slowly varying functions of time at the scale of the optical period $2\pi/\omega_0$ and of the position at the scale of the wavelength.

In the following, we constrain the basis selection to cases that are factorized into transverse and longitudinal components. Specifically, when describing light emitted during a brief time period between t_m and $t_m+\tau$ by sources with fixed properties within this time, located in the plane $z=0$, one can use the mode basis in the form:

$$v_m(\vec{r},t) = u_m(x,y,z)\ \mathrm{rect}_{[t_m,t_m+\tau]}(t-z/c), \tag{1.16}$$

where we define a normalized rectangular function as

$$\mathrm{rect}_{[a,b]}(x) = \begin{cases} 0, & \text{if } x<a \\ \dfrac{1}{\sqrt{b-a}}, & \text{if } a<x<b \\ 0, & \text{if } x>b. \end{cases} \tag{1.17}$$

The function $u_m(x,y,z)$ defines the spatial shape of mode $v_m(\vec{r},t)$, while the function rect $[t_m,t_m+\tau](t-z/c)$ describes its temporal shape and propagation along the z-axis. The set of functions $\{u_m(x,y,z)\}$ is not linearly independent. For example, the two modes $v_{m_1}(\vec{r},t)$ and $v_{m_2}(\vec{r},t)$ can have the same spatial structure $u_{m_1}(x,y,z) = u_{m_2}(x,y,z)$ but be localized at different time moments $|t_{m_1}-t_{m_2}| \geq \tau$.

To identify modes satisfying the constraints outlined in eq. (1.8), one can choose an orthonormal set of functions $f_m(x,y)$ in any transverse plane $z=z_0$. These modes need to exhibit slow variations on the wavelength scale, enabling the use of the paraxial approximation. It is often convenient to normalize this function to 1. The propagation of such modes along the z-axis in the paraxial approximation can be characterized using the Huygens-Fresnel integral [Born 1999; Goodman 2005]

$$u_m(x,y,z) \propto \int dx_0\, dy_0\, f_m(x_0,y_0) \exp\left[ik_0 \frac{(x-x_0)^2+(y-y_0)^2}{2(z-z_0)}\right], \tag{1.18}$$

with proper normalization, adhering to the conditions specified in eq. (1.9).

Therefore, to establish a basis describing the electromagnetic wave emanating from a narrow-band source with time-independent properties using a paraxial approximation, it suffices to select a set of *transverse modes* denoted as $f_m(x,y)$. These modes propagate in the direction of the z-axis according to eq. (1.18) and evolve in time as governed by eq. (1.16) and eq. (1.15). Due to the linearity of the connection between $f_m(x,y)$ and $v_m(\vec{r},t)$ one can find the unitary transformation matrix between the mode basis constructed from the functions $\{f_m(x,y)\}$ and another mode basis constructed from the functions $\{g_m(x,y)\}$ (given that they are defined in the same plane $z=z_0$) as

$$U_{nm} = \int dxdy\ g_n(x,y) f_m^*(x,y). \tag{1.19}$$

A practical strategy entails selecting a set of transverse modes based on the specific geometry of the situation, encompassing only those modes that are excited.

1.1.3 Classical coherence

In the cases when the classical field is deterministically defined, its normalized spatiotemporal shape can be considered as the initial mode in a mode basis. The decomposition in eq. (1.10) then consists of a single term, as any perfectly coherent classical field essentially represents a single mode. However, in most practical scenarios, fields are not perfectly mastered; they exhibit some level of randomness or "incoherence" through amplitude and phase fluctuations in both space and time. In such cases, the complex coefficients $\mathcal{E}_m$ in eq. (1.10) become stochastic variables.

The degree of coherence of the field is characterized by the probability distributions of the complex amplitudes $\mathcal{E}_m$ and the correlations between different amplitudes $\mathcal{E}_m$ and $\mathcal{E}_{m'}$.

The elements $\Gamma^{(1)\mathcal{E}}_{mn}$ of the *first-order coherency matrix*, as defined by Wiener [Wiener 1928], in the mode basis $\{\vec{f}_m(\vec{r},t)\}$ are given by

$$\Gamma^{(1)\mathcal{E}}_{mn} = \overline{\mathcal{E}^*_m \mathcal{E}_n} \tag{1.20}$$

with the bar indicating an ensemble average. The diagonal elements of this matrix represent the energy of the modes, while non-diagonal elements indicate correlations between different modes.

When considering an arbitrary pair of modes $\vec{f}_1(\vec{r},t)$ and $\vec{f}_2(\vec{r},t)$ we define the normalized correlations between them as the *mutual coherence*:

$$\gamma = \frac{\Gamma^{(1)\mathcal{E}}_{12}}{\sqrt{\Gamma^{(1)\mathcal{E}}_{11}\Gamma^{(1)\mathcal{E}}_{22}}}. \tag{1.21}$$

This characteristic resides within the unit circle of the complex plane, i.e., $|\gamma| \leq 1$. When $|\gamma| = 1$, the coherency matrix is degenerate, and we characterize the pair of modes as *fully coherent*. In such instances, the field can be accurately described using only one mode, often called *principle mode*

$$\vec{g}(\vec{r},t) \propto \sqrt{\Gamma^{(1)\mathcal{E}}_{11}}\,\vec{f}_1(\vec{r},t) + e^{i\phi}\sqrt{\Gamma^{(1)\mathcal{E}}_{22}}\,\vec{f}_2(\vec{r},t), \tag{1.22}$$

with $e^{i\phi} = \gamma$. This holds true, for example, when amplitudes $\mathcal{E}_1$ and $\mathcal{E}_2$ are precise (non-stochastic) numbers or stochastic numbers that are entirely correlated with each other. In this case, the field configuration can be obtained from a single source using only linear operations (see section 1.2.4).

Now let us consider the modes $\vec{f}_1(\vec{r},t)$ and $\vec{f}_2(\vec{r},t)$ being two subsequent rectangular temporal modes with the same spatial structure, i.e. $u_1(x,y,z) = u_2(x,y,z)$ and $t_1 + \tau = t_2$ in eq. (1.16). If these two modes are mutually coherent ($\gamma = e^{i\phi}$), one can describe the field using only one mode $\vec{g}(\vec{r},t)$ as in eq. (1.22). When the radiation source maintains constant properties within the described period, implying identical intensity $\Gamma^{(1)\mathcal{E}}_{11} = \Gamma^{(1)\mathcal{E}}_{22}$ and no phase jumps $\gamma = e^{i\phi} = 1$, the principal mode

takes the form of a longer rectangular temporal mode:

$$g(\vec{r},t) = u_1(x,y,z)\ \mathrm{rect}_{[t_1,t_1+2\tau]}(t - z/c). \tag{1.23}$$

This process can be iterated several times, at the time range of deterministic behavior of the radiation source (absence of random phase or amplitude jumps), i.e. within the coherence time of the source τ_c, resulting in a temporal mode with a duration $\tau = \tau_c$. It is important to remember that two modes with uncorrelated complex amplitudes cannot be represented as a single mode. Consequently, in practical situations, it is often advantageous to select temporal modes with durations matching the coherence time of the field, since a description with longer non-overlapping temporal modes is not possible, and with shorter modes is not practical.

1.2 Quantum light

Let us now explore quantum fields in a vacuum, adopting the conventional approach of quantum electrodynamics [Grynberg 2010]. This approach involves introducing the electric field operator, denoted as $\hat{\vec{E}}^{(+)}(\vec{r},t)$, which serves as the quantum extension of the classical complex field $\vec{E}^{(+)}(\vec{r},t)$. This operator can be expanded into the basis of monochromatic plane wave modes, represented by $\vec{v}_l(\vec{r},t)$:

$$\hat{\vec{E}}^{(+)}(\vec{r},t) = \sum_l \mathcal{E}_l^{(1)} \hat{a}_l \vec{v}_l(\vec{r},t), \tag{1.24}$$

$$\vec{v}_l(\vec{r},t) = \vec{\epsilon}_l e^{i(\vec{k}_l\cdot\vec{r}-\omega_l t)}, \quad \mathcal{E}_l^{(1)} = \sqrt{\frac{\hbar\omega_l}{2\epsilon_0 V}}, \tag{1.25}$$

where ϵ_l is a unit polarization vector, $\mathcal{E}_l^{(1)}$ is the single-photon electric field and $\hat{a}_l$ is the annihilation operator of a photon in the plane wave modem. These operators adhere to the commutation relations:

$$[\hat{a}_m, \hat{a}_{m'}^\dagger] = \delta_{mm'}\ , \quad [\hat{a}_m, \hat{a}_{m'}] = 0, \tag{1.26}$$

where $\hat{A}^\dagger$ denotes the Hermitian conjugate of the operator.

1.2.1 Electric field operator in any mode basis

Let us once again employ the narrow-band approximation, where $\omega_l \approx \omega_0$. In this context, any unitary mode transformation can be expressed as:

$$\vec{f}_m(\vec{r},t) = \sum_l U_{ml} \vec{v}_l(\vec{r},t), \tag{1.27}$$

corresponds to the analogous transformation of the photon creation operators

$$\hat{b}_m^\dagger = \sum_l U_{ml} \hat{a}_l^\dagger, \tag{1.28}$$

where photon creation operator $\hat{b}_m^\dagger$ corresponds to the new mode basis $\vec{f}_m(\vec{r},t)$. The transformation of the photon annihilation operators is found through simple Hermitian conjugation of the operators:

$$\hat{b}_m = (\hat{b}_m^\dagger)^\dagger = \sum_l U_{ml}^* \hat{a}_l. \tag{1.29}$$

Further, we will identify a particular mode either by the associated annihilation operator $\hat{b}_m$ or the spatiotemporal structure $\vec{f}_m(\vec{r},t)$ of the mode (or equivalently its cross-section $f_m(x,y)$ using the approximations discussed in section 1.1.2).

The unitarity of the transformation U_{ml} guarantees the preservation of the operators' commutation relations:

$$[\hat{b}_m, \hat{b}_{m'}^\dagger] = \delta_{mm'}, \quad [\hat{b}_m, \hat{b}_{m'}] = 0. \tag{1.30}$$

The total energy of the field given by

$$\hat{H} = \sum_l \hbar\omega_0 \left(\hat{a}_l^\dagger \hat{a}_l + \frac{1}{2}\right) = \sum_m \hbar\omega_0 \left(\hat{b}_m^\dagger \hat{b}_m + \frac{1}{2}\right), \tag{1.31}$$

is also preserved throughout the mode basis transformation.

1.2.2 Quantum states

Pure states

The state of an isolated quantum system is fully characterized by the normalized vector $|\psi\rangle$ belonging to a Hilbert space. The normalization condition ensures that $\langle\psi|\psi\rangle = 1$. Each physically measurable quantity, referred to as an *observable* in quantum physics, is associated with a self-adjoint operator $\hat{A}^\dagger = \hat{A}$ acting on that Hilbert space. The potential outcomes of a measurement are determined by the eigenvalues λ_i of the operator $\hat{A}$. According to the Born rule, the probability of measuring a specific eigenvalue λ_i is given by the mean value of the projection $\hat{P}_i$ onto the eigenspace of $\hat{A}$ corresponding to λ_i:

$$p(\lambda_i) = \langle\psi|\hat{P}_i|\psi\rangle. \tag{1.32}$$

The simplest example of the state of the optical mode $\hat{a}$ is the minimal energy state, known as the *vacuum state* $|0\rangle$, corresponding to the ground state of the system. No photons can be subtracted from the vacuum state, i.e.

$$\hat{a}|0\rangle = 0. \tag{1.33}$$

Applying the photon creation operators $\hat{a}^\dagger$ to the vacuum state generates a sequence of *Fock states*:

$$|n\rangle = \frac{(\hat{a}^\dagger)^n}{\sqrt{n!}}|0\rangle. \tag{1.34}$$

These states are the eigenstates of the photon number operator $\hat{N} = \hat{a}^\dagger\hat{a}$ and the field Hamiltonian eq. (1.31), representing a fixed number of photons in the mode. Measuring the number of photons in this state will always yield the exact result n.

The set of Fock states constitutes the complete basis of quantum states for the mode:

$$\sum_n |n\rangle\langle n| = \hat{I}, \tag{1.35}$$

where $\hat{I}$ is the identity operator.

Another important example of a quantum state is a *coherent state* $|\alpha\rangle$, describing the light generated by an ideal laser. This state is an eigenstate of the annihilation operator, corresponding to the eigenvalue α, an arbitrary complex number:

$$\hat{a}|\alpha\rangle = \alpha|\alpha\rangle. \tag{1.36}$$

This state can be obtained from a vacuum state with a *displacement operator*:

$$|\alpha\rangle = \hat{D}(\alpha)|0\rangle = \exp\left[\alpha\hat{a}^\dagger - \alpha^*\hat{a}\right]|0\rangle. \tag{1.37}$$

Two different coherent states are not orthogonal to each other:

$$\langle\beta|\alpha\rangle = \exp\left[-\frac{1}{2}(|\beta|^2 + |\alpha|^2 - 2\beta^*\alpha)\right] \neq 0. \tag{1.38}$$

The set of coherent states forms an overcomplete basis:

$$\int d^2\alpha |\alpha\rangle\langle\alpha| = \pi\hat{I}. \tag{1.39}$$

The coherent state can be expressed in the Fock basis as:

$$|\alpha\rangle = e^{-\frac{|\alpha|^2}{2}} \sum_n \frac{\alpha^n}{\sqrt{n!}}|n\rangle. \tag{1.40}$$

Using the Born rule and projectors on the Fock state $\hat{P}_n = |n\rangle\langle n|$, we can determine the photon-number statistics for this state:

$$p(n) = e^{-|\alpha|^2}\frac{(|\alpha|^2)^n}{n!}, \tag{1.41}$$

representing a Poisson distribution with parameter $\lambda = |\alpha|^2$. The mean number of detected photons in this state is $N = |\alpha|^2$, and its variance is $\Delta^2 N = N = |\alpha|^2$. This degree of uncertainty in the number of photons, where the variance equals the mean, is commonly known as *shot noise* or *Poisson noise*.

Mixed states

If the state of the mode is not perfectly controlled, i.e. there is a set of classical probabilities p_k (with $\sum p_k = 1$) of obtaining the mode in the states $|\psi_k\rangle$, the measurement statistics can be described using the *density operator* $\hat{\rho}$:

$$\hat{\rho} = \sum_k p_k |\psi_k\rangle\langle\psi_k|. \tag{1.42}$$

By construction, the density operator $\hat{\rho}$ is Hermitian, positive semi-definite and has a unit trace:

$$\hat{\rho}^\dagger = \hat{\rho}, \quad \hat{\rho} \geq 0, \quad \mathrm{Tr}\,\hat{\rho} = 1. \tag{1.43}$$

The Born rule for the density operator description can be obtained using the full probability formula:

$$p(\lambda_i) = \sum_k p_k \langle \psi_k | \hat{P}_i | \psi_k \rangle = \mathrm{Tr}(\hat{P}_i \hat{\rho}), \tag{1.44}$$

and the mean value for any operator $\hat{A}$ is defined as:

$$\langle A \rangle = \sum_k p_k \langle \psi_k | \hat{A} | \psi_k \rangle = \mathrm{Tr}(\hat{A}\hat{\rho}). \tag{1.45}$$

The trace of the squared density operator is referred to as the *purity* of the quantum state:

$$P = \mathrm{Tr}\,\hat{\rho}^2. \tag{1.46}$$

If the state can be represented as a vector $|\psi_1\rangle$, i.e., $p_1 = 1$ and $\hat{\rho}_1 = |\psi_1\rangle\langle\psi_1|$, it has unit purity:

$$P = \mathrm{Tr}\,\hat{\rho}^2 = \mathrm{Tr}\,\hat{\rho} = 1. \tag{1.47}$$

Such a state is called a pure state. In general, the purity has values smaller or equal to one $P \leq 1$, characterizing how much a state is mixed.

Another example of a situation where the system under study can only be described with the density operator is when it is *entangled* with an external system. If we consider an entangled pair of modes $\hat{a}$ and $\hat{b}$ sharing only one photon:

$$|\psi\rangle_{\hat{a}\hat{b}} = \frac{|0\rangle_{\hat{a}}|1\rangle_{\hat{b}} + |1\rangle_{\hat{a}}|0\rangle_{\hat{b}}}{\sqrt{2}}, \tag{1.48}$$

the mode $\hat{a}$ can be individually described only with a density operator:

$$\hat{\rho}_{\hat{a}} = \mathrm{Tr}_{\hat{b}}\left(|\psi\rangle\langle\psi|_{\hat{a}\hat{b}}\right) = \frac{|0\rangle\langle 0|_{\hat{a}} + |1\rangle\langle 1|_{\hat{a}}}{2}, \tag{1.49}$$

which is a maximally mixed state in the subspace of Fock states $|0\rangle_{\hat{a}}$ and $|1\rangle_{\hat{a}}$, with its purity equal to $P = 1/2$.

An important example of a mixed state is a *thermal state*, which describes the statistics of black body radiation. This state can be represented as a classical mixture of Fock states weighted with probabilities given by the Bose-Einstein distribution:

$$\hat{\rho}^{th} = \frac{1}{N+1} \sum_n \left(\frac{N}{N+1}\right)^n |n\rangle\langle n|, \tag{1.50}$$

where $N = \langle \hat{N} \rangle = \mathrm{Tr}(\hat{\rho}^{th}\hat{N})$ is the mean number of photons in the state, and the photon-number variance is

$$\Delta^2 N = N + N^2. \tag{1.51}$$

The purity of the thermal state equals $P = 1/(1+2N)$.

The photon number statistics, with the variance being higher than the average $\Delta^2 N > N$, is often referred to as *Super-Poissonian* statistics. In the limit $N \ll 1$, i.e., when the average emitted number of photons per coherence time of the thermal source is very small, one can neglect the N^2 term in eq. (1.51). In this case, the statistics of the thermal source is close to Poissonian, i.e., $\Delta^2 N \approx N$, which is called the *Poissonian limit*, or sometimes we will refer to it as *faint source approximation*. Note that natural thermal sources have a very short coherence time, and thus often can be considered using a "faint source" approximation. It's important to clarify that this does not imply a low flux of photons from these sources but rather a small number of photons emitted during the coherence time of the source, i.e. their short coherent time.

Even in the Poissonian limit, where the photon number statistics of thermal light are equivalent to those of coherent light, a fundamental difference persists between these two cases, arising from the phase noise of the thermal state. This difference is described in detail in the following section.

1.2.3 Quantum coherency matrix

The quantum coherency matrix $\Gamma^{(1)\hat{a}}_{mn}$ is an extension of the classical coherency matrix given by eq. (1.20). In what follows, we will refer to it simply as the coherency matrix. In a given mode basis $\vec{u}_l(\vec{r}, t)$ with corresponding field operators $\hat{a}_l$, it has elements:

$$\Gamma^{(1)\hat{a}}_{mn} = \left\langle \hat{a}^\dagger_m \hat{a}_n \right\rangle . \tag{1.52}$$

The diagonal part of this matrix defines the average numbers of photons at each mode $\Gamma^{(1)\hat{a}}_{mm} = \langle \hat{N}_m \rangle = \langle \hat{a}^\dagger_m \hat{a}_m \rangle$, while the non-diagonal elements are responsible for the field correlations.

The mutual coherence between the two modes $\hat{a}_1$ and $\hat{a}_2$ is defined analogously to a classical case eq. (1.21):

$$\gamma = \frac{\Gamma^{(1)\hat{a}}_{12}}{\sqrt{\Gamma^{(1)\hat{a}}_{11}\Gamma^{(1)\hat{a}}_{22}}} = \frac{\langle \hat{a}^\dagger_1 \hat{a}_2 \rangle}{\sqrt{\langle \hat{a}^\dagger_1 \hat{a}_1 \rangle \langle \hat{a}^\dagger_2 \hat{a}_2 \rangle}}. \tag{1.53}$$

Note that γ is a complex number, and its complex phase defines the phase difference between the modes $\hat{a}_1$ and $\hat{a}_2$. Thus, if one swaps the order of modes to the opposite one $\hat{a}_1 \leftrightarrow \hat{a}_2$, the mutual coherence will change to its conjugate $\gamma \to \gamma^*$.

Using eqs. (1.28) and (1.29), one can observe that in another mode basis $\{\hat{b}_l\}$, the coherency matrix takes the form:

$$\Gamma^{(1)\hat{b}}_{mn} = \left\langle \hat{b}^\dagger_m \hat{b}_n \right\rangle = \sum_{m'n'} U_{mm'} \Gamma^{(1)\hat{a}}_{m'n'} U^*_{n'n} \ , \tag{1.54}$$

i.e. it changes like a normal matrix under the transformation $U_{mm'}$.

Now, let us consider a coherency matrix of a two-mode coherent state. If the phase between the modes $\hat{a}$ and $\hat{b}$ is fixed the two-mode coherent state can be written

as $|\psi\rangle_{\hat{a}_1\hat{a}_2} = |\alpha\rangle_{\hat{a}_1}|\beta\rangle_{\hat{a}_2}$. The coherency matrix of this state is

$$\mathbf{\Gamma}^{(1)\hat{a}} = \begin{pmatrix} |\alpha|^2 & \alpha^*\beta \\ \alpha\beta^* & |\beta|^2 \end{pmatrix}. \tag{1.55}$$

This is a degenerate matrix, and mutual coherence between the two modes has a unit absolute value $|\gamma| = 1$. One can find a transformation of the form of eq. (1.22), that diagonalizes the coherency matrix, such that only one mode of a new basis $\{\hat{b}_l\}$ is excited:

$$\mathbf{\Gamma}^{(1)\hat{b}} = \begin{pmatrix} |\alpha|^2 + |\beta|^2 & 0 \\ 0 & 0 \end{pmatrix}. \tag{1.56}$$

Thus a pair of modes $\hat{a}_{1,2}$ in coherent states can be represented as a single mode $\hat{b}_1$, which is called a *principal mode* of the field. The same is true for an arbitrary number of coherent modes, i.e. coherent field is a single mode in essence. In this way the coherent states are always mutually coherent to each other, however, it does not mean that the pair of modes can be mutually coherent without being in the coherent state (see section 1.2.4).

Now let us consider a pair of independent thermal emitters $\hat{\rho}_{\hat{a}_1\hat{a}_2} = \hat{\rho}^{th}_{\hat{a}_1}\hat{\rho}^{th}_{\hat{a}_2}$. The coherency matrix in this case is already diagonal and non-degenerate

$$\mathbf{\Gamma}^{(1)\hat{a}} = \begin{pmatrix} N_1 & 0 \\ 0 & N_2 \end{pmatrix}, \tag{1.57}$$

where $N_{1,2}$ are the mean numbers of photons in the modes $\hat{a}_{1,2}$ respectively. One can see, that the pair of thermal sources does not have mutual coherence, i.e. $\gamma = 0$, and this field configuration cannot be represented as a single mode.

Thus, the faint coherent state $|\alpha\rangle_{\hat{a}}$ and the faint thermal state $\hat{\rho}^{th}_{\hat{a}}$ can have almost identical photon number statistics. But having a phase reference or a pair of each state one can observe the difference between these two cases, coming from the coherence properties.

1.2.4 Optical elements and detection

In this section, we explore some fundamental optical elements and their quantum optics descriptions.

Phase shifter

The most basic passive linear optical element is a phase shifter, characterized by the unitary transformation:

$$\hat{R}(\phi) = e^{i\phi\hat{a}^\dagger\hat{a}}. \tag{1.58}$$

In the Heisenberg picture, the action of this element leads to the following operator transformation:

$$\hat{a} \to \hat{R}^\dagger(\phi)\hat{a}\hat{R}(\phi) = e^{i\phi}\hat{a}, \quad \hat{a}^\dagger \to \hat{R}^\dagger(\phi)\hat{a}^\dagger\hat{R}(\phi) = e^{-i\phi}\hat{a}^\dagger. \tag{1.59}$$

Consequently, the phase shift introduces a complex phase factor to the field operators.

Beamsplitter

Another important optical element is a linear beamsplitter. This device has two input and two output ports. If two modes $\hat{a}$ and $\hat{b}$ have the same spatial and temporal profile, their interaction on a beamsplitter is described by the operator

$$\hat{T}_{ab} = e^{i\theta(\hat{a}^\dagger \hat{b} + \hat{a}\hat{b}^\dagger)}, \tag{1.60}$$

where the parameter θ characterizes the beamsplitter. This operation leads to the following operator transformation in the Heisenberg picture:

$$\begin{aligned} \hat{a}_{out} &= \hat{T}_{ab}^\dagger \hat{a} \hat{T}_{ab} = \hat{a}\cos\theta + i\hat{b}\sin\theta, \\ \hat{b}_{out} &= \hat{T}_{ab}^\dagger \hat{b} \hat{T}_{ab} = \hat{b}\cos\theta + i\hat{a}\sin\theta. \end{aligned} \tag{1.61}$$

It is important to note that in practical situations, the phase shifts of the transmitted and reflected beams depend on the material and geometry of the beamsplitter. With the freedom to choose a phase reference for each mode, it is often practical to choose it in a way so that the transformation takes the form of a rotation in the space of field operators:

$$\begin{aligned} \hat{a}_{out} &= \hat{a}\cos\theta + \hat{b}\sin\theta, \\ \hat{b}_{out} &= \hat{b}\cos\theta - \hat{a}\sin\theta. \end{aligned} \tag{1.62}$$

In the following, we will stick to this choice of the phase reference.

Let us consider one important example: the combination of a pair of coherent states $|\alpha\rangle_a$ and $|\beta\rangle_b$ on a beamsplitter. Since coherent states are eigenstates of annihilation operators, they transform as follows under a beamsplitter transformation:

$$\hat{T}_{ab}|\alpha\rangle_a|\beta\rangle_b = |\alpha\cos\theta + \beta\sin\theta\rangle_a |\beta\cos\theta - \alpha\sin\theta\rangle_b. \tag{1.63}$$

This corresponds also to the transformation of classical field amplitudes on the beamsplitter.

It's important to note that the interaction of two coherent states on a beamsplitter results in two independent coherent states, represented as a direct product of states of individual modes. This signifies that any measurements conducted on modes $\hat{a}_{out}$ and $\hat{b}_{out}$ yield independent (uncorrelated) outcomes. This is a distinctive characteristic of coherent states and does not hold for the interaction of arbitrary states on a beamsplitter.

Optical losses (attenuation)

When a light beam propagates through a medium and encounters various optical elements, a portion of it can be lost through undesired reflection and absorption. If we focus only on the light that successfully passes through the optical system, the transformation describing the losses does not take the form of a unitary operation

since the system is opened in this case. In this context, losses are often characterized by the Lindblad Master Equation [Manzano 2020].

However, this non-unitary process can be represented as unitary in an expanded space of modes [Fabre 2023]. Intuitively, this can be grasped by imagining that all the lost light is reflected into some unobserved mode. Thus, by incorporating the reflected light into the model, one gets a closed system with unitary dynamics.

Let us consider the pure state $|\psi\rangle_a$ experiencing losses, characterized by the *transmissivity* factor $\kappa < 1$. It means that the mean energy of the state gets multiplied by κ:

$$E_0 = {}_a\langle\psi|\hat{H}|\psi\rangle_a \xrightarrow{loss} E_f = \kappa E_0. \tag{1.64}$$

This process can be represented in a unitary manner by introducing an additional mode $\hat{b}$ to the scenario. Assuming that all light from mode $\hat{a}$ is reflected into mode $\hat{b}$, this can be described by a straightforward beamsplitter transformation given by eq. (1.60):

$$|\psi_f\rangle_{ab} = \hat{T}_{ab}(\theta)|\psi_0\rangle_a|0\rangle_b, \tag{1.65}$$

where $\cos^2\theta = \kappa$. Typically, the mode $\hat{b}$ is considered to be initially in the vacuum state. However, if the optical system's temperature is high enough to have a non-negligible number of thermal photons, or if there is strong ambient light entering the system, the initial state of mode $\hat{b}$ may be considered as, for example, a thermal state. This process is called *thermalisation*.

It is important at this point to exclude mode $\hat{b}$ from the consideration. This exclusion is necessitated by the unavailability of lost light, making it impossible to measure or manipulate. Consequently, the state of the mode after the losses can be expressed as:

$$\hat{\rho}_a = \mathrm{Tr}_b\left[|\psi_f\rangle\langle\psi_f|_{ab}\right]. \tag{1.66}$$

In the general scenario, this state is not pure due to the presence of entanglement between modes $\hat{a}$ and $\hat{b}$ in the state $|\psi_f\rangle_{ab}$. In the exceptional case where the initial state is a coherent state $|\alpha\rangle_a$, it remains a pure coherent state $|\alpha\sqrt{\kappa}\rangle_a$ after the losses. The discussed approach is readily extendable to initially mixed states.

Often, the most convenient method of describing loss involves representing the beamsplitter transformation in the Heisenberg picture, as outlined in eq. (1.62) — that is, describing it as the transformation of the field operator rather than the states:

$$\hat{a}_{out} = \sqrt{\kappa}\hat{a} + \sqrt{1-\kappa}\hat{b}. \tag{1.67}$$

This allows to describe the transformation of the Fock state $|N_0\rangle_a$ (1.34) under losses as follows:

$$|N_0\rangle_a = \frac{(\hat{a}^\dagger)^{N_0}}{\sqrt{N_0!}}|0\rangle_a \xrightarrow{loss} |\psi_f\rangle_{ab} = \frac{(\sqrt{\kappa}\hat{a}^\dagger + \sqrt{1-\kappa}\hat{b}^\dagger)^{N_0}}{\sqrt{N_0!}}|0\rangle_{ab}. \tag{1.68}$$

Tracing out the mode $\hat{b}$, as outlined in eq. (1.66), we obtain the output state of mode $\hat{a}$ as:

$$\hat{\rho}_a = \sum_{m=0}^{N_0}\binom{m}{N_0}\kappa^m(1-\kappa)^{N_0-m}|m\rangle\langle m|_a, \tag{1.69}$$

where $\binom{m}{n} = m!/(n!(m-n)!)$ represents binomial coefficients.

Thus starting from the arbitrary state $\hat{\rho}_0$ with photon number statistics

$$p_n = \langle n|\hat{\rho}_0|n\rangle, \tag{1.70}$$

after the losses, one obtains the state $\hat{\rho}_f$ with the photon number statistics given by [Kelley 1964; Vogel 2006]

$$q_n = \langle n|\hat{\rho}_f|n\rangle = \sum_{m=n}^{\infty} \binom{m}{n} \kappa^n (1-\kappa)^{m-n} p_m, \tag{1.71}$$

This transformation is often referred to as the *Bernoulli transformation* [Kiss 1995]. It has a clear physical interpretation: for each term in the sum, starting from m photons (with corresponding probability p_m), it is multiplied by κ^n — the probability of n photons being transmitted, and $(1-\kappa)^{(m-n)}$ — the probability of $m-n$ photons being lost. The binomial coefficients take into account that photons are indistinguishable and that there are $\binom{m}{n}$ possibilities to select $m-n$ lost and n transmitted photons from m.

Photodetection

Finally, we delve into the process of light detection. For now, our focus will be on direct photodetection, while field measurements such as homodyne and heterodyne detection will be addressed in the next section. Today, a diverse range of technologies for photon detection is available, including classical semiconductor photodiodes, avalanche photodiodes, photomultiplier tubes, charge-coupled devices, superconducting nanowire single-photon detectors, quantum dot detectors, transition edge sensors, bolometers, and more. Each technology comes with its set of advantages and excels in specific conditions and requirements, such as flux and spectrum of the detected light, detection efficiency and noise level, necessary temporal and spatial resolution, operating temperature, and so forth. Nevertheless, the ultimate goal of these devices in most cases is to measure the energy of the field, essentially to count the number of photons. Many detectors used in quantum optics are so-called *on-off detectors*, meaning they do not resolve the number of photons. The response of these detectors does not depend on the number of absorbed photons if it is not zero. However, in this study, we primarily focus on photon number-resolving detectors, which produce different responses when absorbing different amounts of photons. Examples include photodiodes in the linear regime for intense light, charge-coupled devices [Lesser 2015], superconducting nanowire single-photon detectors [Esmaeil Zadeh 2021; Tao 2019], and transition edge sensors [Morais 2022], among others. In cases where the analyzed field is very weak and its two-photon component is negligible, on-off detectors do not become saturated and can be used on par with photon-number-resolving detectors.

The observable $\hat{N}(\vec{r},t)$ associated with photodetection conducted on a small area around point $\vec{r}$ and at time t on a given beam is proportional to $\hat{E}(\vec{r},t)^\dagger \hat{E}(\vec{r},t)$ [Glauber 1963; Fabre 2020]. Then, the number of photons detected in volume V_D (given by the spatial resolution of the detector) within the time interval $t \in [0, t_D]$

(temporal resolution of the detector) is given by

$$\hat{N} = \frac{1}{V}\int_{V_D} d\vec{r}\int_0^{t_D} dt \sum_{m,n} \hat{b}_m^\dagger \hat{b}_n f_m^*(\vec{r},t) f_n(\vec{r},t). \tag{1.72}$$

It's important to note that this observable generally contains cross-terms such as $\hat{b}_m^\dagger \hat{b}_n$. Therefore, if the detector has sufficient spatial or temporal resolution, the photodetection signal is sensitive to the coherence between different modes.

Now, let us consider the situation where the set of modes with numbers $m \in \mathbb{D}$ is entirely contained within the detection volume (V_D, t_D), while the rest of the modes are localized outside of the detection volume. In this case, the orthogonality of the modes leads to the detected number of photons given by

$$\hat{N} = \sum_{m\in\mathbb{D}} \hat{b}_m^\dagger \hat{b}_m. \tag{1.73}$$

This implies that the measurement result in this scenario solely comprises the number of photons in the modes belonging to $\mathbb{D}$ and does not depend on any correlations. Therefore, if one aims to directly measure the energy of the subset of modes $\mathbb{D}$, it is necessary to localize it in a spatial-temporal volume away from the other modes. This sorting procedure is referred to as *mode demultiplexing*. In this study, we are interested in spatial (transverse) mode demultiplexing, often referred to as SPADE [Tsang 2016], the methods of performing this operation are discussed in section 3.2.3.

In practical scenarios, the statistics of measurement outcomes may not perfectly follow the quantum statistics of the measured state. Firstly, various fluctuations within the measurement setup introduce additional noise. Secondly, detectors typically exhibit finite photodetection efficiency, which is equivalent to experiencing losses before the detection. Additionally, detectors may generate false detection events due to ambient light or thermal fluctuations in the electronics. All these effects may result in a difference in the mean detected signal, from the mean number of photons in the state. However, by appropriately calibrating the detector and assuming its parameters remain stable, one can always correct for this difference (coming from losses, dark counts, etc.) and rescale the mean measured number of photons to approximate the mean number of photons in the state, i.e., $N_{\text{detected}} = \langle \hat{N} \rangle$. It's noteworthy that this compensation often amplifies the noise in the measurement results.

Therefore, the primary challenge lies in the additional noise present in the detection statistics:

$$\Delta^2 N_{\text{detected}} = \Delta^2 N + \Delta^2 N_{EN}. \tag{1.74}$$

The term $\Delta^2 N_{EN}$ is often referred to as *detection noise* or *electronic noise*. Properly accounting for this noise is often critical in metrological applications.

1.3 Continuous variable representation of quantum states

In a general scenario, the density operator $\hat{\rho}$ (or the state vector $|\psi\rangle$) characterizing the quantum state of the mode possesses infinite dimensions when expressed in a

discrete basis. Consequently, it may not always be the most practical tool, particularly in numerical computations. Rather than employing a discrete basis, one can opt for the eigenvectors of operators with a continuous spectrum. In this section, we introduce such operators, commonly utilized in quantum optics, and elucidate the continuous variable (CV) representation of quantum states.

1.3.1 Quadrature operators

In classical optics, the state of the mode $\vec{f}_m(\vec{r},t)$ is entirely described by the complex amplitude $\mathcal{E}_m$ introduced in eq. (1.10), encompassing its real and imaginary parts. In quantum optics, corresponding observables known as *quadrature operators* are constructed:

$$\hat{q}_m = \hat{a}_m + \hat{a}_m^\dagger, \tag{1.75}$$

$$\hat{p}_m = i(\hat{a}_m^\dagger - \hat{a}_m). \tag{1.76}$$

Using these notations, the electric field operator takes the form:

$$\hat{\vec{E}}^{(+)}(\vec{r},t) = \sum_l \mathcal{E}_l^{(1)} \frac{\hat{q}_m + i\hat{p}_m}{2} \vec{u}_l(\vec{r},t) \tag{1.77}$$

The quadrature operators are Hermitian and correspond to observables. Each of them has a continuous spectrum

$$\hat{q}_m |q\rangle_m = q|q\rangle_m, \tag{1.78}$$

that constitutes a full continuous orthonormal basis

$${}_m\langle q|q'\rangle_m = \delta(q-q'), \tag{1.79}$$

$$\int dq\ |q\rangle\langle q|_m = \mathbb{1}. \tag{1.80}$$

Two quadrature operators of the same mode do not commute:

$$[\hat{q}_m, \hat{p}_m] = \hat{q}_m\hat{p}_m - \hat{p}_m\hat{q}_m = 2i. \tag{1.81}$$

Thus, following the Heisenberg uncertainty principle, they cannot be defined simultaneously:

$$\Delta q_m^2 \Delta p_m^2 \geq \frac{1}{4} |\langle [\hat{q}_m, \hat{p}_m] \rangle|^2 = 1. \tag{1.82}$$

The quadrature operators of different orthogonal modes commute, i.e. for $m \neq n$:

$$[\hat{q}_m, \hat{p}_n] = [\hat{q}_m, \hat{q}_n] = [\hat{p}_m, \hat{p}_n] = 0. \tag{1.83}$$

The generalized quadrature operator can defined as

$$\hat{q}_m^\theta = \hat{q}_m \cos\theta + \hat{p}_m \sin\theta = \hat{a}_m e^{-i\theta} + \hat{a}_m^\dagger e^{i\theta}. \tag{1.84}$$

The commutation relation for this case can be written as

$$[\hat{q}_m^{\theta}, \hat{q}_m^{\theta+\pi/2}] = 2i. \tag{1.85}$$

One can observe that the phase shift $\hat{R}(\theta)$ defined in eq. (1.58), transforms the $\hat{q}$ quadrature into a generalized quadrature $\hat{q}_m^{\theta}$. In other words, a phase shift induces a rotation in the phase space.

For a single-mode coherent state $|\alpha\rangle$, the mean values of the quadratures are:

$$\langle \alpha | \hat{q} | \alpha \rangle = 2 \operatorname{Re} \alpha, \langle \alpha | \hat{p} | \alpha \rangle = 2 \operatorname{Im} \alpha. \tag{1.86}$$

Note that different normalization conventions for the definition of quadratures exist. Here, we intentionally use the one that leads to the variances of the two quadratures of the coherent state being

$$\Delta^2 q = \langle \alpha | \hat{q}^2 | \alpha \rangle - \langle \alpha | \hat{q} | \alpha \rangle^2 = 1, \tag{1.87}$$

$$\Delta^2 p = \langle \alpha | \hat{p}^2 | \alpha \rangle - \langle \alpha | \hat{p} | \alpha \rangle^2 = 1. \tag{1.88}$$

As evident, the variances of both quadratures are unity and independent of the state's amplitude in the case of a coherent state. Therefore, the coherent state can be fully characterized through the mean values of its quadratures. However, this is not the case for more general states.

Besides, the coherent states achieve the saturation of the inequality (1.82), signifying the minimum attainable product of quadrature uncertainties. These findings also encompass the vacuum state $|0\rangle$ since it can be regarded as a coherent state with zero amplitude. Consequently, the variances of both quadratures in a vacuum state also equal 1. Remarkably, for some states, the variance of certain quadratures can be less than 1, indicating that the quadrature noise can be smaller than in a vacuum state. Such states are referred to as *squeezed states*.

Squeezed states

In a broad sense, any state with the smallest variance of generalized quadrature below the vacuum noise level is called a squeezed state:

$$\min_{\theta} \Delta^2 q^{\theta} = \min_{\theta} \left[\langle (\hat{q}^{\theta})^2 \rangle - \langle \hat{q}^{\theta} \rangle^2 \right] < 1. \tag{1.89}$$

The corresponding minimum quadrature q^{θ_s} is termed a squeezed quadrature. This advantageous property always comes at the expense of the orthogonal quadrature, which becomes "antisqueezed", ensuring the uncertainty inequality:

$$\Delta^2 q^{\theta_s} \Delta^2 q^{\theta_s+\pi/2} \geq 1. \tag{1.90}$$

To generate a squeezed state, one must leverage a non-linear process, such as

optical parametric amplification [Andersen 2016]. The squeezing process can be effectively described using the Stoler squeeze operator [Stoler 1970]:

$$\hat{S}(\xi) = \exp\left[-\frac{1}{2}\left(\xi(\hat{a}^\dagger)^2 - \xi^*\hat{a}^2\right)\right], \tag{1.91}$$

where the complex number $\xi = re^{i2\theta_s}$ determines the degree of squeezing $r > 0$ and the orientation of the squeezed quadrature θ_s. Since the operator $\hat{S}(\xi)$ is unitary, it preserves the purity of the state. The state

$$|\xi\rangle = \hat{S}(\xi)|0\rangle \tag{1.92}$$

is known as the *squeezed vacuum*. The mean quadratures of this state are zero,

$$\langle\xi|\hat{q}|\xi\rangle = 0, \quad \langle\xi|\hat{p}|\xi\rangle = 0, \tag{1.93}$$

similar to a vacuum state. However, the variances of the quadratures differ from the vacuum noise:

$$\Delta^2 q = \left\langle \xi \left|\hat{q}^2\right| \xi\right\rangle - \left\langle \xi \left|\hat{q}\right| \xi\right\rangle^2 = e^{-2r}, \tag{1.94}$$

$$\Delta^2 p = \left\langle \xi \left|\hat{p}^2\right| \xi\right\rangle - \left\langle \xi \left|\hat{p}\right| \xi\right\rangle^2 = e^{2r}, \tag{1.95}$$

where the direction of squeezing is chosen to be $\theta_s = 0$. Since $r > 0$ the $\hat{q}$ quadrature is squeezed, while the $\hat{p}$ quadrature is antisqueezed. The product of the variances satisfies the Heisenberg inequality, establishing the squeezed vacuum as a *minimum uncertainty state*. It's essential to note that the squeezing operation does not conserve the energy of the system:

$$\langle\xi|\hat{a}^\dagger\hat{a}|\xi\rangle = \sinh^2 r \neq 0, \tag{1.96}$$

indicating that the squeezing process requires an additional source of energy.

To analyze a general squeezed state, it is instructive to examine the action of the Stoler operator $\hat{S}^\dagger$ on the quadrature operators. Assuming $\theta_s = 0$, the following transformation is observed:

$$\hat{S}^\dagger(\xi)\hat{q}\hat{S}(\xi) = e^{-r}\hat{q}, \quad \hat{S}^\dagger(\xi)\hat{p}\hat{S}(\xi) = e^{r}\hat{p}. \tag{1.97}$$

This indicates that the operator $\hat{S}(\xi)$ squeezes the phase space in one direction and expands it in another direction.

1.3.2 Wigner function

Utilizing a complete continuous basis of states with well-defined quadrature $\hat{q}$ allows for the decomposition of the density operator $\hat{\rho}$. However, this approach is not frequently employed in practical applications. The basis vectors $|q\rangle$ do not represent physical states (as they do not belong to a Hilbert space). In experiments, measurements typically involve not only the quadrature $\hat{q}$ but also the quadrature $\hat{p}$ and their linear combinations. Therefore, a robust representation should symmetrically incorporate all quadratures to enhance operational significance. This advantage is

commonly attained by employing a quasiprobability distribution to characterize the quantum state.

In the single-mode case, the quasiprobability distribution is a function of real coordinates (q, p) in the phase space. One of the most useful quasiprobability distributions is the Wigner function. It can be mapped with the density operator with the following transformation

$$W_{\hat{\rho}}(q,p) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ivp}\,\langle q-v|\,\hat{\rho}\,|q+v\rangle\,dv \tag{1.98}$$

where the state $|q+v\rangle$ is an eigenstate of the quadrature operator $\hat{q}$ with corresponding eigenvalue $q+v$.

The Wigner function is real and normalized to 1:

$$\int dq\,dp\,W_{\hat{\rho}}(q,p) = Tr\hat{\rho} = 1. \tag{1.99}$$

Its marginal distributions give the probability distributions of the quadratures q and p

$$\int dp\,W_{\hat{\rho}}(q,p) = \langle q|\hat{\rho}|q\rangle = P(q) \tag{1.100}$$

$$\int dq\,W_{\hat{\rho}}(q,p) = \langle p|\hat{\rho}|p\rangle = P(p). \tag{1.101}$$

Note that both quadratures cannot be measured at the same time, since their operators do not commute. Thus, the Wigner function is a quasiprobability but not a joint probability distribution. For some states $W_{\hat{\rho}}(q,p)$ can get negative values.

In general, the mean value of an arbitrary operator $\hat{A}$ can be calculated as

$$\langle\hat{A}\rangle_{\hat{\rho}} = \mathrm{Tr}(\hat{\rho}\hat{A}) = 2\pi\int_{\mathbb{R}^2} W_{\hat{\rho}}(q,p)W_{\hat{A}}(q,p)\,dq\,dp, \tag{1.102}$$

where $W_{\hat{\rho}}(q,p)$ and $W_{\hat{A}}(q,p)$ are defined according to eq. (1.98). This property can be used to determine the elements of the density matrix ρ_{mn} in any basis $\{|v_m\rangle\}$ from the Wigner function using $\hat{A} = |m\rangle\langle n|$.

The Wigner function of a coherent state $|\alpha\rangle$ is given by

$$W_{\hat{\rho}}(q,p) = \frac{1}{2\pi}e^{-((q-2\,\mathrm{Re}\,\alpha)^2+(p-2\,\mathrm{Im}\,\alpha)^2)/2}. \tag{1.103}$$

It exhibits the shape of a symmetric two-dimensional Gaussian bell distribution with unit variance, having a mean at $(2\,\mathrm{Re}\,\alpha, 2\,\mathrm{Im}\,\alpha)$. States characterized by Gaussian Wigner functions are referred to as *Gaussian states*. Thermal states and squeezed vacuum states exhibit Gaussian Wigner functions, as illustrated in fig. 1.1. Operations that maintain the Gaussianity of the state are termed *Gaussian operations*. This encompasses linear operations (phase shifts, displacements, and beamsplitters) as well as squeezing. Losses also maintain Gaussianity, since they can be represented as a beamsplitter operation with another Gaussian state (typically vacuum).

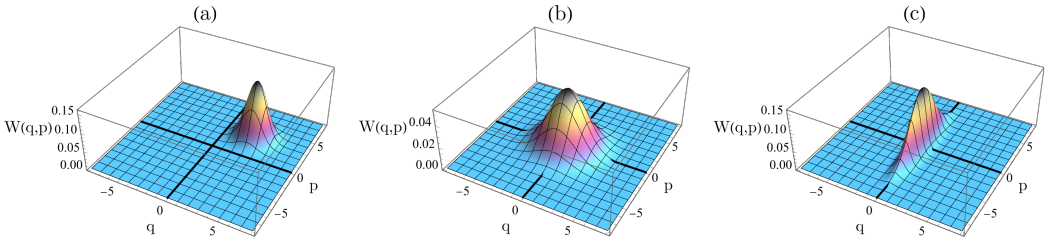

FIGURE 1.1: Wigner functions for (a) coherent state, (b) thermal state and (c) squeezed vacuum state.

In the K-mode case, the Wigner function depends on $2K$ real coordinates $\vec{R} = (q_1, q_2, ...q_K, p_1, p_2, .., p_K)^T$ of the phase space and can be defined as [Walschaers 2021]

$$W_{\hat{\rho}}(\vec{R}) = \frac{1}{(2\pi)^K} \int_{\mathbb{R}^{2K}} \mathrm{Tr}\left(\hat{\rho}\, e^{i\vec{R}'^T \vec{\hat{Q}}}\right) e^{-i\vec{R}^T\vec{R}'}\, d\vec{R}', \tag{1.104}$$

with $\vec{\hat{Q}}$ being the vector composed of the quadrature operators

$$\vec{\hat{Q}} = (\hat{q}_1, \hat{q}_2, ...\hat{q}_K, \hat{p}_1, \hat{p}_2, .., \hat{p}_K)^T. \tag{1.105}$$

In a multimode scenario, a general Gaussian state is associated with the following Wigner function:

$$W_{\hat{\rho}}(\vec{R}) = \frac{1}{(2\pi)^K \sqrt{\det \mathbf{\Gamma}_Q}} \exp\left[-\left(\vec{R} - \left\langle \vec{\hat{Q}} \right\rangle\right)^T \mathbf{\Gamma}_Q^{-1} \left(\vec{R} - \left\langle \vec{\hat{Q}} \right\rangle\right)\right]. \tag{1.106}$$

The matrix $\mathbf{\Gamma}_Q$ is commonly known as the *quadrature covariance matrix* and is defined in the next section. Hence, any Gaussian state is fully characterized by the mean values of quadratures $\left\langle \vec{\hat{Q}} \right\rangle$, commonly known as the *mean field*, and the quadrature covariance matrix $\mathbf{\Gamma}_Q$. The purity P of a Gaussian state is simply related to the covariance matrix by

$$P = \frac{1}{\det \mathbf{\Gamma}_Q}. \tag{1.107}$$

1.3.3 Quadrature covariance matrix

The quadrature covariance matrix [Simon 1994] is real $2K \times 2K$ matrix defined on a given mode basis as

$$\mathbf{\Gamma}_Q = \frac{1}{2}\left\langle \vec{\hat{Q}}\vec{\hat{Q}}^T + \left(\vec{\hat{Q}}\vec{\hat{Q}}^T\right)^T \right\rangle, \tag{1.108}$$

with $\vec{\hat{Q}}$ defined in eq. (1.105). Using this notation, the multimode version of the Heisenberg inequality can be formulated as

$$\mathbf{\Gamma}_Q + \begin{pmatrix} \mathbf{0}_K & i\,\mathbb{1}_K \\ -i\,\mathbb{1}_K & \mathbf{0}_K \end{pmatrix} > 0, \tag{1.109}$$

where $\mathbf{0}_K$ and $\mathbb{1}_K$ represent the zero and identity matrices of size $K \times K$, and the inequality indicates that the matrix is positive-definite.

Due to the structure of the covariance matrix, combining the Williamson [Arvind 1995] and Bloch-Messiah [Bloch 1962; Braunstein 2005] reductions, it can be represented as

$$\mathbf{\Gamma}_Q = \mathbf{O}_1 \mathcal{K} \mathbf{O}_2 \mathbf{\Gamma}_W \mathbf{O}_2^T \mathcal{K} \mathbf{O}_1^T, \tag{1.110}$$

where the diagonal matrix

$$\mathbf{\Gamma}_W = \mathrm{diag}(\kappa_1, ..., \kappa_K, \kappa_1, ..., \kappa_K), \tag{1.111}$$

is composed from the Williamson eigenvalues $\kappa_i \geq 1$. A quadrature covariance matrix equal to $\mathbf{\Gamma}_W$ would correspond to the thermal state of all modes, with the parameters κ_i defining the thermal energy of each mode. In the case of Gaussian states this matrix defines the purity of the state $P = 1/\prod_i \kappa_i$.

The matrix $\mathcal{K}$ is also diagonal

$$\mathcal{K} = \mathrm{diag}(\sigma_1, ..., \sigma_K, \sigma_1^{-1}, ..., \sigma_K^{-1}), \tag{1.112}$$

with real and positive elements $\sigma_i > 0$ responsible for the squeezing of quadratures. It is evident that the transformation $\mathcal{K}$ squeezes one of the quadratures for each mode and expands the orthogonal quadrature. In essence, it describes a squeezing operation.

The orthogonal matrices $\mathbf{O}_{1,2}$ are defined as

$$\mathbf{O}_{1,2} = \begin{pmatrix} \mathrm{Re}\ \mathbf{U}_{1,2} & \mathrm{Im}\ \mathbf{U}_{1,2} \\ -\mathrm{Im}\ \mathbf{U}_{1,2} & \mathrm{Re}\ \mathbf{U}_{1,2} \end{pmatrix}, \tag{1.113}$$

where $\mathbf{U}_{1,2}$ are the unitary matrices $K \times K$. These matrices describe general basis transformation between the modes and can be represented with a set of beamsplitters and phaseshifts.

In the case of a single mode ($K = 1$), the Williamson matrix $\mathbf{\Gamma}_W$ becomes proportional to the identity matrix, and as a result, it commutes with the matrix $\mathbf{O}_2$. Consequently, the transformation $\mathbf{O}_2$ does not impact the quadrature covariance matrix in this scenario. Thus, the quadrature covariance matrix of the single mode can be parameterized using these three variables:

- Williamson eigenvalue κ, which characterizes the purity (or thermality) of the state.
- Parameter σ, responsible for the squeezing. For convenience, further we will use the parameter $s = \sigma^2$.
- Parameter ϕ_s (arising from the 1×1 unitary $\mathbf{U}_1 = e^{i\phi_s}$), defining the direction of squeezing.

1.3.4 Quadrature measurements

Another essential type of measurement in quantum optics is field measurements, providing direct access to the quadratures of the optical field. Originating initially in

the early 20th century for radio-frequency applications, this measurement technique became exceptionally valuable in the optical domain with the invention of coherent optical sources, the lasers.

Homodyne detection

The most widely employed experimental procedure for conducting quadrature measurements on an optical state is *homodyne detection*. The fundamental scheme of homodyne detection is illustrated in fig. 1.2. It involves combining the state to be measured (mode $\hat{a}$) with a reference field (mode $\hat{b}$) on a balanced beamsplitter and subsequently calculating the difference between intensity measurements at both outputs.

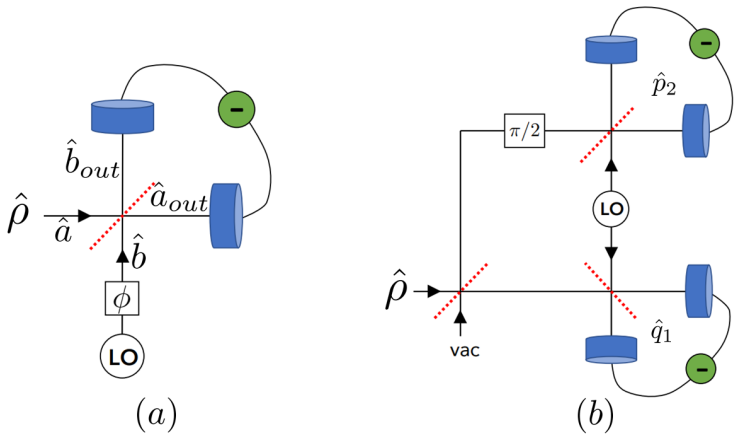

FIGURE 1.2: Schemes of the phase space measurements. (a) Homodyne detection, (b) double homodyne (heterodyne) detection. The dashed red lines represent beamsplitters, blue cylinders — the photodetectors, LO stands for Local Oscillator.
Adapted from [Chabaud 2021].

The observable operator for this measurement is given by:

$$\delta\hat{i} = \hat{a}^\dagger_{out}\hat{a}_{out} - \hat{b}^\dagger_{out}\hat{b}_{out} = \hat{a}^\dagger\hat{b} + \hat{a}\hat{b}^\dagger, \tag{1.114}$$

where the beamsplitter transformation (1.62) was employed. If the reference field, also known as a *local oscillator* (LO), is prepared in the coherent state $|\beta\rangle_b$, with $\beta = |\beta|e^{i\phi}$, then the mean value of the observable $\delta\hat{i}$ is proportional to the mean value of the generalized quadrature $\hat{q}^\phi_a$:

$$\langle\delta\hat{i}\rangle = |\beta|\langle\hat{a}^\dagger e^{i\phi} + \hat{a}e^{-i\phi}\rangle = |\beta|\langle\hat{q}^\phi_a\rangle. \tag{1.115}$$

Thus, by adjusting the phase ϕ of the LO, one can select which quadrature is being measured.

Moreover, for strong LOs $|\beta| \gg 1$ higher moments of the observable $\delta\hat{i}$ are approximately proportional to the corresponding moments of the generalized quadrature $\hat{q}_a^\phi$:

$$\langle \delta\hat{i}^n \rangle \approx |\beta|^n \langle (\hat{q}_a^\phi)^n \rangle. \tag{1.116}$$

This observation verifies that homodyne detection with a strong LO not only accurately measures the mean of the quadrature but also its variance and higher moments[1]. In essence, it projects the state onto the eigenstate $|q^\phi\rangle_a$ of the quadrature operator $\hat{q}_a^\phi$. Homodyne detection finds widespread practical applications, especially in scenarios involving squeezed states.

Double homodyne detection

Double homodyne detection, sometimes also referred to as *heterodyne detection*, involves measuring both quadratures of a single-mode state $\hat{\rho}$. This process includes splitting the state on a beamsplitter and conducting two homodyne detections, one on each output with a $\pi/2$ phase shift between them, as illustrated in fig. 1.2(b). This procedure comes at a cost — a vacuum entering the system introduces additional noise to the measurements. The measured quadrature $\hat{q}_1$ is related to the quadrature $\hat{q}$ of the state as:

$$\hat{q}_1 = \frac{\hat{q} + \hat{q}_v}{\sqrt{2}}, \tag{1.117}$$

where $\hat{q}_v$ is a quadrature of the vacuum state, i.e. $\langle \hat{q}_v \rangle = 0$, $\langle \hat{q}_v \hat{p}_v \rangle = 0$ and $\langle \hat{q}_v^2 \rangle = 1$. To account for losses on the beamsplitter, one needs to multiply the measurement results by $\sqrt{2}$:

$$\hat{q}_1' = \sqrt{2}\hat{q}_1 = \hat{q} + \hat{q}_v. \tag{1.118}$$

While the mean value of this observable coincides with the mean value of the quadrature $\hat{q}$:

$$\langle \hat{q}_1' \rangle = \langle \hat{q} \rangle, \tag{1.119}$$

the measured noise is higher than the original quadrature noise:

$$\Delta^2 q_1' = \langle \hat{q}^2 + 2\hat{q}\hat{q}_v + \hat{q}_v^2 \rangle - \langle \hat{q} \rangle^2 = \Delta^2 q + 1. \tag{1.120}$$

Similarly, it can be shown that

$$\hat{p}_2' = \sqrt{2}\hat{p}_2 = \hat{p} - \hat{p}_v, \text{ and } \Delta^2 p_2' = \Delta^2 p + 1. \tag{1.121}$$

Operators $\hat{q}_1'$ and $\hat{p}_2'$ commute, and these observables can be measured (and are measured) simultaneously.

Note that the measured quadrature covariance comes without an extra noise:

$$\mathrm{cov}(q_1', p_2') = \langle \hat{q}\hat{p} - \hat{q}\hat{p}_v + \hat{q}_v\hat{p} - \hat{q}_v\hat{p}_v \rangle - \langle \hat{q} + \hat{q}_v \rangle \langle \hat{p} - \hat{p}_v \rangle = \mathrm{cov}(q, p). \tag{1.122}$$

[1] This property is often described by the approximate equality of the operators $\delta\hat{i} \approx |\beta|\hat{q}_a^\phi$, which holds true only when averaged over strong coherent states of the LO, where $|\beta| \gg 1$.

Therefore, the measured covariance matrix $\mathbf{\Gamma}(\hat{q}_1', \hat{p}_2')$ differs from the quadrature covariance matrix by the identity matrix $\mathbf{I}$:

$$\mathbf{\Gamma}(\hat{q}_1', \hat{p}_2') = \mathbf{\Gamma}_Q + \mathbf{I}. \tag{1.123}$$

Each measurement outcome (q_1', p_2') of double homodyne detection can be combined into a complex number $\alpha = (q_1' + ip_2')/2$, effectively projecting the measured state onto a coherent state $|\alpha\rangle$. Consequently, it facilitates the sampling of the *Husimi Q function* of the state, defined as:

$$Q(\alpha) = \frac{1}{\pi}\langle\alpha|\hat{\rho}|\alpha\rangle. \tag{1.124}$$

The Husimi function is another quasiprobability distribution that fully characterizes the state $\hat{\rho}$. It is linked to the Wigner function through the transformation:

$$Q(\alpha) = \int d^2\beta \, W_{\hat{\rho}}(\beta) W_0(\beta - \alpha), \tag{1.125}$$

where $W(\alpha) = 4W_{\hat{\rho}}(q, p)$, and $\alpha = (q + ip)/2$. Here, W_0 represents the Wigner function of the vacuum state, which can be derived from eq. (1.103).

The advantage of double homodyne detection, as opposed to homodyne detection, lies in its ability to directly sample the Husimi function, which contains all the information about the state. In contrast, for homodyne detection, the Wigner function holds all the information about the state, but it cannot be directly sampled. Homodyne detection samples its marginals instead, and to fully characterize the state, one needs to vary the phase of the LO, i.e. change the quadrature being measured, as shown in section 2.4.

Chapter 2

Parameter estimation theory

Contents

The objective of numerous physical experiments is to learn specific physical characteristics of the studied system. In many cases, the characteristics of interest cannot be directly accessed and must be inferred indirectly from the measurement results. It is crucial to bear in mind the presence of inevitable noise in real measurement outcomes. This noise arises both from imperfect control of experimental conditions and the quantum nature of measurements. While the former can, in principle, be mitigated through improved experimental techniques, the latter stems from intrinsic fundamental properties of nature. However, optimizing the probing strategy can also contribute to reducing quantum noise.

In this chapter, we provide an overview of some aspects of *parameter estimation theory*. Broadly, it can be defined as a branch of statistics and information theory that addresses the estimation of unknown parameters in statistical models based on observed data. In conjunction with metrology, parameter estimation theory analyzes

potential sources of noise, and looks for optimal strategies for probing, measuring, and data processing.

The first section of this chapter introduces some general concepts of parameter estimation. The second section, drawing from the textbook [Kay 1998], articles [Gessner 2019; Gessner 2020b], and some original studies, presents key bounds in classical parameter estimation theory. Here, we introduce and discuss the primary tool employed in this research — the method of moments for parameter estimation. The third section delves into quantum parameter estimation theory, following the detailed review [Liu 2019]. In the final section of this chapter, we provide a brief overview of approaches to the characterization of quantum states, focusing particularly on continuous-variable tomography [Lvovsky 2009]. We demonstrate the similarity between quantum state characterization and parameter estimation problems, enabling us to address these issues using the same set of instruments in chapter 4.

2.1 General parameter estimation scheme

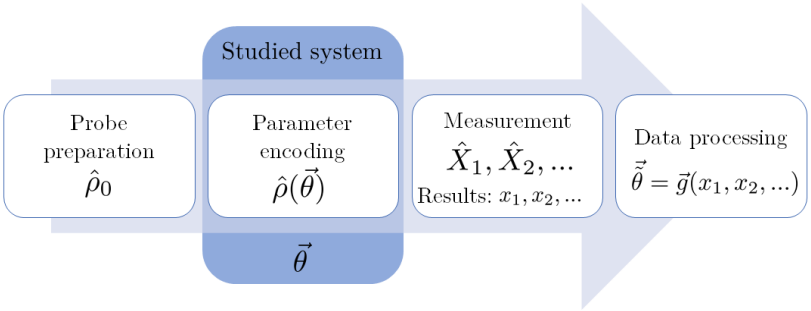

FIGURE 2.1: The fundamental steps of inferring parameters.

In a general scenario, to determine the values of parameters $\vec{\theta} = (\theta_1, \theta_2, \ldots)^T$ characterizing a physical system, a series of operations are typically performed, as schematically depicted in fig. 2.1:

- In the next step, the probe interacts with the studied system. As a result, parameters of interest $\vec{\theta}$ are encoded in the state of the probe: $\hat{\rho}_0 \to \hat{\rho}(\vec{\theta})$. The actual values of parameters $\vec{\theta}$ are referred to as the *ground truth*. We presuppose that the parameters $\vec{\theta}$, although unknown, are deterministically defined: they are fixed and can be exactly determined. Working with random parameters, which will not be covered here, is the subject of Bayesian estimation theory [Kay 1998; Trees 2001]. Nevertheless, the Bayesian approach can also provide useful techniques for estimating deterministic parameters when prior information is accessible.

- After the interaction, the probe undergoes a set of commuting measurements $\vec{\hat{X}} = (\hat{X}_1, \hat{X}_2, \ldots)^T$. It is important to note that the measurement result $\vec{x} = (x_1, x_2, \ldots)^T$ is a stochastic vector as it includes measurement noise. Choosing

an appropriate set of measurements can significantly improve the performance of the scheme. The ultimate limit for the estimation precision, which can not be overcome by detection optimization is given by the Quantum Cramér–Rao bound discussed in section 2.3.

- In the final step, the values of parameters need to be estimated from the measurement results $\vec{x}$. To achieve this, a data processing algorithm $\vec{g}(\vec{x})$ is employed. The resulting values $\vec{\tilde{\theta}} = \vec{g}(\vec{x})$ of the data processing are termed an *estimator*. Note that as the measurement results $\vec{x}$ are inherently random, the value of the estimator $\vec{\tilde{\theta}} = \vec{g}(\vec{x})$ is also random. The main purpose of the parameter estimation theory is to find an estimator that takes the measured data to a set of parameters that lies as close as possible to the ground truth.

2.1.1 Characteristics of the estimators

Let us consider the main characteristics of the estimator $\vec{\tilde{\theta}}$. Its mean deviation from the ground truth

$$\vec{B}(\vec{\tilde{\theta}}) = \langle \vec{\tilde{\theta}} - \vec{\theta} \rangle \tag{2.1}$$

is termed the *bias* of the estimator. If $\vec{B}(\vec{\tilde{\theta}}) = 0$, the estimator $\vec{\tilde{\theta}}$ is labeled as *unbiased*. As the size of the measurement sample increases, unbiased estimators tend to converge to the ground truth, whereas this is not always the case for biased estimators.

The fluctuations of the estimator are determined by the covariance matrix

$$\operatorname{cov} \vec{\tilde{\theta}} = \left\langle \left(\vec{\tilde{\theta}} - \langle \vec{\tilde{\theta}} \rangle \right) \left(\vec{\tilde{\theta}} - \langle \vec{\tilde{\theta}} \rangle \right)^T \right\rangle. \tag{2.2}$$

The diagonal elements of this matrix represent the variances of each estimated parameter:

$$\Delta^2 \tilde{\theta}_\alpha = \left[\operatorname{cov} \vec{\tilde{\theta}} \right]_{\alpha\alpha} = \left\langle \left(\tilde{\theta}_\alpha - \langle \tilde{\theta}_\alpha \rangle \right)^2 \right\rangle, \tag{2.3}$$

which signifies how much the estimations of θ_α are spread around the average value. Here and throughout, Greek indices (α, β,...) are utilized to enumerate the parameters of interest. The off-diagonal elements govern the correlation between the estimates of θ_α and θ_β.

A low variance of a biased estimator does not necessarily indicate the high quality of the estimator. For instance, the trivial estimator $\vec{\tilde{\theta}} = \vec{0}$ has zero variance, but it does not utilize measured data, and its values are not correlated with the ground truth $\vec{\theta}$.

To assess the overall efficiency of estimating parameter θ_α, both the bias and the variance can be combined into a single metric, known as the *mean squared error* (MSE):

$$\operatorname{MSE} \tilde{\theta}_\alpha = \left\langle (\tilde{\theta}_\alpha - \theta_\alpha)^2 \right\rangle = \Delta^2 \tilde{\theta}_\alpha + B^2(\tilde{\theta}_\alpha). \tag{2.4}$$

If one's primary concern is parameter θ_α, the goal should be to minimize its MSE. However, it is essential to recognize that minimizing the MSE for one parameter may come at the cost of increasing the MSE for others. In some cases, the MSE for

all parameters can be aggregated into a single value:

$$\operatorname{MSE}\vec{\tilde{\theta}} = \sum_{\alpha} \operatorname{MSE}\tilde{\theta}_{\alpha}, \tag{2.5}$$

which characterizes the overall efficiency of the estimator $\vec{\tilde{\theta}}$. It is important to note that this metric is defined up to an arbitrary normalization of each parameter θ_α, especially if different parameters have different physical senses and units, and thus, should be used with caution.

It is worth noting that in certain instances, biased estimators may offer advantages. For example, considering physical limitations on the parameters often leads to biased estimators. The resulting MSE of such an estimator is typically smaller than the MSE of its unbiased counterparts [Eldar 2008]. However, throughout this research, we primarily concentrate on unbiased estimators, as their mathematical description is simpler and more firmly established in general cases. Thus, the main characteristic of our unbiased estimator is the covariance matrix (2.2).

2.2 Classical parameter estimation theory

Classical parameter estimation theory focuses on studying estimators, which are based on the measurement results $\vec{x}$ of given observables $\vec{\hat{X}}$. One key finding in classical parameter estimation is the minimal bound on the noise of unbiased estimators — the Cramér–Rao bound (CRB).

2.2.1 Fisher information and Cramér–Rao bound

We initiate our exploration with the single-parameter estimation problem. Let $p(\vec{x}|\theta)$ be the probability density function for observing measurement outcome $\vec{x}$ given the parameter value θ, with normalization

$$\int d\vec{x}\; p(\vec{x}|\theta) = 1. \tag{2.6}$$

This function $p(\vec{x}|\theta)$ is often referred to as the *statistical model*. To quantify the information that observables $\vec{\hat{X}}$ carry about the unknown parameter θ in a given setup, one can leverage the concept of *Fisher information* (FI) [Fisher 1925]:

$$\boxed{\mathcal{F}(\theta, \vec{\hat{X}}) = \int d\vec{x}\; \frac{1}{p(\vec{x}|\theta)} \left(\frac{\partial p(\vec{x}|\theta)}{\partial \theta}\right)^2 = \int d\vec{x}\; p(\vec{x}|\theta) \left(\frac{\partial \ln p(\vec{x}|\theta)}{\partial \theta}\right)^2.} \tag{2.7}$$

The FI quantifies how quickly the probability distribution $p(\vec{x}|\theta)$ changes with small variations $\delta\theta$ in the parameter θ [Braunstein 1994]:

$$\mathcal{F}(\theta, \vec{\hat{X}}) = 8 \lim_{\delta\theta \to 0} \frac{H^2\Big(p(\vec{x}|\theta), p(\vec{x}|\theta + \delta\theta)\Big)}{\delta\theta^2}, \tag{2.8}$$

where $H^2(p(\vec{x}), q(\vec{x}))$ is the *Hellinger distance*, quantifying the similarity between the probability distributions $p(\vec{x})$ and $q(\vec{x})$ [Hellinger 1909]. It is defined as:

$$H^2(p(\vec{x}), q(\vec{x})) = \frac{1}{2}\int d\vec{x}\ \left(\sqrt{p(\vec{x})} - \sqrt{q(\vec{x})}\right)^2 = 1 - \int d\vec{x}\ \sqrt{p(\vec{x})q(\vec{x})}. \tag{2.9}$$

The physical significance of FI becomes particularly evident through the CRB, which establishes a lower limit for the variance of unbiased estimators [Cramér 1946; Rao 1994; Kay 1998]

$$\boxed{\Delta^2\tilde{\theta} \geq \frac{1}{\mu\mathcal{F}(\theta, \vec{\hat{X}})},} \tag{2.10}$$

where μ is the number of repetitions of the parameter estimation procedure outlined in fig. 2.1, i.e. using μ probes.

Proof. To prove the inequality (2.10), we start with the unbiasedness condition of the estimator $\tilde{\theta} = g(\vec{x})$:

$$B(\tilde{\theta}) = \int d\vec{x}\ (g(\vec{x}) - \theta)p(\vec{x}|\theta) = 0. \tag{2.11}$$

Taking the derivative with respect to θ, we obtain:

$$\int d\vec{x}\ (g(\vec{x}) - \theta)\frac{\partial p(\vec{x}|\theta)}{\partial\theta} - \int d\vec{x}\ p(\vec{x}|\theta) = 0. \tag{2.12}$$

Using the probability normalization condition eq. (2.6) and the chain-derivative property $\frac{\partial p}{\partial\theta} = p\,\frac{\partial \ln p}{\partial\theta}$ we get:

$$\int d\vec{x}\ (g(\vec{x}) - \theta)p(\vec{x}|\theta)\frac{\partial \ln p(\vec{x}|\theta)}{\partial\theta} = 1. \tag{2.13}$$

At this point, we can use a Cauchy-Schwarz inequality in the form

$$\left(\int d\vec{x}\ f(\vec{x})h(\vec{x})\right)^2 \leq \int d\vec{x}\ f^2(\vec{x}) \int d\vec{x}\ h^2(\vec{x}), \tag{2.14}$$

with $f(\vec{x}) = (g(\vec{x}) - \theta)\sqrt{p(\vec{x}|\theta)}$ and $h(\vec{x}) = \sqrt{p(\vec{x}|\theta)}\frac{\partial \ln p(\vec{x}|\theta)}{\partial\theta}$. This results in the following inequality:

$$\int d\vec{x}\ (g(\vec{x}) - \theta)^2 p(\vec{x}|\theta)\ \int d\vec{x}\ p(\vec{x}|\theta)\left(\frac{\partial \ln p(\vec{x}|\theta)}{\partial\theta}\right)^2 \geq 1. \tag{2.15}$$

The first integral corresponds to the MSE of the estimator, as defined in eq. (2.4). The MSE coincides with the variance when dealing with unbiased estimators. The second integral represents the FI as defined in eq. (2.7). Thus, we have successfully demonstrated that for a single iteration of the estimation procedure, the lower bound for the variance of an unbiased estimator is given by:

$$\Delta^2\tilde{\theta} \geq \frac{1}{\mathcal{F}(\theta, \vec{\hat{X}})}. \tag{2.16}$$

Now, let's examine the FI associated with two independent sets of measurements, denoted as $\vec{X}$ and $\vec{Y}$. Independence, in this context, implies that the measurement outcomes form independent random vectors $\vec{x}$ and $\vec{y}$. Consequently, the probability density of obtaining the outcomes $\vec{x}$ and $\vec{y}$ can be expressed as follows:

$$p(\vec{x},\vec{y}|\theta) = p_X(\vec{x}|\theta)p_Y(\vec{y}|\theta). \tag{2.17}$$

Computing the FI for this pair of measurements yields:

$$\begin{aligned}\mathcal{F}(\theta,(\vec{X},\vec{Y})) &= \int d\vec{x}d\vec{y}\; p_X(\vec{x}|\theta)p_Y(\vec{y}|\theta)\left(\frac{\partial \ln p_X(\vec{x}|\theta)}{\partial\theta}+\frac{\partial \ln p_Y(\vec{y}|\theta)}{\partial\theta}\right)^2 \\ &= \mathcal{F}(\theta,\vec{X})+\mathcal{F}(\theta,\vec{Y})+2\int d\vec{x}p_X(\vec{x}|\theta)\frac{\partial \ln p_X(\vec{x}|\theta)}{\partial\theta}\int d\vec{y}p_Y(\vec{y}|\theta)\frac{\partial \ln p_Y(\vec{y}|\theta)}{\partial\theta}.\end{aligned} \tag{2.18}$$

One of the last integrals can be transformed as

$$\int d\vec{x}p_X(\vec{x}|\theta)\frac{\partial \ln p_X(\vec{x}|\theta)}{\partial\theta} = \int d\vec{x}\frac{\partial p_X(\vec{x}|\theta)}{\partial\theta} = \frac{\partial}{\partial\theta}\int d\vec{x}p_X(\vec{x}|\theta) = \frac{\partial}{\partial\theta}1 = 0. \tag{2.19}$$

Thus the FI of independent measurements is additive:

$$\mathcal{F}(\theta,(\vec{X},\vec{Y})) = \mathcal{F}(\theta,\vec{X})+\mathcal{F}(\theta,\vec{Y}). \tag{2.20}$$

When employing μ probes, each followed by a set of measurements denoted as $\vec{X}$, the FI for the overall result becomes $\mu\mathcal{F}(\theta,\vec{X})$, since the measurements over distinct probes give statistically independent results. This characteristic, coupled with the inequality (2.16), proves the CRB as expressed in eq. (2.10). □

Multiparameter Cramér–Rao bound

In the context of multiparameter estimation, the information within the measurement set $\vec{X}$ is quantified using the FI matrix, defined by elements [Kay 1998]

$$\boxed{\mathcal{F}_{\alpha\beta}(\vec{\theta},\vec{X}) = \int d\vec{x}\;\frac{1}{p(\vec{x}|\vec{\theta})}\frac{\partial p(\vec{x}|\vec{\theta})}{\partial\theta_\alpha}\frac{\partial p(\vec{x}|\vec{\theta})}{\partial\theta_\beta}.} \tag{2.21}$$

The FI matrix is a symmetric positive-valued and positive-semidefinite matrix.

The CRB in the multiparameter scenario establishes a limit for the estimator covariance matrix, as expressed in eq. (2.2):

$$\boxed{\mathrm{cov}(\vec{\tilde{\theta}}) \geq \frac{1}{\mu}\boldsymbol{\mathcal{F}}^{-1}(\vec{\theta},\vec{X}),} \tag{2.22}$$

where the matrix inequality $\mathbf{A} \geq \mathbf{B}$ implies that $\mathbf{A}-\mathbf{B}$ is positive-semidefinite. Since the diagonal elements of positive-semidefinite matrix are non-negative, the

variances of the estimators $\tilde{\theta}_\alpha$ are limited by the diagonal elements of the FI matrix:

$$\Delta^2\tilde{\theta}_\alpha \geq \frac{1}{\mu}[\boldsymbol{\mathcal{F}}^{-1}]_{\alpha\alpha}. \tag{2.23}$$

The total MSE eq. (2.5) of an unbiased estimator $\vec{\tilde{\theta}}$ is constrained by the trace of the inverse FI:

$$\mathrm{MSE}\,\vec{\tilde{\theta}} \geq \frac{1}{\mu}\,\mathrm{Tr}\,\boldsymbol{\mathcal{F}}^{-1}. \tag{2.24}$$

If the FI matrix is degenerate, meaning there is no inverse matrix, it suggests that the selected set of observables $\vec{\hat{X}}$ is not adequate for estimating all parameters $\vec{\theta}$. In this case, all the measurement results are independent of some linear combination of the parameters $\vec{\theta}$, thus all parameters can not be uniquely identified simultaneously. To solve this issue, expanding the set of observables $\vec{\hat{X}}$ is necessary. This expansion may involve adding more compatible observables to $\vec{\hat{X}}$ in some cases. Alternatively, it might be more practical to select an additional set of observables $\vec{\hat{Y}}$ (or multiple sets) to be measured independently (on independent probes). Then, by utilizing the additivity property (2.20), the degeneracy of the FI matrix can be resolved.

The CRB serves as a lower bound on the variance of an unbiased estimator, offering a benchmark against which the performance of any unbiased estimator can be assessed. Achieving the CRB indicates that an estimator is *efficient*, meaning it attains the best possible precision given the available information in the data.

While finding an efficient estimator is generally challenging for finite datasets [Kay 1998; Trees 2001], the CRB can always be asymptotically (in the limit of an infinite number of measurements) saturated by the *maximum likelihood estimator* (MLE):

$$\vec{\tilde{\theta}}_{ML} = \arg\max_{\vec{\theta}} \ln p(\vec{x}|\vec{\theta}), \tag{2.25}$$

where $\arg\max_{\vec{\theta}}$ denotes the values of the parameters $\vec{\theta}$ where the function $\ln p(\vec{x}|\vec{\theta})$ has its maximum, i.e. the probability of a given outcome $\vec{x}$ is maximal for the parameters $\vec{\tilde{\theta}}_{ML}$.

The MLE is a powerful and widely used method in statistics due to its many strengths. Nevertheless, it is not without drawbacks and constraints. A primary limitation lies in its computational complexity: determining the MLE necessitates addressing intricate optimization problems, and the associated computational load can be substantial. Addressing this challenge becomes especially noticeable in high-dimensional parameter spaces, where the likelihood function (2.25) may display multiple maxima or, in some instances, lack a maximum altogether.

Another difficulty of the MLE approach lies in the complexity of building the full statistical model $p(\vec{x}|\vec{\theta})$ and computing the FI itself, in case of a large space of outcomes $\vec{x}$. Therefore, the general approach to parameter estimation, utilizing the MLE for parameter extraction and the CRB for precision benchmarking, is not universally applicable or convenient in practical scenarios. In numerous cases, opting for simpler estimators, accompanied by their own bounds, can be advantageous. A widely adopted approach with a simpler estimation procedure is the *method of moments*.

2.2.2 Method of moments

The method of moments (MoM) constructs parameter estimators from the statistical moments derived from the measured data [Kay 1998]. In our description, we focus solely on the first moments of the set of operators $\vec{\hat{X}}$, that are measured simultaneously. This set may include not only operators of physical observables but also their higher powers and various products. This approach enables us to extract all necessary statistical moments of physical observables from the first moment of $\vec{\hat{X}}$.

In the following sections, we analyze the bounds for the MoM estimators in the single- and multiparameter estimation scenarios and provide a practical way to saturate these bounds (i.e. achieve the minimal variance of the estimation).

Single observable MoM

We begin our exploration of the MoM by considering the simplest scenario: estimating a single parameter θ based on the measurement outcomes of a single observable $\hat{Y}$. When utilizing μ probes, the measurement yields outcomes $\vec{y} = (y_1, ..., y_\mu)^T$. The sample mean of these data is calculated as follows:

$$\overline{y}^{(\mu)} = \frac{1}{\mu} \sum_{i=1}^{\mu} y_i. \tag{2.26}$$

The anticipated mean value $\langle \hat{Y} \rangle_\theta$ for the observable $\hat{Y}$ can be determined theoretically through

$$\langle \hat{Y} \rangle_\theta = \mathrm{Tr}[\hat{\rho}(\theta)\hat{Y}], \tag{2.27}$$

where subscript θ reflects dependence on the parameter. Notably, finding $\langle \hat{Y} \rangle_\theta$ typically does not necessitate determining the complete measurement statistic $p(y|\theta)$. The function $\langle \hat{Y} \rangle_\theta$ is commonly referred to as the *calibration curve*, and it can also be determined experimentally by measuring a large dataset of $\hat{Y}$ across a range of unknown parameter values.

The MoM constructs an estimator $\tilde{\theta}$ by seeking a parameter value such that the expected value of the observable $\langle \hat{Y} \rangle_{\tilde{\theta}}$ matches the measured sample mean:

$$\langle \hat{Y} \rangle_{\tilde{\theta}} = \overline{y}^{(\mu)}. \tag{2.28}$$

This equality is the central core of the MoM. It is often credited to Pafnuty Chebyshev who introduced this approach in his proof of the central limit theorem back in 1887 [Fischer 2010].

The law of large numbers states that as the sample size increases, the sample mean converges to the expected value:

$$\lim_{\mu \to \infty} \overline{y}^{(\mu)} = \langle \hat{Y} \rangle_\theta. \tag{2.29}$$

This ensures that the moment-based estimator $\tilde{\theta}$ is asymptotically unbiased. However, it is essential to recognize that the unbiasedness of the moment-based estimator is not guaranteed for finite statistics μ.

Note that eq. (2.28) may have multiple solutions in general case. Hence, we typically assume that if the calibration curve is non-monotonic, then we possess preliminary information about the parameter's value range, where only one solution of eq. (2.28) exists. The degeneracy of the eq. (2.28) by itself and methods to resolve it constitute an interesting topic for further research, and are only slightly addressed in this thesis.

Finding the solution $\tilde{\theta} = g(\overline{y}^{(\mu)})$ of the algebraic equation (2.28) may require the use of numerical methods. Hence, the MoM does not always yield an explicit estimator as an analytical function of the measured data $\vec{y}$. However, the mathematical complexity of solving this single algebraic equation is typically much lower than tackling the optimization problem for the MLE (2.25) , especially when dealing with a large dataset size μ.

To assess the efficiency of the MoM estimator, we examine the propagation of uncertainty in eq. (2.28). Assuming the calibration function $\langle \hat{Y} \rangle_\theta$ is approximately linear within the confidence interval $\tilde{\theta} \pm \Delta\tilde{\theta}$, the variance of the estimator is given by:

$$\Delta^2\tilde{\theta} = \frac{\Delta^2\overline{y}^{(\mu)}}{\left(\frac{\partial\langle \hat{Y} \rangle_\theta}{\partial\theta}\right)^2}. \tag{2.30}$$

In the following, we assume that the dataset is sufficiently large, so, following the law of large numbers, the variance of the sample mean is:

$$\Delta^2\overline{y}^{(\mu)} = \frac{\Delta^2 Y}{\mu}, \tag{2.31}$$

where $\Delta^2 Y$ represents the variance of the observable $\hat{Y}$:

$$\Delta^2 Y = \mathrm{Tr}[\hat{\rho}(\theta)\hat{Y}^2] - \mathrm{Tr}[\hat{\rho}(\theta)\hat{Y}]^2. \tag{2.32}$$

Therefore, the variance of the estimator is given by:

$$\Delta^2\tilde{\theta} = \frac{1}{\mu}\frac{\Delta^2 Y}{\left(\frac{\partial\langle \hat{Y} \rangle_\theta}{\partial\theta}\right)^2} = \frac{1}{\mu\, M(\theta, \hat{X})}, \tag{2.33}$$

where we introduce

$$M(\theta, \hat{X}) = \frac{\left(\frac{\partial\langle \hat{Y} \rangle_\theta}{\partial\theta}\right)^2}{\Delta^2 Y} \tag{2.34}$$

which we refer to as the *moment-based sensitivity*, or simply *sensitivity*. In the case where only one observable is considered, it corresponds directly to the signal-to-noise ratio. This measure determines the variance of the moment-based estimator $\tilde{\theta} = g(\overline{y}^{(\mu)})$.

Note that the assumed linearity of the calibration curve within the confidence interval $\tilde{\theta} \pm \Delta\tilde{\theta}$ is assured for a sufficiently large μ. Hence, due to the convergence and unbiasedness of the estimator, the derivative $\frac{\partial\langle \hat{Y} \rangle_\theta}{\partial\theta}$ can be calculated either at the estimated parameter value $\tilde{\theta}$ or at the true value of parameter θ, depending on the specific needs. In the subsequent discussions, all expressions dependent on θ are

generally presumed to be evaluated at the ground truth values of the parameters, unless explicitly stated otherwise.

Multiobservable MoM with single parameter

The next step is to generalize the estimation with the MoM to the case where multiple observables $\vec{\hat{X}} = (\hat{X}_1, \hat{X}_2, ...)^T$ are being measured simultaneously for each probe, and a single parameter θ is estimated from the measurement results. In this context, we continue to view the estimator as a function of the sample means $\overline{x}_j^{(\mu)}$ of all observables:

$$\tilde{\theta} = f(\overline{x}_1^{(\mu)}, \overline{x}_2^{(\mu)}, ...). \tag{2.35}$$

Our objective is to construct a straightforward estimator, formulated as the solution to a single algebraic equation, akin to eq. (2.28). To achieve this, we create a linear combination of the measured sample means expressed as:

$$\overline{y}^{(\mu)} = \sum_j c_j \overline{x}_j^{(\mu)}. \tag{2.36}$$

Subsequently, we optimize the linear coefficients c_j to attain the highest precision for the estimator

$$\tilde{\theta} = g(\overline{y}^{(\mu)}). \tag{2.37}$$

In section 2.2.3 we demonstrate that this estimator, with the optimal choice of coefficients c_j, is not less precise than the general MoM estimator (2.35), in the asymptotic limit.

Due to linearity of quantum mechanics, the linear combination $\overline{y}^{(\mu)}$ can be regarded as a sample mean of observable $\hat{Y}$ given by

$$\hat{Y} = \sum_j c_j \hat{X}_j = \vec{c}^T \vec{\hat{X}}, \tag{2.38}$$

where $\vec{c} = (c_1, c_2, ...)^T$. The expectation value of $\hat{Y}$ is expressed as:

$$\langle \hat{Y} \rangle_\theta = \vec{c}^T \langle \vec{\hat{X}} \rangle_\theta, \tag{2.39}$$

and its variance is given by:

$$\Delta^2 Y = \langle \vec{c}^T \vec{\hat{X}} \vec{\hat{X}}^T \vec{c} \rangle_\theta - \langle \vec{c}^T \vec{\hat{X}} \rangle_\theta \langle \vec{\hat{X}}^T \vec{c} \rangle_\theta = \vec{c}^T \mathbf{\Gamma} \vec{c}, \tag{2.40}$$

where $\mathbf{\Gamma}$ is the covariance matrix of the observables, defined by its elements:

$$\Gamma_{jk} = \langle \hat{X}_j \hat{X}_k \rangle_\theta - \langle \hat{X}_j \rangle_\theta \langle \hat{X}_k \rangle_\theta. \tag{2.41}$$

Thus, the problem is reduced to the previous example of a single observable MoM, with the estimation variance defined in eq. (2.33):

$$\Delta^2\tilde{\theta} = \frac{1}{\mu}\frac{\Delta^2 Y}{\left(\frac{\partial\langle\hat{Y}\rangle_\theta}{\partial\theta}\right)^2} = \frac{1}{\mu}\frac{\vec{c}^T\mathbf{\Gamma}\vec{c}}{\left(\vec{c}^T\frac{\partial\langle\vec{\hat{X}}\rangle_\theta}{\partial\theta}\right)^2}. \tag{2.42}$$

The whole estimation procedure with the MoM is schematically represented in fig. 2.2: observables $\vec{\hat{X}}$ are measured μ times, yielding measurement results for each observable $\hat{X}_m$, which are aggregated to form sample means $\overline{x}_m^{(\mu)}$. These sample means are then combined with weights c_m to obtain the value $\overline{y}^{(\mu)}$. Subsequently, the parameter θ is estimated from this value using the calibration curve $\langle\hat{Y}\rangle_\theta$.

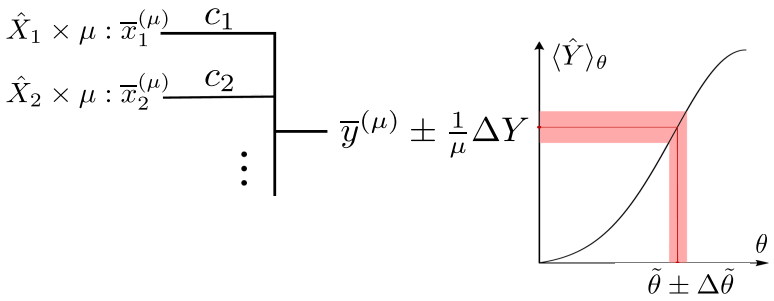

FIGURE 2.2: Schematic of the MoM estimation: observables $\vec{\hat{X}}$ are measured μ times, yielding measurement results for each observable $\hat{X}_m$, which are aggregated to form sample means $\overline{x}_m^{(\mu)}$. These sample means are then combined with weights c_m to obtain the value $\overline{y}^{(\mu)}$. Subsequently, the parameter θ is estimated from this value using the calibration curve $\langle\hat{Y}\rangle_\theta$. The red area visualizes the propagation of noise from $\overline{y}^{(\mu)}$ to the estimator $\tilde{\theta}$.
Scheme adapted from [Sorelli 2021a].

Now, to analyze the eq. (2.42), and find minimal achievable noise of the estimation, we use Cauchy-Schwarz inequality in the form

$$|\vec{u}^T\vec{v}|^2 \leq |\vec{u}^T\vec{u}||\vec{v}^T\vec{v}|, \tag{2.43}$$

and introduce the vectors

$$\vec{u} = \mathbf{\Gamma}^{1/2}\vec{c}, \quad \vec{v} = \mathbf{\Gamma}^{-1/2}\frac{\partial\langle\vec{\hat{X}}\rangle_\theta}{\partial\theta} \tag{2.44}$$

Consequently, the variance of the estimator can be expressed as:

$$\Delta^2\tilde{\theta} = \frac{1}{\mu}\frac{|\vec{u}^T\vec{u}|}{|\vec{u}^T\vec{v}|^2} \geq \frac{1}{\mu}\frac{1}{|\vec{v}^T\vec{v}|} = \frac{1}{\mu}\frac{1}{\frac{\partial\langle\vec{\hat{X}}^T\rangle_\theta}{\partial\theta}\mathbf{\Gamma}^{-1}\frac{\partial\langle\vec{\hat{X}}\rangle_\theta}{\partial\theta}}. \tag{2.45}$$

We refer to the denominator as the *moment-based sensitivity*, or simply *sensitivity*:

$$\boxed{M(\theta,\vec{\hat{X}}) = \frac{\partial\langle\vec{\hat{X}}^T\rangle_\theta}{\partial\theta}\mathbf{\Gamma}^{-1}\frac{\partial\langle\vec{\hat{X}}\rangle_\theta}{\partial\theta}.} \tag{2.46}$$

Inequality (2.45) establishes the precision bound for estimators of the form in eq. (2.37), represented as:

$$\boxed{\Delta^2\tilde{\theta} \geq \frac{1}{\mu}\frac{1}{M(\theta,\vec{\hat{X}})}.} \tag{2.47}$$

A higher value of the sensitivity $M(\theta,\vec{\hat{X}})$ corresponds to a more precise (less noisy) estimator, and vice versa. In section 2.2.3, we demonstrate that this bound is applicable to a more general estimator, as defined in eq. (2.35).

Analyzing the expression (2.46), it is evident that the sensitivity M has non-negative values due to the positivity of the covariance matrix $\mathbf{\Gamma}$. One can also show that the sensitivity of uncorrelated measurement sets $\vec{\hat{X}}^{(1)}$ and $\vec{\hat{X}}^{(2)}$ is additive, owing to the block-diagonal structure[1] of the covariance matrix $\mathbf{\Gamma}$ (and, consequently, of its inverse $\mathbf{\Gamma}^{-1}$):

$$M(\theta,(\vec{\hat{X}}^{(1)},\vec{\hat{X}}^{(2)})) = M(\theta,\vec{\hat{X}}^{(1)}) + M(\theta,\vec{\hat{X}}^{(2)}). \tag{2.48}$$

This property is especially useful in cases where measurements $\vec{\hat{X}}^{(1)}$ and $\vec{\hat{X}}^{(2)}$ are performed over independent sets of probes. This can be necessary in situations where these two sets of measurements are incompatible and cannot be measured simultaneously.

One can note, that inequality (2.47) resembles the CRB (2.10), with the sensitivity $M(\theta,\vec{\hat{X}})$ serving as an analog to the FI. However, the MoM approach considers a limited class of moment-based estimators (2.37), while the CRB encompasses all possible estimators. Thus the following inequality holds true:

$$M(\theta,\vec{\hat{X}}) \leq \mathcal{F}(\theta,\vec{\hat{X}}). \tag{2.49}$$

The next step is to find the optimal linear coefficients $\vec{c}^{opt}$, that minimize the noise of the estimation. The Cauchy-Schwarz inequality (2.43) is saturated (i.e., attains its equality) when two vectors are collinear, meaning that $\vec{u} = C_0\vec{v}$, where C_0 is any real number. This constant factor C_0 does not affect the estimator, so it is often practical to choose $C_0 = 1$. Consequently, the minimum variance of the estimator $\tilde{\theta}$ is achieved

[1] $\mathrm{cov}(\hat{X}_j^{(1)},\hat{X}_k^{(2)}) = 0, \forall j,k$

for the linear combination $\overline{y}^{(\mu)}$ with coefficients

$$\vec{c}^{opt} = \mathbf{\Gamma}^{-1} \frac{\partial \langle \vec{\hat{X}} \rangle_\theta}{\partial \theta}. \tag{2.50}$$

The estimator $\tilde{\theta}$ is constructed as a solution to the algebraic equation $\langle \hat{Y} \rangle_{\tilde{\theta}} = \overline{y}^{(\mu)}$, where the calibration curve $\langle \hat{Y} \rangle_\theta$ is defined in eq. (2.39). The experimental results are combined in a linear combination $\overline{y}^{(\mu)}$ as per eq. (2.36).

It is important to note that the optimal linear coefficients $\vec{c}^{opt}$, as defined in eq. (2.50), depend on the value of the unknown parameter θ. Therefore, the estimation procedure typically begins by establishing an initial assumption $\tilde{\theta}^{(0)}$ regarding the parameter θ. Subsequently, the linear coefficients are determined as

$$\vec{c}^{opt}(\tilde{\theta}^{(0)}) = \left(\mathbf{\Gamma}^{-1} \frac{\partial \langle \vec{\hat{X}} \rangle_\theta}{\partial \theta} \right) \Bigg|_{\tilde{\theta}^{(0)}}. \tag{2.51}$$

The coefficients $\vec{c}^{opt}(\tilde{\theta}^{(0)})$ are generally different from the optimal coefficients $\vec{c}^{opt}(\tilde{\theta})$ calculated with the true value of the parameter θ, and the corresponding estimator will not saturate the bound in eq. (2.47). However, this estimator is unbiased, and with sufficient measurement statistics μ, it produces the result $\tilde{\theta}^{(1)}$ close to the true value of parameter θ. Subsequently, one can repeat the estimation using $\tilde{\theta}^{(1)}$ as the initial assumption on the parameter value and obtain a more precise estimator value $\tilde{\theta}^{(2)}$. It is noteworthy that the same experimental data are used in both steps, requiring no additional measurements. This procedure can be repeated several times until the sequence $\tilde{\theta}^{(i)}$ converges to some value $\tilde{\theta}$ and stops alternating, typically within a very few iterations. The observed value of $\tilde{\theta}$ is then considered as the result of the iterative procedure, and its variance asymptotically saturates the bound in eq. (2.33). Note that we do not assume that the initial guess $\tilde{\theta}^{(0)}$ is close to the ground truth θ, its value only affects the number of iterations before the convergence.

Thus the MoM not only establishes the bound (2.47) for the moment-based single parameter estimator (2.35), but also offers a straightforward approach to reach this bound by combining the measured sample means $\overline{x}_j^{(\mu)}$ with optimal coefficients $\vec{c}^{opt}$ (2.50).

Multiparameter MoM

The next step is to consider the most general case: estimation of N_p parameters $\vec{\theta} = (\theta_1, ..., \theta_{N_p})^T$ from the measurement results of multiple observables $\vec{\hat{X}}$. In this situation, the quality of the estimation procedure is described by the covariance matrix of the estimator $\operatorname{cov} \vec{\tilde{\theta}}$ defined in eq. (2.2). The covariance matrix of the moment-based multiparameter estimators

$$\vec{\tilde{\theta}} = \vec{f}(\overline{x}_1^{(\mu)}, \overline{x}_2^{(\mu)}, ...) \tag{2.52}$$

is bounded by the following inequality:

$$\boxed{\mathbf{\Sigma} = \operatorname{cov}\vec{\theta} \geq \frac{1}{\mu}\boldsymbol{M}^{-1}.} \tag{2.53}$$

The *moment matrix* $\boldsymbol{M}$ is defined by its elements [Gessner 2020b]

$$\boxed{M_{\alpha\beta}(\vec{\theta}, \vec{\hat{X}}) = \frac{\partial\langle\vec{\hat{X}}^T\rangle_{\vec{\theta}}}{\partial\theta_\alpha}\mathbf{\Gamma}^{-1}\frac{\partial\langle\vec{\hat{X}}\rangle_{\vec{\theta}}}{\partial\theta_\beta},} \tag{2.54}$$

where the measurement covariance matrix $\mathbf{\Gamma}$ is defined in eq. (2.40). The proof of the inequality (2.53) is provided in section 2.2.3. The diagonal elements of this inequality give the bounds for the variances of the estimators

$$\Delta^2\tilde{\theta}_\alpha \geq \frac{1}{\mu}\left[\boldsymbol{M}^{-1}\right]_{\alpha\alpha} = \frac{1}{\mu}\frac{1}{S_\alpha}, \tag{2.55}$$

where $S_\alpha = 1/\left[\boldsymbol{M}^{-1}\right]_{\alpha\alpha}$ is referred to as sensitivity to the parameter θ_α. For the single parameter estimation $S_1 = M_{11}$, while in multiparameter scenario the sensitivity to the parameter θ_α is decreased due to possible correlation of the estimator $\tilde{\theta}_\alpha$ to other unknown parameters [Suzuki 2020]:

$$S_\alpha = \frac{1}{[\boldsymbol{M}^{-1}]_{\alpha\alpha}} \leq M_{\alpha\alpha}. \tag{2.56}$$

Now we are looking for an optimal moment-based estimator $\vec{\tilde{\theta}}$ that saturates the bound (2.53). For this one needs to build a set of linear combinations

$$\overline{y}_\alpha^{(\mu)} = \sum_j c_{\alpha j}\overline{x}_j^{(\mu)}, \tag{2.57}$$

with coefficients

$$\vec{c}_\alpha = \mathbf{\Gamma}^{-1}\frac{\partial\langle\vec{\hat{X}}\rangle_{\vec{\theta}}}{\partial\theta_\alpha}. \tag{2.58}$$

The linear combinations eq. (2.57) corresponds to sample mean of the observables $\hat{Y}_\alpha = \vec{c}_\alpha^T\vec{\hat{X}}$, which have a covariance matrix:

$$\operatorname{cov}(\hat{Y}_\alpha, \hat{Y}_\beta) = \vec{c}_\alpha^T\mathbf{\Gamma}\vec{c}_\beta = \frac{\partial\langle\vec{\hat{X}}^T\rangle_{\vec{\theta}}}{\partial\theta_\alpha}\mathbf{\Gamma}^{-1}\,\mathbf{\Gamma}\,\mathbf{\Gamma}^{-1}\frac{\partial\langle\vec{\hat{X}}\rangle_{\vec{\theta}}}{\partial\theta_\beta} = M_{\alpha\beta}. \tag{2.59}$$

Following the law of large numbers:

$$\operatorname{cov}(\overline{y}_\alpha^{(\mu)}, \overline{y}_\beta^{(\mu)}) = \frac{1}{\mu}\operatorname{cov}(\hat{Y}_\alpha, \hat{Y}_\beta) = \frac{1}{\mu}M_{\alpha\beta}. \tag{2.60}$$

To build an estimator $\vec{\tilde{\theta}}$ one needs to solve the system of equations:

$$\overline{y}_{\alpha}^{(\mu)} = \langle \hat{Y}_{\alpha} \rangle_{\vec{\tilde{\theta}}}, \quad \alpha = 1...N_p. \tag{2.61}$$

Considering the right part of this equation as a function of random variables $\vec{\tilde{\theta}}$, we use the noise propagation formula in the linear approximation, and arrive to the following expression for the covariances:

$$\operatorname{cov}(\overline{y}_{\alpha}^{(\mu)}, \overline{y}_{\beta}^{(\mu)}) = \sum_{\alpha',\beta'} \frac{\partial \langle \hat{Y}_{\alpha} \rangle_{\vec{\tilde{\theta}}}}{\partial \theta_{\alpha'}} \operatorname{cov}(\tilde{\theta}_{\alpha'}, \tilde{\theta}_{\beta'}) \frac{\partial \langle \hat{Y}_{\beta} \rangle_{\vec{\tilde{\theta}}}}{\partial \theta_{\beta'}} = [\mathbf{J}\, \mathbf{\Sigma}\, \mathbf{J}^T]_{\alpha\beta}, \tag{2.62}$$

where the Jacobian matrix $\mathbf{J}$

$$J_{\alpha\alpha'} = \frac{\partial \langle \hat{Y}_{\alpha} \rangle_{\vec{\tilde{\theta}}}}{\partial \theta_{\alpha'}} = \vec{c}_{\alpha}^{T} \frac{\partial \langle \vec{\hat{X}} \rangle_{\vec{\tilde{\theta}}}}{\partial \theta_{\alpha'}} = \frac{\partial \langle \vec{\hat{X}}^T \rangle_{\vec{\theta}}}{\partial \theta_{\alpha}} \mathbf{\Gamma}^{-1} \frac{\partial \langle \vec{\hat{X}} \rangle_{\vec{\theta}}}{\partial \theta_{\alpha'}} = M_{\alpha\alpha'}(\vec{\theta}, \vec{\hat{X}}), \tag{2.63}$$

equals the moment matrix $\boldsymbol{M}$.

Combining equations (2.60–2.63), we arrive at

$$\frac{1}{\mu} \boldsymbol{M} = \boldsymbol{M}\, \mathbf{\Sigma}\, \boldsymbol{M}^T, \tag{2.64}$$

and finally, considering that the matrix $\boldsymbol{M}$ is symmetric and non-degenerate, we establish that the estimator, derived from optimal linear combinations in eq. (2.57), asymptotically achieves the limit defined in eq. (2.53) :

$$\operatorname{cov} \vec{\tilde{\theta}} = \mathbf{\Sigma} = \frac{1}{\mu} \boldsymbol{M}^{-1}. \tag{2.65}$$

2.2.3 Relation between the FI and the MoM sensitivity

The multiparameter CRB, defined in eq. (2.22), limits the quality of an arbitrary estimator, which is based on the measurement results of observables $\vec{\hat{X}}$. To analyze the performance limit of the moment-based estimator eq. (2.52) with the CRB we consider the statistics of the sample means $\overline{x}_j^{(\mu)}$. Following the central limit theorem, for sufficiently large statistics $\mu \gg 1$ the sample means $\overline{x}_j^{(\mu)}$ follow the multivariate Gaussian distribution $\mathcal{N}(\langle \vec{\hat{X}} \rangle_{\vec{\theta}}, \frac{1}{\mu}\mathbf{\Gamma})$. The probability density of the K-dimensional normal distribution $\mathcal{N}(\vec{m}, \mathbf{C})$ is defined as

$$p(\vec{x}) = \frac{1}{(2\pi)^K \det \mathbf{C}} \exp\left[-\frac{1}{2} (\vec{x} - \vec{m})^T \mathbf{C}^{-1} (\vec{x} - \vec{m}) \right]. \tag{2.66}$$

The FI matrix for the observation with such statistics is given by the following expression [Kay 1998]:

$$\mathcal{F}_{\alpha\beta} = \frac{\partial \vec{m}^T}{\partial \theta_{\alpha}} \mathbf{C}^{-1} \frac{\partial \vec{m}}{\partial \theta_{\alpha}} + \frac{1}{2} \operatorname{Tr}\left[\mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial \theta_{\alpha}} \mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial \theta_{\beta}} \right], \tag{2.67}$$

where the derivatives of the vectors and matrices are defined elementwise.

For the measurement statistics $\mathcal{N}(\langle\vec{\hat{X}}\rangle_{\vec{\theta}}, \frac{1}{\mu}\mathbf{\Gamma})$ it takes the form:

$$\mathcal{F}^{(\mu)}_{\alpha\beta}(\vec{\theta},\vec{\tilde{x}}) = \frac{\partial\langle\vec{\hat{X}}^T\rangle_{\vec{\theta}}}{\partial\theta_\alpha}\left(\frac{1}{\mu}\mathbf{\Gamma}\right)^{-1}\frac{\partial\langle\vec{\hat{X}}\rangle_{\vec{\theta}}}{\partial\theta_\beta} + \frac{1}{2}\,\mathrm{Tr}\left[\left(\frac{1}{\mu}\mathbf{\Gamma}\right)^{-1}\frac{\partial\left(\frac{1}{\mu}\mathbf{\Gamma}\right)}{\partial\theta_\alpha}\left(\frac{1}{\mu}\mathbf{\Gamma}\right)^{-1}\frac{\partial\left(\frac{1}{\mu}\mathbf{\Gamma}\right)}{\partial\theta_\beta}\right] =$$
$$= \mu M_{\alpha\beta}(\vec{\theta},\vec{\hat{X}}) + \frac{1}{2}\,\mathrm{Tr}\left[\mathbf{\Gamma}^{-1}\frac{\partial\mathbf{\Gamma}}{\partial\theta_\alpha}\mathbf{\Gamma}^{-1}\frac{\partial\mathbf{\Gamma}}{\partial\theta_\beta}\right]. \tag{2.68}$$

For $\mu \gg 1$ the contribution of the first term dominates over the second which consequently can be neglected, thus the FI matrix of the sample mean of μ samples equals the moment matrix of these measurements

$$\mathcal{F}^{(\mu)}_{\alpha\beta}(\vec{\theta},\vec{\tilde{x}}) = \mu M_{\alpha\beta}(\vec{\theta},\vec{\hat{X}}). \tag{2.69}$$

It is practical to calculate the FI matrix per one probe, i.e.

$$\boldsymbol{\mathcal{F}}(\vec{\theta},\vec{\tilde{x}}) = \frac{1}{\mu}\boldsymbol{\mathcal{F}}^{(\mu)}(\vec{\theta},\vec{\tilde{x}}) = \boldsymbol{M}(\vec{\theta},\vec{\hat{X}}). \tag{2.70}$$

This means, that inequality (2.53) in its essence is the CRB for any first-moment based estimator eq. (2.52). Moreover, the estimator based on N_p optimal linear combinations of the sample means

$$\vec{\tilde{\theta}} = \vec{g}(\overline{y}^{(\mu)}_1, \overline{y}^{(\mu)}_2, ..., \overline{y}^{(\mu)}_{N_p}), \tag{2.71}$$

saturates this bound, as shown in eq. (2.65). Thus, no estimator (2.52), based on the measured first moment of observables $\vec{\hat{X}}$, can provide higher accuracy than the estimator (2.71) based on optimal linear combinations $\vec{\overline{y}}^{(\mu)}$.

In general case, considering higher statistical moments of the measured data may enhance the quality of estimation, i.e. the following inequality holds between the FI matrix and the moment matrix [Gessner 2020b]:

$$\boldsymbol{\mathcal{F}}(\vec{\theta},\vec{\hat{X}}) \geq \boldsymbol{M}(\vec{\theta},\vec{\hat{X}}). \tag{2.72}$$

The rigorous proof of this inequality is provided in [Stein 2014], where the matrix (2.54) was introduced as a lower bound for the FI. The MoM sensitivity offers several advantageous aspects:

- It serves as a lower bound for the FI, often being significantly easier to compute. Additionally, it provides a lower bound for the quantum Fisher information, as elaborated further in section 2.3.
- The MoM sensitivity establishes a bound for the estimator based on the sample mean. In some practical situations, for example when the detection system is slow and does not resolve individual probes, one does not have access to individual measurement results but only observes the sample means.

- The MoM approach provides a clear algorithm for an optimal estimator that saturates the MoM bound (2.53) and does not rely on any optimization procedures.

It is not always obvious, whether the first moment of the observables $\vec{\hat{X}}$ contains all the useful information about the parameters $\vec{\theta}$ and if the inequality (2.72) is saturated, or higher statistical moments should be included in the consideration to increase the estimation accuracy. One of the approaches to this question lies in the context of quantum metrology and is discussed in section 2.3. Here, to demonstrate the difference between the FI and the MoM sensitivity we consider the problem of loss estimation.

2.2.4 Example: loss estimation.

In this section, we apply the concepts introduced above to the specific problem of loss estimation. This example aims to provide intuition regarding the differences and similarities between the CRB (2.22) and the MoM bound (2.53).

Estimating the loss parameter of a bosonic channel is a crucial practical task, closely linked to the separation estimation problem [Lupo 2016] (and imaging in general [Chiuri 2022]), which constitutes the main focus of this thesis. The principal scheme of this procedure is illustrated in fig. 2.3: probe light in the state $\hat{\rho}_0$ is transmitted through the lossy channel with an unknown transmissivity κ. As a consequence of the losses, the state of the probe is altered to $\hat{\rho}(\kappa)$. The probe is then detected using a photon number resolving detector, where the observable $\hat{N} = \hat{a}^\dagger \hat{a}$ is measured. Subsequently, the parameter κ is estimated from the measurement results.

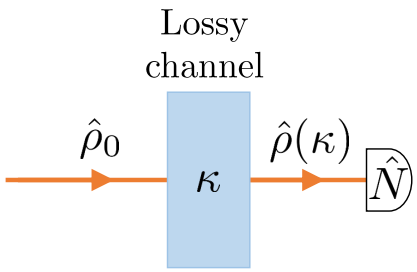

FIGURE 2.3: Scheme for the estimation of the loss parameter of a bosonic channel probed by the state $\hat{\rho}_0$. The channel is characterized by the transmissivity κ. The output state $\hat{\rho}(\kappa)$ is measured with a photon number resolving detector.

The photon number statistics of the probe state is defined by the diagonal elements of the density matrix $\hat{\rho}_0$ in the Fock basis:

$$p_n = \langle n|\hat{\rho}_0|n\rangle. \tag{2.73}$$

After the losses, the photon number statistics p_n transforms according to the Bernoulli transformation (1.71) into q_n:

$$q_n = \langle n|\hat{\rho}(\kappa)|n\rangle = \sum_{m=n}^{\infty} \binom{m}{n} \kappa^n (1-\kappa)^{m-n} p_m, \tag{2.74}$$

that correspond to the probabilities of detecting n photons.

Loss estimation with the MoM

We begin the analysis of this problem with the MoM approach, based on the first moment of the observable $\hat{N}$. To construct the MoM estimator, one only needs to know the mean number of photons in the state $\hat{\rho}(\kappa)$:

$$\langle \hat{N} \rangle_\kappa = \mathrm{Tr}\left[\hat{\rho}(\kappa)\hat{N}\right] = \kappa N_0, \tag{2.75}$$

where

$$N_0 = \mathrm{Tr}\left[\hat{\rho}_0 \hat{N}\right] = \sum_n n p_n \tag{2.76}$$

represents the average number of photons in the probe.

To evaluate the variance of the estimator, one also needs to know the second moment of the observable $\hat{N}$. Using eq. (2.74) one can show that:

$$\Delta^2 N = \mathrm{Tr}\left[\hat{\rho}(\kappa)\hat{N}^2\right] - \langle \hat{N} \rangle_\kappa^2 = \sum_n n^2 q_n - \langle \hat{N} \rangle_\kappa^2 = \kappa^2 \Delta^2 N_0 + \kappa(1-\kappa) N_0, \tag{2.77}$$

with

$$\Delta^2 N_0 = \mathrm{Tr}\left[\hat{\rho}_0 \hat{N}^2\right] - N_0^2 = \sum_n n^2 p_n - N_0^2, \tag{2.78}$$

being the photon number variance in the probe state. Eq. (2.77) can be readily demonstrated using the Heisenberg picture $\hat{a} \to \sqrt{\kappa}\hat{a} + \sqrt{1-\kappa}\hat{v}$ for the description of losses.

The MoM sensitivity, being simple error propagation in this case, takes the form

$$M_{\hat{\rho}_0}(\kappa, \hat{N}) = \left(\frac{\partial \langle \hat{N} \rangle_\kappa}{\partial \kappa}\right)^2 \frac{1}{\Delta^2 N} = \frac{N_0^2}{\kappa^2 \Delta^2 N_0 + \kappa(1-\kappa) N_0}. \tag{2.79}$$

Indeed, one of the significant advantages of MoM in this case is its ability to analyze estimation sensitivity for arbitrary input states, given only the knowledge of the first two moments of the probe's photon number statistics. I.e. it is not even necessary to know the full photon statistics p_n of the probe.

Upon examining eq. (2.79), it becomes evident that for MoM estimation, an optimal probe should exhibit low photon-number noise $\Delta^2 N_0$ and a high mean number of photons N_0. Therefore, the Fock state $|N_0\rangle$ serves as the most suitable choice for estimating the losses [Adesso 2009].

In fig. 2.4 we compare the MoM sensitivity of loss estimation with photon number resolving detection for different probe states:

- Fock state $|N_0\rangle$ with zero photon-number variance $\Delta^2 N_0 = 0$.

- Coherent state $|\alpha\rangle$ with $|\alpha|^2 = N_0$. The corresponding noise is $\Delta^2 N_0 = N_0$.
- Thermal state $\hat{\rho}_{th}(N_0)$ with $\Delta^2 N_0 = N_0^2 + N_0$.
- Superposition of the vacuum and Fock states $|\psi_s\rangle = \frac{|0\rangle + |2N_0\rangle}{\sqrt{2}}$ with $\Delta^2 N_0 = N_0^2$.

In all cases, the mean number of photons is set to $N_0 = 2$.

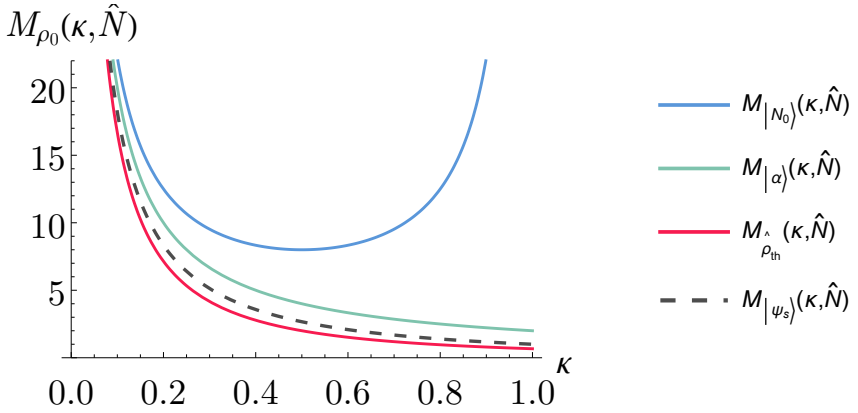

FIGURE 2.4: MoM sensitivity of loss estimation for different probe states: Fock state $|N_0\rangle$ (blue line), coherent state $|\alpha\rangle$ (green line), thermal state $\hat{\rho}_{th}(N_0)$ (red line), and superposition of the vacuum and Fock states $|\psi_s\rangle$ (black dashed line). For all curves, $N_0 = 2$.

One can observe that estimating small losses (i.e., the transmissivity $\kappa \approx 1$) with the MoM estimator poses a challenge with most probes, except for the Fock state. It is interesting to note the behavior of sensitivity with the Fock probe: it decreases for intermediate losses while remaining high for small $\kappa \approx 1$ and high $\kappa \approx 0$ losses. This occurs because the signal $\frac{\partial \langle \hat{N} \rangle_\kappa}{\partial \kappa} = N_0$ does not depend on κ, while the noise $\Delta^2 N$ is minimal for small and large losses, where the output state $\hat{\rho}(\kappa)$ is close to Fock states $|N_0\rangle$ and $|0\rangle$ respectively.

CRB for loss estimation

To determine the CRB for loss estimation from the results of the measurement we employ the standard formula for the FI for discrete statistics:

$$\mathcal{F}_{\hat{\rho}_0}(\kappa, \hat{N}) = \sum_n \frac{1}{q_n} \left(\frac{\partial q_n}{\partial \kappa} \right)^2, \tag{2.80}$$

with detection probabilities q_n defined in eq. (2.74). This expression does not simplify further in the general case; therefore, we examine several specific examples of probe statistics:

- For the probe in the Fock state $|N_0\rangle$ the photon number statistics is $p_n = \delta_{nN_0}$, the FI (2.80) coincides with the MoM sensitivity (2.79)

$$\mathcal{F}_{|N_0\rangle}(\kappa, \hat{N}) = M_{|N_0\rangle}(\kappa, \hat{N}) = \frac{N_0}{\kappa(1-\kappa)}. \tag{2.81}$$

This result suggests that using the Fock probe and measuring only the mean energy, for example with a slow detector that does not resolve individual probes, suffices to construct the MoM estimator and saturate the CRB.

- The photon number statistics of a probe in a coherent state $|\alpha\rangle$ is given by eq. (1.41). The resulting FI for a coherent probe also coincides with the MoM sensitivity:
$$\mathcal{F}_{|\alpha\rangle}(\kappa, \hat{N}) = M_{|\alpha\rangle}(\kappa, \hat{N}) = \frac{N_0}{\kappa}. \tag{2.82}$$

- The photon number statistics of the thermal probe $\hat{\rho}_{th}(N_0)$ is defined by the diagonal part of the density operator represented in the Fock basis (1.50). In this case, the FI is:
$$\mathcal{F}_{\hat{\rho}_{th}}(\kappa, \hat{N}) = M_{\hat{\rho}_{th}}(\kappa, \hat{N}) = \frac{N_0}{\kappa(1 + N_0\kappa)}. \tag{2.83}$$
It is also saturated by the MoM sensitivity. At this point, one may assume, that it is always enough to measure the mean energy of the output state $\hat{\rho}(\kappa)$ to optimally estimate the channel transmissivity κ. However, the next example disproves such a hasty conclusion.

- The superposition of the vacuum and Fock states $|\psi_s\rangle = \frac{|0\rangle + |2N_0\rangle}{\sqrt{2}}$ has photon number statistics $p_n = \frac{1}{2}(\delta_{n\,0} + \delta_{n\,2N_0})$. The FI calculated with eq. (2.80) yields:
$$\mathcal{F}_{|\psi_s\rangle}(\kappa, \hat{N}) = \frac{N_0}{\kappa(1-\kappa)} \left(1 - 2N_0 \frac{\kappa}{(1-\kappa)} \frac{(1-\kappa)^{2N_0}}{1 + (1-\kappa)^{2N_0}}\right). \tag{2.84}$$

The FI for the probe $|\psi_s\rangle$ is smaller than that of the optimal (Fock state) probe: $\mathcal{F}_{|\psi_s\rangle} < \mathcal{F}_{|N_0\rangle}$. However, for small losses $\kappa \approx 1$, the performance of the superposition probe is nearly optimal (see fig. 2.5). This occurs because the state $|2N_0\rangle$ after small losses $\kappa \approx 1$ has almost no vacuum component. Thus, by filtering out detection events with zero photons (i.e. around 50% of the outcomes) one obtains the measurement results that correspond to the probe $|2N_0\rangle$, resulting in average information per probe being

$$\mathcal{F}_{|\psi_s\rangle}(\kappa|_{\kappa\approx 1}, \hat{N}) \approx \frac{1}{2}\mathcal{F}_{|2N_0\rangle}(\kappa, \hat{N}) = \mathcal{F}_{|N_0\rangle}(\kappa, \hat{N}), \tag{2.85}$$

since the FI $\mathcal{F}_{|N_0\rangle}$ in this case is linear with respect to N_0.

The MoM estimator relies solely on the mean transferred energy and does not have access to individual measurement outcomes. Consequently, post-selection of non-zero detected photon numbers is unavailable within this approach. As a result, in the case of the superposition probe $|\psi_s\rangle$ and weak losses $\kappa \approx 1$, the MoM estimator loses a significant portion of information:

$$\mathcal{M}_{|\psi_s\rangle}(\kappa, \hat{N}) = \frac{N_0}{\kappa(1-\kappa)} \left(\frac{1-\kappa}{1 + \kappa(N_0 - 1)}\right) < \mathcal{F}_{|\psi_s\rangle}(\kappa, \hat{N}). \tag{2.86}$$

Its comparison with the FI is illustrated in fig. 2.5.

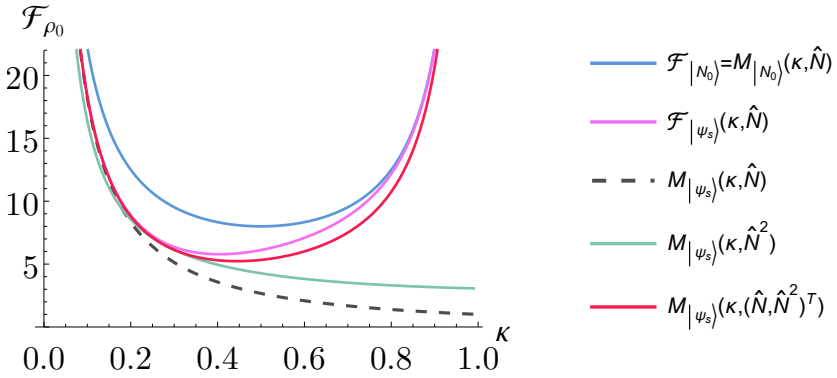

FIGURE 2.5: Loss estimation with the superposition probe state $|\psi_s\rangle$: the FI (purple line), the MoM sensitivity with the first moment $\hat{X} = \hat{N}$ (dashed line), with the second moment $\hat{X} = \hat{N}^2$ (green line), and the first two moments $\vec{\hat{X}} = (\hat{N}, \hat{N}^2)^T$ (red line). The FI for the Fock probe $|N_0\rangle$ is shown in blue (coincides with the MoM sensitivity).

To enhance the performance of the MoM estimation, we also explore estimators based on the second moment $\hat{X} = \hat{N}^2$ (green line in fig. 2.5) and the first two moments $\vec{\hat{X}} = (\hat{N}, \hat{N}^2)^T$ (red line) of the photon number operator $\hat{N}$. To calculate the MoM sensitivity in these cases, we use eq. (2.46). Interestingly, neither the first nor the second moment alone provides good sensitivity for small losses $\kappa \approx 1$. However, the combination of the first two moments yields a sensitivity $M_{|\psi_s\rangle}(\kappa, (\hat{N}, \hat{N}^2)^T)$ close to the FI of the full measurement statistics $\mathcal{F}_{|\psi_s\rangle}(\kappa, \hat{N})$. This suggests that useful information in the case of small losses lies in the correlation of the different moments rather than in the first or second moments individually.

It is noteworthy that the photon number statistics of the superposition state $|\psi_s\rangle$ is identical to the statistics of the classical mixture $\hat{\rho}_s = \frac{1}{2}(|0\rangle\langle 0| + |2N_0\rangle\langle 2N_0|)$. Therefore, these two probes perform identically for the loss estimation problem with photon number measurement. The mixed state $\hat{\rho}_s$ corresponds to the source of the probe pulses, emitting the Fock state $|2N_0\rangle$ with 50% probability. Hence, as indicated by the previous analysis, with such a probabilistic source of the Fock state, it is necessary to detect each probe individually to construct an efficient estimator in the case of small losses. Conversely, for a deterministic source of the Fock (or coherent/thermal) states, the detection system does not have to resolve individual probes, as the estimator based on the mean transferred energy saturates the CRB.

2.2.5 FI and MoM in the Poissonian limit

In this section, we explore another significant example: Poisson statistics of measurement, and we compare the FI with the MoM sensitivity in this scenario. This case holds considerable importance in quantum optical metrology. Given any initial state $\hat{\rho}_0$, after undergoing sufficiently strong losses ($\kappa \ll 1$ and $\kappa \,\mathrm{Tr}[\hat{N}\hat{\rho}_0] \ll 1$), it can be approximated within the zero and single photon sub-space. This approximation neglects contributions from terms with higher photon numbers. The photon number

statistics in such a case can be represented as follows:

$$p_1 = 1 - p_0, \;\; p_1 \ll p_0 < 1, \;\; \text{and } p_{n>1} \approx 0. \tag{2.87}$$

According to the Poisson limit theorem, these statistics can be approximated by the Poisson distribution [Papoulis 2002]:

$$p_n(\lambda) = e^{-\lambda}\frac{\lambda^n}{n!}, \tag{2.88}$$

with $\lambda = p_1 \ll 1$. If photon number measurements are conducted in multiple orthogonal modes, the outcomes of measurements in individual modes can always be treated as independent in this scenario.

One more important example of Poissonian statistics of the observables is metrology employing a single-mode coherent state probe $\hat{\rho}_0 = |\alpha\rangle\langle\alpha|$. Under linear operations, this state remains coherent and maintains Poisson photon number statistics. Consequently, if the system is linear, the photon number statistics of the output state $\hat{\rho}(\theta)$ are also Poissonian, even in cases of weak losses ($\kappa \lesssim 1$). This outcome holds across any measurement basis. The detection events in orthogonal modes are always uncorrelated because coherent states remain a product state under linear transformations, as demonstrated in eq. (1.63).

We begin our exploration with single-parameter estimation derived from the measurement outcomes of the single observable $\hat{N}$. Assuming the measurement statistics follow the Poisson distribution (2.88), with the parameter of interest θ encoded into $\lambda = \langle\hat{N}\rangle_\theta$, the FI of this measurement, computed using a discrete version of eq. (2.7), is expressed as:

$$\mathcal{F}(\theta, \hat{N}) = \sum_n \frac{1}{p_n(\lambda)}\left(\frac{\partial p_n(\lambda)}{\partial\theta}\right)^2 = \frac{1}{\langle\hat{N}\rangle_\theta}\left(\frac{\partial\langle\hat{N}\rangle_\theta}{\partial\theta}\right)^2. \tag{2.89}$$

For a collection of independent observables $\hat{N}_m$ with Poisson statistics, the FI accumulates as follows:

$$\mathcal{F}(\theta, \vec{\hat{N}}) = \sum_m \frac{1}{\langle\hat{N}_m\rangle_\theta}\left(\frac{\partial\langle\hat{N}_m\rangle_\theta}{\partial\theta}\right)^2. \tag{2.90}$$

The covariance matrix of these observables is diagonal:

$$\Gamma_{mn} = \delta_{mn}\,\Delta^2 N_m = \delta_{mn}\langle\hat{N}_m\rangle_\theta. \tag{2.91}$$

Consequently, the MoM sensitivity, calculated with eq. (2.46), gives

$$M(\theta, \vec{\hat{N}}) = \sum_m \frac{1}{\langle\hat{N}_m\rangle_\theta}\left(\frac{\partial\langle\hat{N}_m\rangle_\theta}{\partial\theta}\right)^2 = \mathcal{F}(\theta, \vec{\hat{N}}), \tag{2.92}$$

yielding the same result as the FI. Thus, in cases of Poisson statistics of observables, an optimal estimator can be constructed based on the first moment of observables $\vec{\hat{N}}$.

Now to develop better intuition on why the sensitivity based on the first moment saturates the FI in this case, let us revisit a single observable problem to consider

it in more detail. If μ probes are used, the measurement results are given by the vector $(n_1, ..., n_\mu)^T$. Interestingly, even though the sample mean $\overline{n}^{(\mu)}$ contains all useful information about the state with Poissonian statistics, one can still devise an estimator solely based on the second moment. The sensitivity of this estimation is given by:

$$M(\theta, \hat{N}^2) = \frac{1}{\langle \hat{N}^4 \rangle_\theta - \langle \hat{N}^2 \rangle_\theta^2} \left(\frac{\partial \langle \hat{N}^2 \rangle_\theta}{\partial \theta} \right)^2 = \tag{2.93}$$

$$= \left(\frac{1}{\langle \hat{N} \rangle_\theta} - \frac{2}{1 + 6\langle \hat{N} \rangle_\theta + 4\langle \hat{N} \rangle_\theta^2} \right) \left(\frac{\partial \langle \hat{N} \rangle_\theta}{\partial \theta} \right)^2 < M(\theta, \vec{\hat{N}}) \tag{2.94}$$

which is smaller than that for the first moment estimator, yet the difference is at most 20%.

Now, it is interesting to have a look at the estimator based on the first two moments $\vec{\hat{X}} = (\hat{N}, \hat{N}^2)^T$. The covariance matrix of these observables is given by

$$\mathbf{\Gamma} = \langle \hat{N} \rangle_\theta \begin{pmatrix} 1 & 1 + 2\langle \hat{N} \rangle_\theta \\ 1 + 2\langle \hat{N} \rangle_\theta & 1 + 6\langle \hat{N} \rangle_\theta + 4\langle \hat{N} \rangle_\theta^2 \end{pmatrix}. \tag{2.95}$$

This covariance matrix is non-degenerate, indicating that the moments $\langle \hat{N} \rangle_\theta$ and $\langle \hat{N}^2 \rangle_\theta$ are not fully linearly correlated (although this does not guarantee their statistical independence). However, the sensitivity of the first two moments,

$$M(\theta, (\hat{N}, \hat{N}^2)^T) = M(\theta, \hat{N}), \tag{2.96}$$

does not surpass the sensitivity of the first moment estimation[2]. Moreover, in this scenario, the optimal coefficients (2.50) for linear combination (2.36) are

$$\vec{c} = (1, 0)^T, \tag{2.97}$$

meaning the optimal linear combination of the measurement results solely includes the sample mean and disregards higher moments.

It is also intriguing to examine the MLE (2.25) in the context of Poisson statistics. It can be derived as:

$$\tilde{\theta}_{ML} = \arg\max_\theta \sum_{i=1}^{\mu} \ln p_{n_i}(\lambda(\theta)). \tag{2.98}$$

Utilizing eq. (2.88), we can readily demonstrate that:

$$\frac{\partial}{\partial \theta} \sum_{i=1}^{\mu} \ln p_{n_i}(\lambda(\theta)) = \frac{\partial \lambda(\theta)}{\partial \theta} \sum_{i=1}^{\mu} \left(\frac{n_i}{\lambda(\theta)} - 1 \right), \tag{2.99}$$

that equals 0 when

$$\lambda(\tilde{\theta}_{ML}) = \frac{1}{\mu} \sum_{i=1}^{\mu} n_i = \overline{n}^{(\mu)}. \tag{2.100}$$

[2]This is expected since $M(\theta, \hat{N})$ already saturates the CRB.

Hence, the likelihood maximization approach in the case of Poisson statistics yields an estimator that solely depends on the sample mean of the observable. This fact explains the optimality of the MoM estimator for Poisson statistics.

All conclusions drawn in this section remain valid for the multiparameter estimation case: estimators based on the sample means of Poissonian observables saturate the CRB, and the moment matrix (2.54) aligns with the FI matrix (2.21). This equivalence can be demonstrated by carefully examining the matrix generalizations of eqs. (2.89) and (2.92) in the multiparameter scenario, element by element. Consequently, the FI and MoM approaches are shown to be equivalent for Poissonian measurement statistics.

2.3 Quantum parameter estimation theory

In the previous section, we analyzed the efficiency of different estimators, that are based on the given measurement results, which is a primary task of classical parameter estimation theory. In contrast, quantum parameter estimation, as a component of quantum metrology, establishes the sensitivity bound for arbitrary measurements (and subsequent parameter estimation) conducted on the parameter-dependent state $\hat{\rho}(\vec{\theta})$.

2.3.1 Quantum Fisher information and quantum Cramér–Rao bound.

The fundamental quantity in quantum parameter estimation is the *quantum Fisher information* (QFI) matrix $\boldsymbol{\mathcal{F}}^{\mathrm{Q}}$, defined by its elements [Helstrom 1969; Liu 2019]:

$$\mathcal{F}^{\mathrm{Q}}_{\alpha\beta}(\vec{\theta}) = \frac{1}{2}\left\langle \{\hat{\mathcal{L}}_\alpha, \hat{\mathcal{L}}_\beta\}\right\rangle_{\vec{\theta}}, \tag{2.101}$$

where $\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}$ is the anti-commutator and $\hat{\mathcal{L}}_\alpha$ is the *symmetric logarithmic derivative* (SLD), implicitly defined as

$$\frac{\partial\hat{\rho}(\vec{\theta})}{\partial\theta_\alpha} = \frac{1}{2}\{\hat{\rho}(\vec{\theta}), \hat{\mathcal{L}}_\alpha\}. \tag{2.102}$$

Similarly to its classical counterpart, the QFI matrix is real, symmetric, positive semidefinite, and additive for independent probes that are measured independently. It establishes the ultimate, fundamental bound for any arbitrary estimator based on arbitrary measurement:

$$\operatorname{cov}\vec{\hat{\theta}} \geq \frac{1}{\mu}\boldsymbol{\mathcal{F}}^{\mathrm{Q}}(\vec{\theta})^{-1}, \tag{2.103}$$

where μ is the number of repetitions of the parameter estimation procedure. This inequality is referred to as the *quantum Cramér–Rao bound* (QCRB).

Since the QFI is optimized over all possible measurements, the following chain of inequalities holds true:

$$\boldsymbol{\mathcal{F}}^{\mathrm{Q}}(\vec{\theta}) \geq \boldsymbol{\mathcal{F}}(\vec{\theta}, \vec{\hat{X}}) \geq \boldsymbol{M}(\vec{\theta}, \vec{\hat{X}}). \tag{2.104}$$

Thus, on the one hand, the QFI $\boldsymbol{\mathcal{F}}^{\mathrm{Q}}$ serves as an upper bound for both the FI $\boldsymbol{\mathcal{F}}(\vec{\theta}, \vec{\hat{X}})$ and the moment matrix $\boldsymbol{M}(\vec{\theta}, \vec{\hat{X}})$. If a given set of measurements $\vec{\hat{X}}$ saturates the first inequality, i.e. $\boldsymbol{\mathcal{F}}(\vec{\theta}, \vec{\hat{X}}) = \boldsymbol{\mathcal{F}}^{\mathrm{Q}}(\vec{\theta})$, these measurements are called *optimal*. In this case, no further improvement in the estimation quality can be achieved by optimizing the measurement scheme. Only optimizing the probe state $\hat{\rho}_0$ or the parameter-encoding procedure can further enhance the information about the parameters.

In some cases, calculating the FI of specific practical measurements $\vec{\hat{X}}$ can be challenging. For instance, it is known that the statistics $p(\vec{x}|\vec{\theta})$ of multimode photon number-resolving measurements of multi-photon states (which are the main measurement types for this thesis) is hard or even impossible to calculate with classical computer [Aaronson 2014]. However, the ultimate precision bound, given by the QFI, can sometimes be calculated in this case using the specific properties of the state, such as Gaussianity [Sorelli 2022]. In this scenario, calculating the moment matrix and demonstrating that it even saturates the QFI ($\boldsymbol{M}(\vec{\theta}, \vec{\hat{X}}) = \boldsymbol{\mathcal{F}}^{\mathrm{Q}}(\vec{\theta})$) confirms the optimality of the measurement $\vec{\hat{X}}$. Additionally, in this case, the MoM suggests a practical estimation strategy that saturates the ultimate bound. Using this approach in Chapter 3, we demonstrate the optimality of the spatial mode demultiplexing measurement for the source separation estimation problem, considering a wide class of source statistics.

Finally, the moment matrix $\boldsymbol{M}(\vec{\theta}, \vec{\hat{X}})$ can be considered as a lower bound for both the FI and QFI matrices. Depending only on the first two moments of the statistics of the observables $\vec{\hat{X}}$, the moment matrix is often much easier to compute for a general class of probe states $\hat{\rho}_0$, while the computation of the QFI and FI matrices strongly depends on the structure of the state $\hat{\rho}(\vec{\theta})$. Thus, when the QFI is challenging to compute, one can draw useful conclusions by computing its lower bound in the form of the moment matrix.

2.3.2 Geometrical sense of the QFI and some particular case examples

The QFI matrix illustrates the rate at which the state $\hat{\rho}(\vec{\theta})$ changes with variations in the parameters $\vec{\theta}$. This change is captured by the Bures distance , which quantifies the disparity between two quantum states $\hat{\rho}$ and $\hat{\sigma}$. It is defined as [Liu 2019; Sidhu 2020]:

$$D_B(\hat{\rho}, \hat{\sigma}) = \sqrt{2(1 - F(\hat{\rho}, \hat{\sigma}))}, \tag{2.105}$$

where the fidelity $F(\hat{\rho}, \hat{\sigma})$ between the two states is given by:

$$F(\hat{\rho}, \hat{\sigma}) = \mathrm{Tr}\sqrt{\sqrt{\hat{\rho}}\hat{\sigma}\sqrt{\hat{\rho}}}. \tag{2.106}$$

It is worth noting that there are various conventions for defining fidelity and Bures distance. Here, we adhere to the definitions outlined in [Liu 2019].

If we consider two sets of parameters $\vec{\theta}$ and $\vec{\theta} + d\vec{\theta}$ that are infinitesimally close,

and assuming that $\hat{\rho}(\vec{\theta})$ and $\hat{\rho}(\vec{\theta}+d\vec{\theta})$ have the same rank, the squared Bures distance can be expressed as [Liu 2019]:

$$D_B^2\left(\hat{\rho}(\vec{\theta}),\hat{\rho}(\vec{\theta}+d\vec{\theta})\right)=\frac{1}{4}\,d\vec{\theta}^T\boldsymbol{\mathcal{F}}^Q d\vec{\theta}. \tag{2.107}$$

This ratio establishes a connection between the geometry of the manifold of quantum states $\hat{\rho}(\vec{\theta})$ and the QFI matrix, offering deeper insights into the fundamental properties of the QFI matrix. This property also provides an alternative method for calculating the QFI matrix:

$$\mathcal{F}^Q_{\alpha\beta}(\vec{\theta})=-4\frac{\partial^2 F\left(\hat{\rho}(\vec{\theta}),\hat{\rho}(\vec{\Theta})\right)}{\partial\Theta_\alpha\partial\Theta_\beta}\bigg|_{\vec{\Theta}=\vec{\theta}}, \tag{2.108}$$

which reveals that the QFI matrix measures how rapidly the state $\hat{\rho}(\theta)$ changes with alterations in the parameters. The more sensitive the state $\hat{\rho}(\vec{\theta})$ to the change in the parameter θ_α, the higher the corresponding QFI element $\mathcal{F}^Q_{\alpha\alpha}$ is.

This form of QFI computation can also prove valuable for practical purposes, as demonstrated below.

QFI for Gaussian states

The geometric approach becomes particularly valuable in the case of Gaussian states. The fidelity between two Gaussian states, characterized by the mean-field vectors $\vec{Q}^{(1,2)}$ and quadrature covariance matrices $\mathbf{\Gamma}_Q^{(1,2)}$ respectively, can be expressed as [Scutaru 1998]:

$$F^2(\hat{\rho}_1,\hat{\rho}_2)=\frac{2\exp\left[-\frac{1}{2}(\vec{Q}^{(1)}-\vec{Q}^{(2)})^T(\mathbf{\Gamma}_Q^{(1)}+\mathbf{\Gamma}_Q^{(2)})^{-1}(\vec{Q}^{(1)}-\vec{Q}^{(2)})\right]}{\sqrt{\det(\mathbf{\Gamma}_Q^{(1)}+\mathbf{\Gamma}_Q^{(2)})+T}-\sqrt{T}}, \tag{2.109}$$

with $T=\left(1-\det\mathbf{\Gamma}_Q^{(1)}\right)\left(1-\det\mathbf{\Gamma}_Q^{(2)}\right)$. One can use this expression for fidelity, and QFI formula (2.108) to demonstrate that in the case of a single-mode Gaussian state with a mean-field vector $\vec{Q}(\vec{\theta})$ and a 2×2 quadrature covariance matrix $\mathbf{\Gamma}_Q(\vec{\theta})$, the corresponding QFI matrix can be found as [Pinel 2013]:

$$\mathcal{F}^Q_{\alpha\beta}(\vec{\theta})=\frac{1}{2}\frac{1}{1+P^2}\,\mathrm{Tr}\left[\mathbf{\Gamma}_Q^{-1}\frac{\partial\mathbf{\Gamma}_Q}{\partial\theta_\alpha}\mathbf{\Gamma}_Q^{-1}\frac{\partial\mathbf{\Gamma}_Q}{\partial\theta_\beta}\right]+\frac{2}{1-P^4}\frac{\partial P}{\partial\theta_\alpha}\frac{\partial P}{\partial\theta_\beta}+\frac{\partial Q^T}{\partial\theta_\alpha}\mathbf{\Gamma}_Q^{-1}\frac{\partial Q}{\partial\theta_\beta}, \tag{2.110}$$

where the purity is $P=1/\det\mathbf{\Gamma}_Q$. This expression provides an explicit way of calculating the QFI for Gaussian states, which is further analyzed in chapter 4.

QFI for pure states

Another interesting particular case is the case of pure states. If the probe state is pure $\hat{\rho}_0=|\psi_0\rangle\langle\psi_0|$, and it remains pure after parameter encoding $\hat{\rho}(\vec{\theta})=|\psi(\vec{\theta})\rangle\langle\psi(\vec{\theta})|$, then the encoding procedure can be described by a unitary transformation $|\psi(\vec{\theta})\rangle=$

$\hat{U}(\vec{\theta})|\psi_0\rangle$. In a small region of the parameter space, the encoding operator can always be represented in the linear approximation as $\hat{U}(\vec{\theta}) = \exp\left[i\vec{\hat{H}}^T\vec{\theta}\right]$.

The fidelity (2.106) between two pure states $|\psi\rangle$ and $|\phi\rangle$ is simplified to

$$F(|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|) = |\langle\psi|\phi\rangle|. \tag{2.111}$$

Using this property, and linear unitary encoding

$$|\psi(\vec{\theta})\rangle = \exp\left[i\vec{\hat{H}}^T\vec{\theta}\right]|\psi_0\rangle, \tag{2.112}$$

one can show that the QFI matrix in this case is defined as the covariance matrix of the generators $\vec{\hat{H}}$:

$$\mathcal{F}^{Q}_{\alpha\beta}(\vec{\theta}) = 4\left(\frac{1}{2}\langle\psi_0|\{\hat{H}_\alpha, \hat{H}_\beta\}|\psi_0\rangle - \langle\psi_0|\hat{H}_\alpha|\psi_0\rangle\langle\psi_0|\hat{H}_\beta|\psi_0\rangle\right). \tag{2.113}$$

In the single-parameter estimation case, the QFI is simplified to

$$\mathcal{F}^{Q}(\vec{\theta}) = 4\left(\langle\psi_0|\hat{H}^2|\psi_0\rangle - \langle\psi_0|\hat{H}|\psi_0\rangle^2\right) = 4\Delta^2 H. \tag{2.114}$$

The same result can be obtained with the definition of the QFI (2.101), since for the pure states $\hat{\rho}^2(\vec{\theta}) = |\psi(\vec{\theta})\rangle\langle\psi(\vec{\theta})|\psi(\vec{\theta})\rangle\langle\psi(\vec{\theta})| = \hat{\rho}(\vec{\theta})$ and subsequently

$$\frac{\partial\hat{\rho}(\vec{\theta})}{\partial\theta_\alpha} = \frac{\partial\hat{\rho}^2(\vec{\theta})}{\partial\theta_\alpha} = \frac{\partial\hat{\rho}(\vec{\theta})}{\partial\theta_\alpha}\hat{\rho}(\vec{\theta}) + \hat{\rho}(\vec{\theta})\frac{\partial\hat{\rho}(\vec{\theta})}{\partial\theta_\alpha} = \frac{1}{2}\left(\hat{\mathcal{L}}_\alpha\hat{\rho}(\vec{\theta}) + \hat{\rho}(\vec{\theta})\hat{\mathcal{L}}_\alpha\right). \tag{2.115}$$

Thus the SLD for pure state equals to

$$\hat{\mathcal{L}}_\alpha = 2\frac{\partial\hat{\rho}(\vec{\theta})}{\partial\theta_\alpha}, \tag{2.116}$$

which, combined with encoding (2.112) and definition (2.101) gives eq. (2.113).

2.3.3 QFI calculation through state diagonalization

Another commonly used technique for calculating the QFI involves diagonalizing the state $\hat{\rho}(\vec{\theta})$ in its eigenbasis:

$$\hat{\rho}(\vec{\theta}) = \sum_{i=1}^{\operatorname{rank}\hat{\rho}(\vec{\theta})} \lambda_i|u_i\rangle\langle u_i|. \tag{2.117}$$

Assuming that $\hat{\rho}(\vec{\theta})$ is full rank in the Hilbert space containing SLDs, the SLDs can be represented in this eigenbasis as [Liu 2019]:

$$\hat{\mathcal{L}}_\alpha = \sum_{i,j=1}^{\operatorname{rank}\hat{\rho}(\vec{\theta})} \frac{2}{\lambda_i + \lambda_j}\left\langle u_i\middle|\frac{\partial\hat{\rho}(\vec{\theta})}{\partial\theta_\alpha}\middle|u_j\right\rangle |u_j\rangle\langle u_i|. \tag{2.118}$$

Substituting this result into eq. (2.101), one obtains

$$\mathcal{F}^{Q}_{\alpha\beta}(\vec{\theta}) = \sum_{i,j=1}^{\operatorname{rank}\hat{\rho}(\vec{\theta})} \frac{2}{\lambda_i + \lambda_j} \operatorname{Re}\left[\left\langle u_i \middle| \frac{\partial\hat{\rho}(\vec{\theta})}{\partial\theta_\alpha} \middle| u_j \right\rangle \left\langle u_j \middle| \frac{\partial\hat{\rho}(\vec{\theta})}{\partial\theta_\beta} \middle| u_i \right\rangle\right]. \tag{2.119}$$

This expression can be adapted for cases where $\hat{\rho}(\vec{\theta})$ is not full rank. In such instances, one must exclude the diverging summation terms, i.e., those for which $\lambda_i + \lambda_j = 0$.

The diagonalization method is commonly employed in the literature. However, its implementation in practice is often challenging, especially for CV states since they are described by the infinite-size density matrix[3].

2.3.4 Saturability of the QCRB

Having the QCRB (2.103) and the expression (2.101) for the QFI, one may wonder how to find a quantum optimal measurement that saturates the inequality (2.104). To answer this question, we initially examine the single-parameter estimation case. Firstly, we determine the eigenbasis of the SLD operator

$$\hat{\mathcal{L}}|l_k\rangle = l_k|l_k\rangle, \tag{2.120}$$

which constitutes a complete basis since the SLD operator is Hermitian. Subsequently, we construct an observable $\hat{X}$ with identical eigenvectors:

$$\hat{X}|l_k\rangle = a_k|l_k\rangle. \tag{2.121}$$

Then, the measurement statistics can be derived as follows:

$$p(a_k|\theta) = \langle l_k|\hat{\rho}(\theta)|l_k\rangle, \tag{2.122}$$

and the corresponding FI is given by [Liu 2019]:

$$\mathcal{F}(\theta, \hat{X}) = \sum_k \frac{1}{p(a_k|\theta)} \left(\frac{\partial p(a_k|\theta)}{\partial\theta}\right)^2 = \tag{2.123}$$

$$= \sum_k \frac{1}{\langle l_k|\hat{\rho}(\theta)|l_k\rangle} \left\langle l_k \middle| \frac{\hat{\mathcal{L}}\hat{\rho}(\theta) + \hat{\rho}(\theta)\hat{\mathcal{L}}}{2} \middle| l_k \right\rangle^2 = \sum_k l_k^2 \langle l_k|\hat{\rho}(\theta)|l_k\rangle. \tag{2.124}$$

It is straightforward to show, that the FI of observable $\hat{X}$ equals to the QFI:

$$\mathcal{F}^{Q}(\theta) = \operatorname{Tr}\left[\hat{\rho}(\theta)\hat{\mathcal{L}}^2\right] = \sum_k \langle l_k|\hat{\rho}(\theta)\hat{\mathcal{L}}^2|l_k\rangle = \sum_k l_k^2 \langle l_k|\hat{\rho}(\theta)|l_k\rangle. \tag{2.125}$$

Therefore, any observable $\hat{X}$, that is diagonal in the SLD eigenbasis, is considered quantum optimal. This conclusion can be extended to multiple observables, bearing

[3] The density matrix can be truncated, however its dimension for CV states is typically high

in mind the fundamental principle of quantum physics that non-commuting observables cannot be simultaneously measured. If SLDs commutes

$$[\hat{\mathcal{L}}_\alpha, \hat{\mathcal{L}}_\beta] = 0, \tag{2.126}$$

they can all be simultaneously diagonalized. Consequently, observables constructed from their eigenbasis are compatible with each other and saturate QCRB. There exists a relaxed condition for the achievability of the QCRB [Ragy 2016; Demkowicz-Dobrzański 2020]:

$$\mathrm{Tr}\left[\hat{\rho}(\theta)[\hat{\mathcal{L}}_\alpha, \hat{\mathcal{L}}_\beta]\right] = 0. \tag{2.127}$$

However, if the condition in Equation (2.126) is not met, determining optimal measurements generally poses a challenge.

In cases where the QCRB is not saturable, the lower bound for the total mean square error of the estimator can be determined using the Holevo bound [Holevo 1973; Demkowicz-Dobrzański 2020], which is at most twice the trace of the inverse QFI matrix. The achievability of the Holevo bound in the limit of large sample sizes has been demonstrated in [Gill 2013; Demkowicz-Dobrzański 2020].

In summary, quantum metrology establishes a fundamental bound (2.103) for the precision of parameter estimation. When condition (2.126) is satisfied, it also provides observables (2.121) that achieve this bound. However, implementing these specific measurements, suggested by the eigenbasis of the SLDs, is often challenging in practice. Moreover, these optimal measurements can often be parameter-dependent. Consequently, without prior knowledge of the parameter values, an adaptive measurement scheme becomes necessary: initially estimating the parameters with a preliminary non-optimal measurement and then adjusting the measurement scheme to extract more information from subsequent measurements, repeating this process iteratively if needed. Implementing such an approach can be particularly challenging, especially when aiming for a short measurement time. Moreover, the initial iterations in this approach may not yield all possible information.

In reality, an approach employing a fixed measurement scheme to estimate the parameters in a wide range of values may prove more practical. Additionally, practical constraints often limit the class of measurements feasible for a given level of technology. Hence, it is valuable to assess the information content of practically available measurements and compare it with the ultimate QCRB. This comparison constitutes one of the objectives of this thesis.

2.4 Quantum state characterization

The primary challenge in quantum metrology arises from the inherently probabilistic nature of measurement in quantum physics. Even when the state of the system $\hat{\rho}$ is defined up to a single unknown parameter θ, multiple copies of this state $\hat{\rho}(\theta)$ are typically required to conduct measurements, gather outcome statistics, and estimate θ. This challenge intensifies when no prior information about the state $\hat{\rho}$ is accessible, necessitating the complete reconstruction of the state $\hat{\rho}$ from the measurement outcomes. This process is known as *quantum state tomography* [Lvovsky 2009]

or shorter *quantum tomography* [DAriano 2003; Artiles 2005]. The name originates from the fact that the full state $\hat{\rho}$ is reconstructed from its various "projections" — the probability distributions of different measurement results. Consequently, the problem mathematically resembles classical tomography methods, such as CT scans in medical imaging or other 3D imaging techniques, that reconstruct 3D objects from their 2D projections.

In this section, we provide a brief overview of known techniques for realizing quantum state tomography, with a specific emphasis on the states of light. This task holds particular practical importance for the characterization of non-classical light sources. We consider a single-mode scenario where the spatial and temporal mode of the light is known, and the goal is to determine the quantum state $\hat{\rho}$ of this mode. It is important to note, however, that one of the major challenges in quantum tomography is the scaling of the complexity with an increasing number of modes.

There are various measurement strategies employed in quantum tomography. Methods utilizing photon-number resolving detection [Olivares 2019] are particularly suited for the tomography of states localized near vacuum in the Fock space. Other techniques, such as compressed sensing [Smith 2013; Kyrillidis 2018], have proven to be efficient for tomography of quantum states that exhibit sparsity in a certain basis or possess a low-rank structure. In this brief overview, our primary focus is on tomography methods relying on measurements of continuous variables, as this methodology is better suited for characterizing general CV states with high photon-number components [Lvovsky 2009].

One of the known approaches to CV state tomography is based on the double homodyne detection [Chabaud 2021]. A double homodyne (or heterodyne) detection method can be employed to sample the Husimi function $Q(\alpha)$, as briefly mentioned in 1.3.4 [DAriano 2003]. Subsequently, the Husimi function can be transformed into either the Wigner function $W(\alpha)$ or the density operator $\hat{\rho}$, enabling a comprehensive description of the state under study. The main advantage of this approach is the ability to reconstruct the single-mode state using a fixed measurement setup. However, this technique necessitates a sophisticated detection scheme and is highly sensitive to the noise in the measured data.

Another well-established approach to optical state tomography is the *homodyne tomography*. This method reconstructs the state $\hat{\rho}$ from the statistics of homodyne detection. For a comprehensive characterization of the state, it is crucial to perform homodyne detection with different phases of the local oscillator, thereby acquiring statistics of numerous quadratures. The probability distribution for an arbitrary quadrature $\hat{q}^\theta$ (1.84) is obtained by integral projection of the Wigner function along the corresponding direction [Artiles 2005]:

$$P(q^\theta, \theta) = \int dp^\theta \, W_{\hat{\rho}}(q^\theta \cos\theta - p^\theta \sin\theta, q^\theta \sin\theta + p^\theta \cos\theta), \tag{2.128}$$

this integral is known as *Radon transform*.

These marginal distributions $P(q^\theta, \theta)$ can be obtained experimentally. To achieve this, one needs to collect quadrature samples for various phases θ of the local oscillator, i.e. accumulate a set of pairs (q_m, θ_m). By organizing this measured data into

bins, one can estimate the probabilities $P(q^\theta, \theta)$ from the frequencies of different outcomes. In the following sections, we analyze several methods for characterizing the quantum state $\hat{\rho}$ based on this data.

2.4.1 Inverse Radon transformation

The most straightforward way to obtain the Wigner function from the homodyne detection results is to invert the Radon transform (2.128) using the back-projection algorithm [Herman 2009]:

$$W(q,p) = \frac{1}{2\pi^2}\int_0^\pi d\theta \int_{-\infty}^{\infty} dq^\theta P(q^\theta,\theta)K(q^\theta\cos\theta + p^\theta\sin\theta - q^\theta), \tag{2.129}$$

with the integration kernel

$$K(x) = \frac{1}{2}\int_{-\infty}^{\infty} |\xi| \exp[i\xi x] d\xi. \tag{2.130}$$

By applying the inverse Radon transform (2.129) to the measured probabilities $P(q^\theta, \theta)$, the Wigner function of the state can be recovered. If necessary, it can be further transformed to determine the density operator $\hat{\rho}$, for instance, using eq. (1.102).

The intermediate step of binning the data and calculating individual marginal distributions associated with each phase can be bypassed if the phases θ_m are uniformly spread over the 2π interval. It can be achieved by directly using the acquired pairs (q_m, θ_m) in the summation [Lvovsky 2002]:

$$W(q,p) = \frac{1}{2\pi^2 N}\sum_{m=1}^{N} K(q\cos\theta_m + p\sin\theta_m - q_m). \tag{2.131}$$

It is worth noting that the kernel $K(x)$ is infinite at $x = 0$. Therefore, in numerical implementations of the inverse Radon transformation, it undergoes low-pass filtering: the infinite integration limits in eq. (2.130) are replaced by $\pm k_c$, with constant k_c chosen to mitigate numerical artifacts associated with the reconstruction while preserving the main features of the Wigner function. This approach is commonly known as the *filtered back-projection algorithm* [Lvovsky 2009]. It operates under the assumption of a certain smoothness of the Wigner function, effectively filtering out high-frequency components in its spectrum.

Thus, performing a series of homodyne measurements, one can use the transformation (2.131) to reconstruct the Wigner function of the state. However, this approach is highly sensitive to the measurement noise, and it can be challenging to estimate the resulting error of the reconstruction [Butucea 2007]. The inverse Radon transform performs well only when these uncertainties are negligible, i.e., in the limit of a very large number of data points and very precise measurements [Lvovsky 2009]. Otherwise, errors in the right side of the transformation (2.129) can result in inaccurate or unphysical features in the reconstructed state. For instance, negative values may appear on the diagonal of the reconstructed density matrix, and its trace is not guaranteed to equal one. For this reason, the inversion method is

rarely used in modern homodyne tomography, and the statistical inference methods are used instead [Lvovsky 2009].

2.4.2 Statistical reconstruction

Statistical methods in quantum tomography are closely related to parameter estimation theory. Within this framework, the components of the density matrix ρ_{nm} (typically in the Fock basis) are regarded as unknown parameters $\vec{\theta}$. To reduce the number of parameters, the state $\hat{\rho}$ is often assumed to be localized in the chosen basis, and therefore it is fully described by a truncated density matrix of finite size. Subsequently, MLE (2.25) can be employed to determine the density matrix that maximizes the probability of the observed measurement results. Unlike the direct inversion approach, the use of statistical reconstruction enables the specification of the parameter space ρ_{nm} in such a way that it always corresponds to a valid density matrix (i.e., unit-trace and positively defined). Note that constraining the parameters in this manner can lead to biased estimators and render the CRB (2.22) inapplicable for evaluating the accuracy of the state reconstruction.

Maximizing the likelihood function straightforwardly is typically challenging for a large size of the density matrix. The optimization needs to be performed over a considerable number of parameters, while also taking into account the constraints on these parameters. The iterative *Expectation-Maximization* (EM) algorithm is often employed to address this issue [Lvovsky 2009].

To introduce the EM algorithm for homodyne tomography, first, let us consider the measurement of the observables with discrete spectrum and eigenvectors $|\psi_j\rangle$. Assume there are N repetitions of the measurement, with each corresponding outcome occurring f_j times. Then, an iterative procedure can be constructed using an operator [Hradil 1997]

$$\hat{R}(\hat{\rho}) = \frac{1}{N}\sum_j \frac{f_j}{\langle\psi_j|\hat{\rho}|\psi_j\rangle}|\psi_j\rangle\langle\psi_j|. \tag{2.132}$$

Starting from an initial guess $\hat{\rho}^{(0)}$ for the density matrix, we can construct a sequence of operators [DAriano 2003]:

$$\hat{\rho}^{(k+1)} = \mathcal{N}\left[\hat{R}(\hat{\rho}^{(k)})\hat{\rho}^{(k)}\hat{R}(\hat{\rho}^{(k)})\right], \tag{2.133}$$

where the operation $\mathcal{N}$ denotes the normalization to ensure the unit trace of the density operator. This sequence preserves the positivity of the density matrix and demonstrates fast convergence in many cases. However, there exists a counterexample of a state for which the sequence (2.133) diverges [Řeháček 2007]. Therefore, there remains a risk that the algorithm could fail for a particular experiment. In such cases, one can use the so-called "diluted" algorithm [Řeháček 2007]:

$$\hat{\rho}^{(k+1)} = \mathcal{N}\left[\frac{\mathbb{1}+\epsilon\hat{R}(\hat{\rho}^{(k)})}{1+\epsilon}\hat{\rho}^{(k)}\frac{\mathbb{1}+\epsilon\hat{R}(\hat{\rho}^{(k)})}{1+\epsilon}\right], \tag{2.134}$$

which depends on a single parameter ϵ that determines the "length" of the step in the parameter space associated with one iteration. For $\epsilon \to \infty$, we recover the iterative scheme (2.133). On the other hand, in the limit of $\epsilon \to 0$, the likelihood will monotonically increase, and the iterations will converge to the maximum-likelihood state, as shown in [Řeháček 2007]. Thus, if the standard iterative scheme (2.133) does not converge, one can use its diluted generalization (2.134) and choose sufficiently small ϵ to ensure the convergence.

The generalization of this approach to continuous-spectrum homodyne detection is given by [Lvovsky 2009]:

$$\hat{R}(\hat{\rho}) = \sum_m \frac{1}{\langle q^{\theta_m}|\hat{\rho}|q^{\theta_m}\rangle} |q^{\theta_m}\rangle\langle q^{\theta_m}|, \tag{2.135}$$

where $|q^\theta\rangle$ is an eigenvector of the quadrature operator $\hat{q}^\theta$ (1.84) and the summation is performed over each sample (q_m, θ_m) of the homodyne data. This operator, combined with the iterative scheme (2.134), offers a powerful tool for maximum-likelihood homodyne tomography of the density matrix.

In addition to MLE tomography, other parameter estimation techniques, such as least-square-based methods [Haah 2016; Acharya 2019; Guţă 2020] have demonstrated their ability to offer fast and efficient estimators for quantum states.

2.4.3 Other tomographic methods

Having some prior information about the state, one might find a Bayesian approach to quantum tomography [Blume-Kohout 2010; Chapman 2022] to be a valuable tool. This approach, grounded in Bayesian statistics, not only allows for efficient use of prior knowledge about the state but also inherently quantifies uncertainty in the reconstructed density matrix.

While both the Bayesian and maximum-likelihood (particularly EM) approaches to quantum tomography have proven efficient in many practical scenarios, their computational complexity can be significant, and the convergence rate can be low.

Another promising and rapidly evolving technique in quantum tomography is based on neural networks [Tiunov 2020; Lohani 2020; Koutný 2022]. This approach offers several advantages, including computational efficiency and robustness against noise. However, similar to other applications of neural networks, efficient utilization of this technique requires a large training dataset and significant computational resources during the training step. Additionally, neural networks are susceptible to overfitting, and interpreting the results of neural network computations, as well as assessing the quality of these results, can be challenging. On the other hand, by carefully selecting the training set, it is possible to incorporate additional information about the class of studied states. For example, training the network specifically on Gaussian states [Hsieh 2022b] can enhance its performance for this specific problem.

2.4.4 Moment-based quantum state characterization

If the studied states belong to a limited class of states (for example Gaussian states), they can typically be parameterized with a finite number of parameters $\vec{\theta}$. In this

scenario, state tomography can be efficiently formulated as a traditional parameter estimation problem with the measured state $\hat{\rho}(\vec{\theta})$. Having a good parametrization ensures that $\hat{\rho}(\vec{\theta})$ always remains a physical density operator. Limitations on the parameters θ_m are typically set individually, allowing the parameters to be considered independent of each other (unlike the elements of the density matrix).

For instance, a single-mode Gaussian state can be parametrized with five real-valued parameters: the mean quadratures $\langle\hat{q}\rangle$ and $\langle\hat{p}\rangle$, and the three parameters of the quadrature covariance matrix (κ, s, ψ_s), which define the purity, squeezing, and the direction of squeezing, as introduced in section 1.3.3. The physical constraints dictate that $\kappa \geq 1$ and $s > 0$. Therefore, if the parameters are sufficiently far from these boundaries, they can be considered unbounded, allowing for the construction of asymptotically unbiased estimators. In chapter 4 we employ the MoM to develop simple and efficient estimators for these parameters. We demonstrate that this estimator effectively utilizes the homodyne measurement data while requiring relatively low computational power, compared to optimization methods, such as MLE.

Chapter 3

Separation estimation problem

Contents

In this chapter, we introduce and analyze the central problem of this thesis: resolving two point-like sources of light. This problem, representing the simplest version of imaging, provides valuable insights into the role of coherence effects, quantum statistics of light, and the selection of measurement techniques for imaging.

The chapter begins with a general introduction to optical imaging, presenting important concepts such as the point-spread function, numerical aperture, and the

Rayleigh limit, following the classical textbooks on Fourier optics [Goodman 2005] and imaging [Barrett 2004]. Then, we formulate the problem of point source resolving and describe different approaches to the detection of light, namely direct imaging and spatial mode demultiplexing. We provide a review of the studies, that use the spatial mode demultiplexing for various imaging problems.

In the subsequent sections, we employ parameter estimation techniques to analyze the information content of different measurements in various scenarios. First, we study the resolving of a pair of mutually coherent sources, reproducing our results presented in [Karuseichyk 2022]. In the last section, we examine sources with partial mutual coherence, including cases of separation-dependent coherence and brightness of the sources [Karuseichyk 2024].

3.1 Elements of the imaging

3.1.1 Model of the optical system

To describe optical imaging, first, let us consider how light transforms as it travels through free space and lenses. In this study, we focus on analyzing quasi-monochromatic light with a fixed polarization. We start by describing a classical light beam localized near the optical z-axis in the plane $z = 0$, with wave vectors clustered around $\vec{k}_0$, parallel to the z-axis (paraxial approximation). According to Fourier optics, the propagation of light to the plane z can be described by the Huygens-Fresnel diffraction integral [Goodman 2005]:

$$E_z(\vec{\rho}, t) \propto \int d\vec{\rho}' \exp\left[ik_0 \frac{|\vec{\rho}' - \vec{\rho}|^2}{2z}\right] E_0(\vec{\rho}', t - z/c), \tag{3.1}$$

where $\vec{\rho}$ and $\vec{\rho}'$ represent transverse vectors in the planes z and $z = 0$ respectively, and $k_0 = |\vec{k}_0| = 2\pi/\lambda$ denotes the magnitude of the wave vector of the considered light beam. For simplicity, constant factors are omitted. Additionally, we'll disregard the time dependence, as our focus lies on imaging objects with properties that remain constant over time.

When examining propagation over distances significantly greater than the beam size (i.e., $\max(|\vec{\rho}|^2)/(\lambda z) \ll 1$), known as the *far field* regime, the following approximation can be employed [Goodman 2005]:

$$E_z(\vec{\rho}) \propto \exp\left[i\pi \frac{|\vec{\rho}|^2}{\lambda z}\right] \int d\vec{\rho}' \exp\left[-i\, 2\pi \frac{\vec{\rho}\cdot\vec{\rho}'}{\lambda z}\right] E_0(\vec{\rho}') = \\ = \exp\left[i\pi \frac{|\vec{\rho}|^2}{\lambda z}\right] F\left[E_0(\vec{\rho}')\right]\left(\frac{\vec{\rho}}{\lambda z}\right), \tag{3.2}$$

where

$$F[f(\vec{\rho}')](\rho) = \int d\rho' \exp\left[-i\, 2\pi \vec{\rho}\cdot\vec{\rho}'\right] f(\rho') \tag{3.3}$$

represents the two-dimensional Fourier transform. Thus, the paraxial propagation of light to the far field corresponds to the Fourier transform of the field distribution, with an additional quadratic phase factor. The phase factor corresponds to the

curvature of the wavefront and it is often neglected when the field is studied in the vicinity of the optical axis [Barrett 2004].

When a large lens with focal length f is positioned at the plane $z = 0$, it introduces to the propagating field an additional phase factor, that quadratically depends on $|\vec{\rho}'|$:

$$\Delta\phi = -\frac{\pi|\vec{\rho}'|^2}{\lambda f}. \tag{3.4}$$

Consequently, eq. (3.1) transforms to:

$$E_z(\vec{\rho}) \propto \int d\vec{\rho}' \exp\left[i\pi\frac{|\vec{\rho}' - \vec{\rho}|^2}{\lambda z}\right] \exp\left[-i\pi\frac{|\vec{\rho}'|^2}{\lambda f}\right] E_0(\vec{\rho}'). \tag{3.5}$$

If the observation plane aligns with the focal plane of the lens ($z = f$), eq. (3.5) simplifies to eq. (3.2). Therefore, the evolution of the field passing through the lens to the focal plane is also represented by a Fourier transform (3.2), with an additional quadratic phase factor.

Now, let us explore the conceptual imaging scheme illustrated in fig. 3.1. Despite its simplicity, this scheme can effectively describe nearly any passive far-field linear imaging. In this setup, the field distribution $E_O(\vec{\rho}_O)$ in the object plane ($z = 0$) gets imaged by the lens to the image plane ($z = z_1 + z_2$). The lens is positioned in the plane $z = z_1$ and is constrained by the aperture. The aperture modulates the transferred field by a position-dependent loss factor $A(\vec{\rho}_A) \leq 1$.

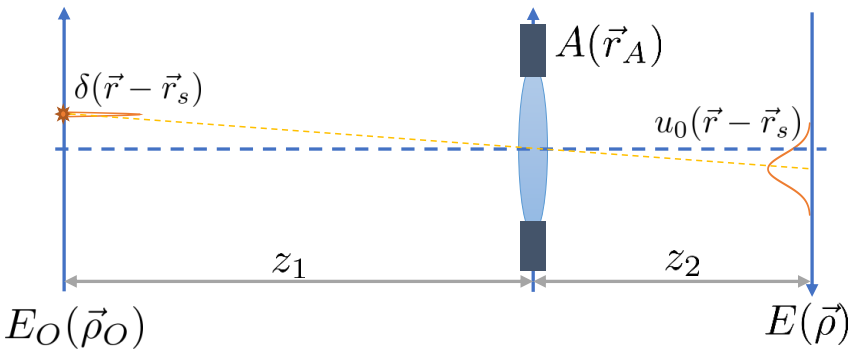

FIGURE 3.1: Fundamental imaging scheme. The optical field, described by the electric field distribution $E_O(\vec{\rho}_O)$ in the object plane $z = 0$, propagates over a distance z_1 in free space. It then passes through a lens, constrained by the aperture $A(\vec{\rho}_A)$, before further propagating over a distance z_2 to reach the image plane. The resulting field distribution in the image plane is denoted by $E(\vec{\rho})$.

The position of the image plane is given by the lens equation

$$\frac{1}{f} = \frac{1}{z_1} + \frac{1}{z_2}. \tag{3.6}$$

Hence, if the object is positioned sufficiently far from the lens ($z_1 \to \infty$)[1], the image is formed in the focal plane of the lens, meaning $z_2 = f$.

In this scenario, the field undergoes transformation (3.2) while propagating through the large distance z_1. It is then modulated by the aperture $A(\vec{\rho}_A)$ before undergoing another transformation (3.5) as it propagates through the lens to the image plane [Goodman 2005]. Neglecting the square phase factors the field distribution in the image plane takes the form

$$E(\vec{\rho}) \propto F\left[A(\vec{\rho}_A)F\left[E_O(\vec{\rho}_O)\right]\left(\frac{\vec{\rho}_A}{\lambda z_1}\right)\right]\left(\frac{\vec{\rho}}{\lambda z_2}\right). \tag{3.7}$$

In what follows, without loss of generality, we use a transformation to the coordinates $\vec{r}$, assuming that the corresponding length units in the object, aperture, and imaging planes are selected to compensate for the magnification factor $M = z_2/z_1$ of the imaging system, as well as for the factors λz_1 and λz_2 in eq. (3.7). Then, using the properties of the Fourier transform, it becomes straightforward to demonstrate that:

$$E(\vec{r}) \propto \int d\vec{r}' u_0(\vec{r} - \vec{r}')E_O(\vec{r}') = u_0(\vec{r}) * E_O(\vec{r}), \tag{3.8}$$

where $*$ denotes the convolution operation. The function

$$u_0(\vec{r}) \propto F\left[A(\vec{\rho}_A)\right](\vec{r}) \tag{3.9}$$

represents the *point spread function* (PSF) of the imaging system. Typically, for convenience, it is normalized as

$$\int d\vec{r}|u_0(\vec{r})|^2 = 1. \tag{3.10}$$

This function describes the image of a point source, meaning that for $E_O(\vec{r}) \propto \delta(\vec{r} - \vec{r}_s)$, the field distribution in the image plane is $E(\vec{r}) \propto u_0(\vec{r} - \vec{r}_s)$. Since each point of the object is imaged to a finite-sized spot, the images $E(\vec{r})$ of the objects $E_O(\vec{r})$ typically appear more blurred and less contrasted than the original objects.

Note that the transformation (3.8) by construction appeared to be translational invariant. This means that for a shifted object $E_O(\vec{r} + \vec{a})$, the corresponding image also shifts as $E(\vec{r} + \vec{a})$ without any distortions. This characteristic is specific to far-field imaging. Outside of the far-field scenario, when objects are sufficiently close to the lens, translational invariance does not hold perfectly all along the object and image planes. In such cases, various aberrations arise when the object is shifted far enough from the optical axis, and the transformation (3.8) no longer accurately describes the imaging process. Nevertheless, for objects localized close to the optical axis, the transformation (3.8) still provides a reasonably accurate description of the imaging, even when the objects are not located in the far field from the imaging setup. Some modifications to the definition of the PSF (3.9) are required in such cases [Goodman 2005]. In what follows, we will focus on far-field translationally-invariant imaging with the PSF defined in eq. (3.9). Note that breaking this approximation and introducing optical aberrations to the model may potentially increase

[1] The following analysis can also be performed for finite z_1, yielding analogous results, as discussed in section 6.3 of [Goodman 2005] and section 9.6 of [Barrett 2004].

the resolution of the system, as it introduces an additional mechanism for parameter encoding [Liang 2023a]. However, this effect lies outside the scope of this research.

3.1.2 Aperture and PSF

First, let us note that for real central-symmetric aperture functions, or at least when $A(-\rho_A) = A^*(\rho_A)$, the PSF (3.9) for far-field imaging is real-valued, $u_0(\vec{r}) \in \mathbb{R}$, which follows from the properties of the Fourier transform (3.3).

In most practical cases, the aperture of the imaging system takes on a circular form:

$$A(\vec{r}) = \mathrm{rect}_{[0,1]}\left(\frac{2|\vec{r}|}{D}\right), \tag{3.11}$$

where the rectangular function is defined in (1.17), and D represents the diameter of the aperture. This type of aperture is often referred to as a *sharp aperture*, owing to a sharp change of transmission at the boundary of the aperture. Due to the rescaling of coordinates in the aperture plane, the diameter in corresponding units is given by

$$D = \frac{2\mathrm{NA}}{\lambda}, \tag{3.12}$$

where NA stands for the *numerical aperture* of the imaging system. The numerical aperture is given by the following expression:

$$\mathrm{NA} = n \sin \alpha, \tag{3.13}$$

with n being the index of refraction of the medium in which the lens is working, and α is the half-angle of the maximum cone of light that can enter or exit the lens.

The corresponding PSF, given by the normalized Fourier transform of the eq. (3.11), is

$$u_0(\vec{r}) = \frac{1}{\sqrt{\pi}} \frac{J_1(2\pi \frac{\mathrm{NA}}{\lambda} |\vec{r}|)}{|\vec{r}|}, \tag{3.14}$$

where $J_1(r)$ is the first order Bessel function of the first kind. This function is often referred to as *Airy PSF*, and its square as *Airy disk*. One can see that for larger values of the numerical aperture, the PSF is more localized, signifying a higher quality imaging system, with smaller blurring of the image.

Rayleigh limit

A point source of light, located in the position $\vec{r}_s$ of the object plane, produces the following intensity distribution in the image plane

$$I_0(\vec{r}) \propto |u_0(\vec{r} - \vec{r}_s)|^2, \tag{3.15}$$

with $|u_0(\vec{r})$ being the PSF defined in (3.9).

In the case of two incoherent point sources, their individual intensity distributions are summed up, since no interference is present. Thus, the image of two

equally bright incoherent sources located in positions $\vec{r}_{1,2} = (\pm d/2, 0)$ is

$$I(\vec{r}) \propto |u_0(\vec{r} - \vec{r}_1)|^2 + |u_0(\vec{r} - \vec{r}_2)|^2. \tag{3.16}$$

We present the crosssections of this image, taken along the y-axis, for different source separations d on fig. 3.2.

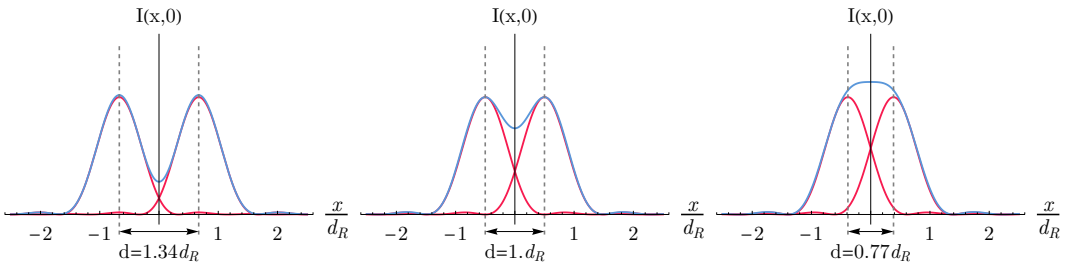

FIGURE 3.2: The cross-section of an image of two incoherent sources with different separations: classically resolved sources (left), sources on the border of Rayleigh criteria (middle), and sources not resolvable according to Rayleigh criterion (right).

As the separation d becomes smaller, the images of the sources merge, making it harder to distinguish individual sources. The heuristic bound, introduced by Lord Rayleigh 145 years ago claims that the smallest resolvable separation is the one for which the central maximum of one source's image coincides with the first minimum of another source's image [Rayleigh 1879; Villiers 2016]. Thus, using the first root of the Bessel function $J_1(3.83...) = 0$, one can show that the *Rayleigh limit* is given by

$$d_R = 1.22 \frac{\lambda}{2\,\mathrm{NA}}. \tag{3.17}$$

This separation corresponds to $\sim 20\%$ dip in the intensity distribution, as shown on the middle plot in fig. 3.2.

The Rayleigh criterion is based on limitations to the resolving capabilities of the human visual system and to provide sufficient contrast for an observer to distinguish two separate objects in the image. The Rayleigh criterion is therefore not a fundamental physical law and instead a somewhat arbitrarily defined value. This was clearly stated by Rayleigh himself in 1879 [Rayleigh 1879]:

> This rule is convenient on account of its simplicity; and it is sufficiently accurate in view of the necessary uncertainty as to what exactly is meant by resolution.

There are other visual heuristic criteria. For example, the *Sparrow limit* [Sparrow 1916], initially introduced for spectroscopy, is defined as the separation between two point emitters when the total PSF exhibits no dip in intensity at the midpoint, but instead displays an intensity plateau. The Sparrow limit is, therefore, smaller than that defined by Rayleigh and is given by:

$$d_S = 0.94 \frac{\lambda}{2\,\mathrm{NA}} = 0.77\, d_R. \tag{3.18}$$

The image of the sources separated by d_S is presented on the right plot of the fig. 3.2.

Another example of a resolution criterion is *Abbe limit* [Abbe 1873], defined simply as

$$d_A = \frac{\lambda}{2\,\mathrm{NA}} = 0.82\,d_R. \tag{3.19}$$

A more practical definition of imaging resolution is to use the *full width at half maximum* (FWHM) of the point spread function, in the case of Airy PSF (3.14) it is equal to:

$$d_{FWHM} = 1.02\frac{\lambda}{2\,\mathrm{NA}} = 0.84\,d_R. \tag{3.20}$$

The advantage of using the FWHM is that it can be easily measured in the lab by imaging a pseudo-point emitter, thus serving as a comparison metric for real microscope systems.

Note that all these visual criteria are defined for incoherent sources of equal brightness. For a pair of sources sharing arbitrary mutual coherence γ defined in eq. (1.21), the intensity distribution in the image plane takes the form:

$$I(\vec{r}) \propto |u_0(\vec{r}-\vec{r}_1)|^2 + |u_0(\vec{r}-\vec{r}_2)|^2 + 2\,\mathrm{Re}[\gamma\, u_0(\vec{r}-\vec{r}_1)u_0^*(\vec{r}-\vec{r}_2)]. \tag{3.21}$$

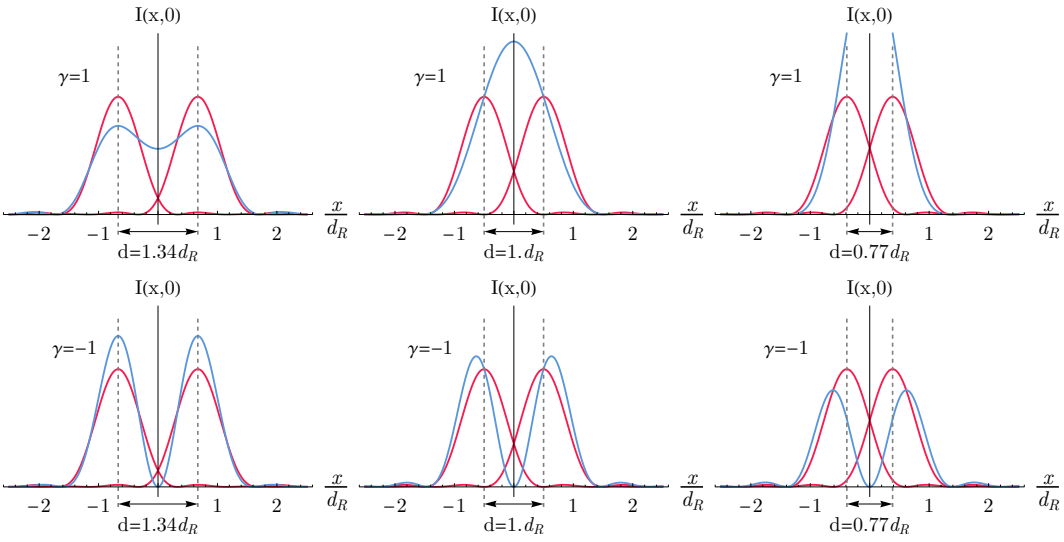

FIGURE 3.3: The cross-section of an image of two mutually coherent sources with different separations. The top row depicts in-phase sources ($\gamma = 1$), and the bottom row depicts anti-phase sources ($\gamma = -1$). Red lines correspond to the individual images of the sources, and blue lines represent the images of both sources interfering with each other.

We present several examples of images of mutually coherent sources in fig. 3.3. As one can see from this figure, it is more challenging to visually resolve sources that constructively interfere, i.e., when $\gamma = 1$. In this case, an analogous criterion to the Rayleigh limit for coherent sources can be formulated to give approximately the

same 20% intensity dip between the two peaks of intensity, which corresponds to:

$$d_R^{(\gamma=1)} = 1.64\frac{\lambda}{2\,\mathrm{NA}} = 1.34\,d_R. \tag{3.22}$$

The analogous analysis can be performed for any value of mutual coherence γ. For example, sources with an imaginary value of mutual coherence $\gamma \in \mathbb{I}$ (corresponding to a phase shift of $\pi/2$) result in the same intensity distribution as for incoherent sources. If the sources have opposite phases ($\gamma = -1$), they interfere destructively, and there is always a dark line separating their images (see the bottom row of plots in fig. 3.3). Thus, from the perspective of visual resolution criteria, anti-phase coherent sources are always resolvable. However, this conclusion does not account for the decrease in intensity due to destructive interference, indicating that visual criteria may fall short in this example.

The Rayleigh criterion, like other visual criteria, describes the perception of the image by the human eye. It does not imply any post-processing of the image or alternative ways of measuring the light in the image plane, different from straightforward intensity detection. By approaching this problem with a wider array of instruments, one may aim to resolve sources that are considered unresolvable by visual criteria. This problem is traditionally referred to as *sub-Rayleigh imaging* or *super-resolution imaging*. One of the ways to surpassing the Rayleigh limit is to consider the imaging problem as a parameter estimation problem, which is the main approach used in this chapter.

Gaussian PSF

The widely used approximation for the PSF in imaging is a Gaussian approximation. Since the central spot of the Airy disk contains most of the light, and its shape resembles the Gaussian function, researchers often employ the Gaussian model of the PSF [Zhang 2007]. This Gaussian PSF is mathematically represented by the equation:

$$u_0(\vec{r}) = \sqrt{\frac{1}{2\pi\sigma^2}}\exp\left[-\frac{|\vec{r}|^2}{4\sigma^2}\right], \tag{3.23}$$

where

$$\sigma = 0.42\frac{\lambda}{2\,\mathrm{NA}} = 0.34\,d_R. \tag{3.24}$$

A comparison between the Airy PSF (3.14) and the Gaussian PSF (3.23) is depicted in fig. 3.4.

The Gaussian PSF model is widely used in the literature due to its numerous advantageous mathematical properties, which simplifies analytical calculations. Within this thesis, many results are obtained for an arbitrary (yet real and central symmetric) shape of the PSF, but final calculations for plots are mostly performed using the Gaussian PSF. The Gaussian PSF model is often referred to as *soft aperture approximation* because the aperture responsible for this PSF has a Gaussian profile, making it smoother compared to a sharp circular aperture (3.11). Note that numerical aperture is strictly defined only for sharp apertures. In what follows we use the width

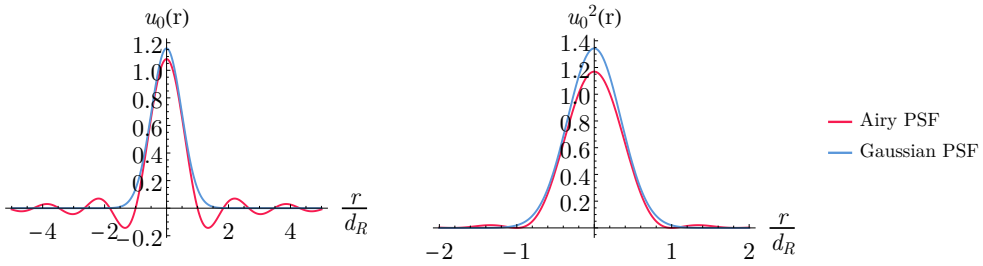

FIGURE 3.4: The comparison between the Airy PSF (red line) and the Gaussian PSF (blue line).

of the Gaussian PSF σ as the primary characteristic of the optical system's resolution (given a particular wavelength).

PSF factorization

One of the advantageous properties of the Gaussian function is its factorizability:

$$u_0(x,y) = \sqrt{\frac{1}{2\pi\sigma^2}} \exp\left[-\frac{(x^2+y^2)}{4\sigma^2}\right] = u_0(x)\, u_0(y), \tag{3.25}$$

where

$$u_0(x) = \frac{1}{(2\pi\sigma^2)^{1/4}} \exp\left[-\frac{x^2}{4\sigma^2}\right]. \tag{3.26}$$

Thus, when considering imaging of point sources located on the x-axis, the field distribution over the y-axis will always be represented as a separate factor $u_0(y)$. Consequently, the problem can consistently be reduced to a single dimension x, as we demonstrate in more detail in section 3.3.

3.1.3 Losses in the imaging system

When light propagates through a diffraction-limited imaging system, a part of it gets absorbed by the aperture. Typically, we consider the imaging of point-like sources that emit light uniformly in all directions. Consequently, the energy transmissivity κ of the imaging setup is determined by the solid angle of the aperture Ω_A, and can be expressed as

$$\kappa = \frac{\Omega_A}{4\pi} = \sin^2\frac{\alpha}{2} \approx \frac{\mathrm{NA}^2}{4}, \tag{3.27}$$

where we assume that imaging is conducted in air or vacuum (i.e., the index of refraction $n \approx 1$), the point source is situated sufficiently close to the optical axis, and the numerical aperture of the lens is small ($\mathrm{NA} \ll 1$).

In a general scenario, the reflection and absorption of light on the lenses can introduce additional losses. Moreover, the finite efficiency of the detection system can also be considered at this stage by integrating the detectors' efficiency into κ. Thus, in the following, we simply use constant κ to account for all the losses in the system. Since we employ a translationally symmetric model of imaging, the

transmissivity κ remains constant regardless of the position of the point source. It is interesting to note that the total transferred energy may still depend on the position of point sources when several mutually coherent sources are imaged (see section 3.3).

Now, we can formulate the description of the imaging transformation using the language of quantum physics. If a point source located at position $\vec{r}_s$ emits light into mode $\hat{s}$ with the spatial transverse profile $\delta(\vec{r} - \vec{r}_s)$, it excites mode $\hat{c}$ in the image plane with the spatial profile $u_0(\vec{r} - \vec{r}_s)$. This implies that losses on the aperture not only transform the state of light but also alter the spatial shape of modes. Referring to the description of losses introduced in section 1.2.4, passing through the imaging system corresponds to the following transformation of the field operators [Lupo 2016]:

$$\hat{c} = \sqrt{\kappa}\hat{s} + \sqrt{1-\kappa}\hat{v}, \tag{3.28}$$

where $\hat{v}$ corresponds to the field operator of the vacuum mode. Hereafter we refer to the mode $\hat{s}$ as the *source mode*, and the mode $\hat{c}$ as the *image mode*.

3.2 Formulation of the source resolving problem

Now we have all the necessary tools to formulate the central problem of this thesis: the imaging of two point-like sources of light. The conceptual scheme of this problem is illustrated in fig. 3.5.

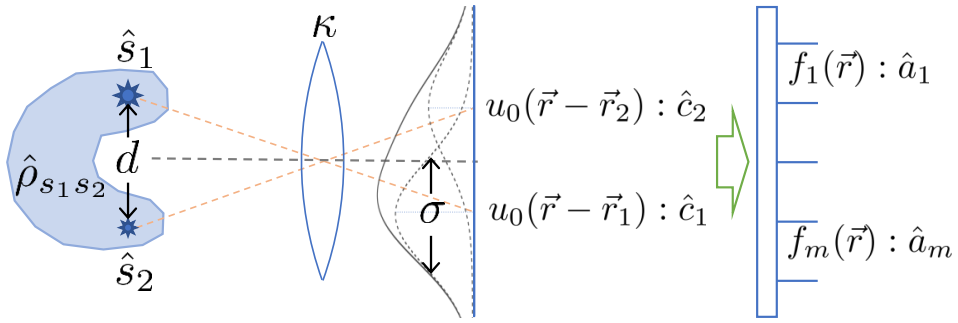

FIGURE 3.5: The imaging of two point sources. The sources emit light into the modes $\hat{s}_{1,2}$, with the joint quantum state of the emitted light denoted as $\hat{\rho}_{s_1 s_2}$. The emitted light passes through a diffraction-limited imaging system with transmissivity κ and PSF $u_0(\vec{r})$. The resulting state of light is represented in the pair of non-orthogonal image modes $\hat{c}_{1,2}$, or in the set of orthogonal measurement modes $\{\hat{a}_m\}$.

In the considered imaging scheme, the two point sources are located at positions $\vec{r}_{1,2}$. The parameter of primary interest is the separation between these sources $d = |\vec{r}_2 - \vec{r}_1|$. In the subsequent discussion, we refer to the determination of the separation d as the *resolving* of the sources. This terminology slightly deviates from its traditional meaning, which typically involves choosing between two hypotheses: whether there is one or two sources in the image plane. Although these two problems are closely intertwined and share many common techniques [Lu 2018;

Schlichtholz 2024; Wadood 2024], this research primarily focuses on the task of estimating the separation.

Following the conventional approach, we assume that the position of the *centroid* of the sources, denoted by $\vec{r}_c = (\vec{r}_1 + \vec{r}_2)$, is known [Tsang 2016; Lupo 2016; Sorelli 2021a]. For the sake of simplicity in the presentation, we further assume that the sources lie along the x-axis, i.e., $\vec{r}_{1,2} = (x_{1,2}, 0)^T$.

3.2.1 State of the emitted light

The sources emit light into two spherical modes $\hat{s}_{1,2}$, which are orthogonal to each other. The temporal profile of the modes $\hat{s}_{1,2}$ is defined by the coherence time τ_c of the sources. The joint quantum state of the emitted light, represented by $\hat{\rho}_{s_1 s_2}$, determines both the individual statistics of the sources and their correlations. The coherency matrix (1.52) of the emitted field is Hermitian and has dimensions of 2x2, allowing it to be parameterized with four real parameters:

$$\Gamma^{(1)\hat{s}}_{jk} = \left\langle \hat{s}^\dagger_j \hat{s}_k \right\rangle_{\rho_{s_1 s_2}} = \begin{pmatrix} N_S\, p & N_S\sqrt{p(1-p)}\,\gamma_0 e^{i\phi} \\ N_S\sqrt{p(1-p)}\,\gamma_0 e^{-i\phi} & N_S(1-p) \end{pmatrix}_{jk}. \tag{3.29}$$

Here, $N_S = \operatorname{Tr} \mathbf{\Gamma}^{(1)\hat{s}}$ represents the total number of emitted photons per coherence time τ_c, $0 \leq p \leq 1$ denotes the power imbalance between the sources (where they are equally bright if $p = 1/2$), and $\gamma = \gamma_0 e^{i\phi}$ is the mutual coherence (1.53) between the sources. We naturally refer to the parameter $\gamma_0 = |\gamma|$ as the absolute value of the mutual coherence and to ϕ as the mutual phase between the sources, even if they are not in a coherent state.

As we demonstrate in the following sections, all the field characteristics after passing through the imaging system contain the parameter N_S only in combination with the transmissivity factor κ. Thus the important parameter of this problem is the combination κN_S, which defines the brightness[2] of the sources imaged with a given optical system. We define sources with $\kappa N_S \ll 1$ as faint sources. Note, however, that the number of photons emitted by such sources per unit time or per integration time of the detection system can still be large if the coherence time of the sources is short, or if transmissivity κ is low. Unlike most research on this topic, our approach also allows for studying bright sources with $\kappa N_S > 1$, demonstrating the non-trivial scaling of the separation estimation sensitivity with this parameter.

Depending on the problem's setup, some or all of parameters (N_s, p, γ_0, ϕ) of the state may be known before measurement. In certain cases, we treat them as unknown parameters, however, we always assume to know the absolute degree of mutual coherence γ_0[3]. Consequently, the problem naturally divides as follows: fully coherent sources ($\gamma_0 = 1$) are addressed in section 3.3, and partially coherent sources ($0 \leq \gamma_0(d) \leq 1$) in section 3.4.

The coherency matrix (3.29) does not entirely define the state $\hat{\rho}_{s_1 s_2}$. In general, it depends on infinitely many parameters. However, our primary focus lies on passive imaging, where the sources emit light without external illumination. Consequently,

[2] We assume the coherence time of the sources τ_c to be fixed. Thus the number of photons per coherence time is proportional to the number of photons per unit time.

[3] This parameter can also be estimated from the measurement [Liang 2021].

the sources lack any external phase reference, and their state is phase-averaged over the global phase. As a result, all the averages that do not contain an equal amount of creation and annihilation operators vanish:

$$\langle \hat{s}_j \rangle = \langle \hat{s}_j \hat{s}_k \rangle = \langle \hat{s}_j^\dagger \hat{s}_k \hat{s}_l \rangle = ... = 0. \tag{3.30}$$

In section 3.4 we also assume the statistics of the sources to be Gaussian. In this case, the eqs. (3.29) and (3.30) fully define the state $\hat{\rho}_{s_1 s_2}$. In section 3.3 we explore a more general case, accounting for arbitrary statistics of the sources.

3.2.2 Light transformation in the imaging system

Now let us consider how the emitted light passes through a diffraction-limited imaging system with transmissivity κ and PSF $u_0(\vec{r})$. In the image plane two image modes $\hat{c}_{1,2}$ with spatial profiles $u_0(\vec{r} - \vec{r}_{1,2})$ are excited

$$\hat{c}_{1,2} = \sqrt{\kappa}\hat{s}_{1,2} + \sqrt{1-\kappa}\hat{v}_{1,2}. \tag{3.31}$$

Note that the vacuum modes $\hat{v}_{1,2}$ are not generally orthogonal to each other, as the sources experience losses to overlapping modes. Consequently, the modes $\hat{c}_{1,2}$ also become non-orthogonal to each other. This non-orthogonality can also be observed through the overlap of their spatial profiles:

$$\delta = \int d\vec{r}\, u_0(\vec{r} - \vec{r}_1) u_0^*(\vec{r} - \vec{r}_2) \neq 0. \tag{3.32}$$

Thus, while the light in the imaging plane can be represented by the pair of image modes $\hat{c}_{1,2}$, it is not always convenient due to their non-orthogonality and dependence on the positions of the sources. One solution to this challenge is to use symmetric and anti-symmetric combinations of the image modes [Lupo 2016; Sorelli 2021b; Sorelli 2022]:

$$u_\pm(\vec{r}) = \frac{u_0(\vec{r} - \vec{r}_1) \pm u_0(\vec{r} - \vec{r}_2)}{2(1 \pm \delta)}. \tag{3.33}$$

This pair of modes is typically advantageous for calculating the QFI since it constitutes the smallest set of orthogonal modes capable of representing the state of light in the image plane. However, when computing the FI and the MoM sensitivity of specific measurements the state should be represented in the same mode basis as the measurements are performed.

Let us examine the measurements of light in the image plane, conducted within the set of orthogonal modes $\hat{a}_m$ with spatial profiles $f_m(\vec{r})$. These modes are referred to as the *measurement modes*. The imaging system under consideration performs a linear transformation of the field, allowing the field operators $\hat{a}_m$ to be expressed as:

$$\hat{a}_m = z_m^{(1)}\hat{s}_1 + z_m^{(2)}\hat{s}_2 + \hat{v}'_m, \tag{3.34}$$

where $z_m^{(1,2)}$ are complex coefficients, and $\hat{v}'_m$ represents vacuum modes' field operators, which we do not normalize for the simplicity of the presentation.

We determine the coefficients $z_m^{(1)}$ by considering the scenario where mode $\hat{s}_2$ is in the vacuum state. In this case, the field in the image plane is fully described by the single mode $\hat{c}_1$ with the spatial profile $u_0(\vec{r}-\vec{r}_1)$. This mode can be decomposed over the orthogonal set of measurement modes $f_m(\vec{r})$. Hence, following the mode transformation rule (1.29), the field operators of the measurement modes in the case of a single light source $\hat{s}_1$ are represented as:

$$\hat{a}_m^{(1)} = \left(\int d\vec{r}\, f_m^*(\vec{r}) u_0(\vec{r}-\vec{r}_1) \right) \hat{c}_1 + \hat{v}_m^{(1)} = \beta_m(\vec{r}_1)\hat{c}_1 + \hat{v}_m^{(1)}, \tag{3.35}$$

where $\hat{v}_m^{(1)}$ are the field operators of the non-excited (vacuum) modes in the image plane. Substituting expression (3.31) for $\hat{c}_1$ into this equation and comparing the result to eq. (3.34), we arrive at $z_m^{(1)} = \sqrt{\kappa}\beta_m(\vec{r}_1)$, where $\beta_m(\vec{r}_1)$ are the spatial overlaps of the first source image mode with the measurement mode $f_m(\vec{r})$

$$\beta_m(\vec{r}_1) = \int d\vec{r}\, f_m^*(\vec{r}) u_0(\vec{r}-\vec{r}_1). \tag{3.36}$$

Similarly, we find coefficients $z_m^{(2)} = \sqrt{\kappa}\beta_m(\vec{r}_2)$ by considering mode $\hat{s}_1$ in the vacuum state, yielding the final result [Karuseichyk 2022]:

$$\boxed{\hat{a}_m = \sqrt{\kappa}\sum_j \beta_m(\vec{r}_j)\, \hat{s}_j + \hat{v}'_m.} \tag{3.37}$$

The vacuum contribution here arises from two origins: modes $\hat{v}_{1,2}$ enters the system due to the losses on the aperture described in eq. (3.31), and modes $\hat{v}_m^{(1,2)}$ corresponding to non-excited modes in the image plane, as given by eq. (3.35).

The contribution of the vacuum modes in this expression can be eliminated for the averages of the operators in the *normal order*. The normal order of the operator corresponds to the form where all creation operators are to the left of all annihilation operators in the products. For example, the operator $\hat{a}^\dagger\hat{a}$ is normally ordered, and the operator $\hat{a}\hat{a}^\dagger$ is not. The latter can be transformed into the normal order by using the commutation relations (1.26)[4]

$$\hat{a}\hat{a}^\dagger = \hat{a}^\dagger\hat{a} + \hat{I}. \tag{3.38}$$

If the operator in the normal order contains one or several field operators of the vacuum modes, its mean value equals zero, since

$$\hat{v}|0\rangle = 0, \text{ and } \langle 0|\hat{v}^\dagger = 0. \tag{3.39}$$

Thus if one is interested in finding the average of some operator $g(\{\hat{a}_m^\dagger\},\{\hat{a}_n\})$, this operator should be transformed using the commutation relations (1.26) into the normal-ordered form $g_N(\{\hat{a}_m^\dagger\},\{\hat{a}_n\}) = g(\{\hat{a}_m^\dagger\},\{\hat{a}_n\})$. Then, the mean value of this

[4]This transformation to normal order should not be confused with the normal ordering operation $:\hat{X}:$ [Scully 1997]. The latter changes the order of the field operators to normal order, without respecting the commutation rules, i.e., $:\hat{a}\hat{a}^\dagger: = \hat{a}^\dagger\hat{a}$, which, in the general case, changes the operator $:\hat{a}\hat{a}^\dagger: \neq \hat{a}\hat{a}^\dagger$.

operator can be calculated without considering any additional vacuum modes:

$$\boxed{\langle\, g_N(\{\hat{a}_m^\dagger\},\{\hat{a}_n\})\rangle = \langle\, g_N(\{\sqrt{\kappa}\sum_j \beta_m^*(\vec{r}_j)\hat{s}_j^\dagger\},\{\sqrt{\kappa}\sum_j \beta_m(\vec{r}_j)\hat{s}_j\})\rangle_{\rho_{s_1 s_2}},} \tag{3.40}$$

with $\beta_m(\vec{r}_j)$ (3.36) containing all the geometry of the problem. This useful property allows us to build an analytical description of the moment-based parameter estimation throughout this research.

Now let us consider the case when the PSF model $u_0(\vec{r})$ and the measurement modes $f_m(\vec{r})$ are factorable over the coordinates x and y, i.e. $u_0(\vec{r}) = u_0^{(x)}(x)u_0^{(y)}(y)$ and $f_m(\vec{r}) = f_m^{(x)}(x)f_m^{(y)}(y)$. In this case the overlap coefficients (3.36) can be represented as

$$\beta_m(\vec{r}_j) = \left(\int dx\, f_m^{(x)*}(x)u_0^{(x)}(x-x_j)\right)\left(\int dy\, f_m^{(y)*}(y)u_0^{(y)}(y-y_j)\right). \tag{3.41}$$

Then, if the sources are located on the x-axis ($y_j = 0$), and the spatial profile of the measurement modes along the y-axis matches the profile of the PSF $f_m^{(y)}(y) = u_0^{(y)}(y)$, then the second integral in eq. (3.41) always equals 1, and the problem becomes single dimensional.

3.2.3 Measurement techniques

There are two main types of measurements that can be performed on individual optical modes $\hat{a}_m$: intensity measurement (see section 1.2.4) and quadrature measurements (see section 1.3.4). Quadrature measurements require the presence of the local oscillator (LO), i.e., a strong coherent field that is phase locked with the sources. However, as discussed in section 3.2.1, in our case, we usually do not have a phase reference to the sources. Therefore, the emitted light is considered to be phase-averaged, and quadrature measurements are not very efficient in this scenario. Although useful information can still be extracted from the second moment of the measured quadratures, this approach has been shown to be not quantum-optimal even in the simplest cases [Yang 2017; Len 2020; Datta 2021].

In this thesis, our primary focus is on the measurement of light intensity

$$\hat{X}_m = \hat{N}_m = \hat{a}_m^\dagger \hat{a}_m. \tag{3.42}$$

This represents the most fundamental optical measurement, as even quadrature detection involves several of such photodetectors (see fig. 1.2). Moreover, ideal intensity detection, when combined with an optimized choice of measurement modes, has been demonstrated to saturate the QCRB for the separation estimation between incoherent sources [Tsang 2019b; Sorelli 2021b]. Therefore, the next step is to select the measurement mode basis $f_m(\vec{r})$.

Direct imaging

The most straightforward method for examining the light distribution in the image plane is to employ a camera, which consists of an array of small localized photodetectors. This approach is called *direct imaging* (DI). This measurement corresponds to intensity detection in the *pixel modes*, which are rectangular modes of size $l_x \times l_y$:

$$f_m^{DI}(x,y) = \mathrm{rect}_{[x_m^{pix}-l_x/2, x_m^{pix}+l_x/2]}(x)\ \mathrm{rect}_{[y_m^{pix}-l_y/2, y_m^{pix}+l_y/2]}(y), \tag{3.43}$$

where x_m^{pix} and y_m^{pix} represent the coordinates of the center of pixel with index m, and the rectangular function is defined in eq. (1.17). If the size of the pixels is sufficiently smaller than the PSF width, $l_{x,y} \ll \sigma$, the measurement modes can be approximated with delta-functions

$$f_m^{DI}(x,y) = \sqrt{l_x}\delta(x - x_m^{pix})\sqrt{l_y}\delta(y - y_m^{pix}). \tag{3.44}$$

This approximation works since the field does not change significantly on the scale of the pixel size $l_{x,y}$, thus it can be just sampled in the center of the pixel, instead of averaging over the pixel.

In this case, the overlaps (3.36) reduce to

$$\beta_m^{DI}(x_j, y_j) = \sqrt{l_x l_y}\, u_0(x_m^{pix} - x_j, y_m^{pix} - y_j). \tag{3.45}$$

The mean number of photons, detected in the pixel m is given by eq. (3.40):

$$\begin{aligned} N_m = \langle \hat{a}_m^\dagger \hat{a}_m \rangle = \kappa l_x l_y \langle (u_0^*(x_m^{pix} - x_1, y_m^{pix} - y_1)\hat{s}_1^\dagger + u_0^*(x_m^{pix} - x_2, y_m^{pix} - y_2)\hat{s}_2^\dagger) \\ \times (u_0(x_m^{pix} - x_1, y_m^{pix} - y_1)\hat{s}_1 + u_0(x_m^{pix} - x_2, y_m^{pix} - y_2)\hat{s}_2)\rangle. \end{aligned} \tag{3.46}$$

By defining the light intensity as an average detected number of photons per unit square $I(\vec{r}_m^{pix}) = N_m/(l_x l_y)$, we reproduce a classical formula for intensity distribution from two sources with coherency matrix (3.29):

$$I(\vec{r}) = \kappa N_S \left(p\, |u_0(\vec{r} - \vec{r}_1)|^2 + (1-p)|u_0(\vec{r} - \vec{r}_2)|^2 + 2\,\mathrm{Re}[\gamma u_0^*(\vec{r} - \vec{r}_1)u_0(\vec{r} - \vec{r}_2)] \right). \tag{3.47}$$

Often it is convenient to consider DI in the continuous limit of infinitely small pixels $l_{x,y} \to 0$. In this case, the summation over the index m of the measurement mode is changed to the integration over the positions of pixels

$$\sum_m g_1(N_m) = \int d\vec{r}\, \frac{1}{l_x l_y} g_1(I(\vec{r}) l_x l_y) = \int d\vec{r}\, g_1(I(\vec{r})), \tag{3.48}$$

where the size of the pixel $l_x l_y$ cancels out for the first-order homogenous functions $g_1(ax) = a g_1(x)$.

This continuous approximation is often referred to as *continuous direct imaging*. It is important to note that this idealized model may tend to overestimate the performance of real cameras, which have pixels of finite size and are subject to detection

noise. However, it serves as a useful upper bound for the performance of any camera, enabling comparison of other measurement techniques to this idealized DI.

The DI technique is a widely used and well-established approach to imaging. It offers numerous advantages, from the availability of cameras to translational covariance of the measurement in the xy-directions. Thus, we often use the sensitivity of DI as a reference for comparison.

It is important to note that information about the separation of incoherent sources, provided by DI measurements, vanishes as the separation decreases to zero. This phenomenon has been recently termed *Rayleigh's curse* [Tsang 2016]. It is a specific feature of DI measurement and is not replicated in the QFI behavior. In other words, DI is not a quantum-optimal measurement for resolving two closely located incoherent sources [Tsang 2016]. Thus, let us consider more general measurement modes.

Mode demultiplexing and its realization

A more generalized approach to measuring light in the image plane extends beyond localized pixel modes, enabling access to modes with arbitrary spatial shapes. An illustrative example of mode-selective measurement is homodyne detection, wherein the spatial (and temporal) profile of the measured mode is determined by the mode profile of the LO. Thus, by altering the spatial profile of the LO, one can select the mode being measured. Theoretical analyses of this approach to resolving of the point sources have been conducted [Len 2020; Datta 2021; Datta 2020], and experimental implementations have been achieved [Pushkina 2021; Costa Filho 2021]. A limitation of this technique is its ability to access only one spatial mode at a time. Moreover, as previously noted, homodyne measurement is not quantum-optimal for the separation estimation problem.

Thus within this thesis, we are focused on the mode-selective intensity measurements. One can not perform the intensity measurement in the spatially extended modes if they overlap with each other, as we discussed in section 1.2.4. Consequently, it is necessary to perform some operation, that separates different spatial modes in space (or time). This operation is in general referred to as *mode sorting*, or specifically for spatial modes it is called *spatial mode demultiplexing* (SPADE).

One of the types of devices with spatial mode sorting capabilities are the photonic lanterns [Leon-Saval 2013; Birks 2015]. These devices consist of single-mode fiber cores being gradually reduced until they cannot confine the light anymore. The light is then coupled into the multimode waveguide (see fig. 3.6). In the reverse direction it can work as a spatial mode sorter. Note that designing a photonic lantern that can decompose the light in a specific basis of spatial modes can be a challenging task.

Another technology for spatial mode sorting is known as *multi-plane light conversion* [Labroille 2014; Kupianskyi 2023; Zhang 2023b]. It operates on the principle that any spatial unitary transformation can be achieved by cascading a finite number of phase masks and Fourier transforms [Morizur 2010]. Typically, N phase masks are needed to transform N modes. However, with the recently developed "magic mapping," demultiplexing of 1035 modes was achieved using only 14 phase masks [Fontaine 2021].

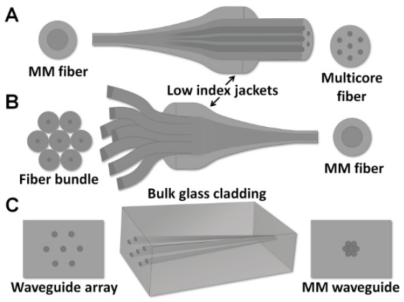

FIGURE 3.6: Photonic lantern implementations [Leon-Saval 2013].

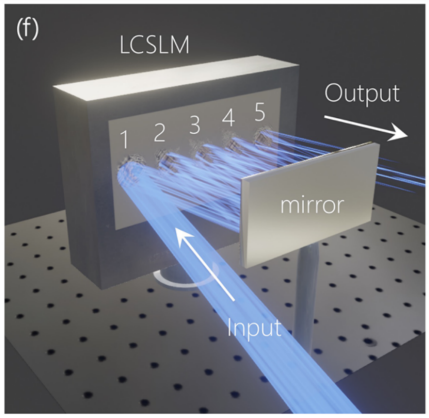

FIGURE 3.7: Multi-plane light converter [Kupianskyi 2023].

Multi-plane light conversion is often implemented using several reflections on a spatial light modulator (SLM), allowing programmable unitary transformations, i.e. adaptive choice of the mode decomposition basis (see fig. 3.7). However, in efforts to reduce losses and price, devices can also use reflective phase plates with fixed properties [Bade 2018].

The multi-plane light conversion technique offers high precision of the mode decomposition, and it can be applied to any spatial-mode transformation. Thanks to these remarkable features, multi-plane light conversion is a preferred tool for various applications, especially in fibered and free-space communications, but also for point sources resolving [Rouvière 2024; Santamaria 2024; Tan 2023b; Boucher 2020].

Another recent promising mode-sorting technique leverages the intricate spatial mode mixing process within a multimode fiber [Defienne 2020] (fig. 3.8). Remarkably, this method demonstrated the decomposition of over 25 spatial modes using just one phase mask. However, it is noteworthy that the reported conversion efficiency of this approach was approximately 20%.

The final approach to SPADE that we mention is rooted in nonlinear optics. It has been shown that by tuning the spatial shape of the pump, mode-selective conversion of light can be achieved through a sum-frequency generation process [Sephton 2019; Kumar 2021]. When combined with spectral filtering, this method can facilitate spatial-mode-selective measurement.

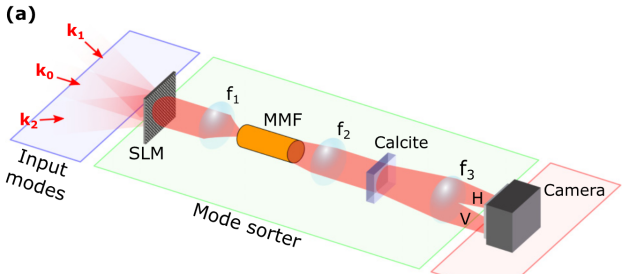

FIGURE 3.8: Spatial mode sorting with SLM and multimode fiber [Defienne 2020].

Mathematical description of SPADE

Now equipped with the tools to conduct SPADE, we can investigate its application to the point source resolving problem. In the subsequent discussions, we consider symmetric, real, and factorable PSF $u_0(x,y) = u_0(x)u_0(y) = u_0^*(x,y)$. If the sources are located along the x-axis, i.e., $\vec{r}_{1,2} = (x_{1,2}, 0)$, the problem can be effectively reduced to one dimension, by choosing the measurement basis as $f_m(x,y) = f_m(x)u_0(y), m = 1...K$, as demonstrated in eq. (3.41).

It has been demonstrated in [Rehacek 2017] that arbitrary measurement basis with the symmetry

$$f_m^{(s)}(-x) = (-1)^m f_m^{(s)}(x),\ m = 1...K \tag{3.49}$$

saturates the QCRB for resolving faint incoherent sources in the limit of $K \to \infty$, i.e. when the measurement basis is full. Note that this measurement basis is efficient only when it is aligned to the centroid of the sources. This means that the sources should be positioned symmetrically with respect to the origin of the chosen coordinate system, i.e. $\vec{r}_1 = -\vec{r}_2$. Following the conventional approach, we maintain this assumption in subsequent calculations. Achieving the sensitivity of the centroid-aligned measurement can be challenging in practical cases, but this example serves as a valuable lower bound for the FI and QFI.

The positions of sources $\vec{r}_{1,2} = (\pm d/2, 0)$ are incorporated into the problem through the overlaps:

$$\beta_m = \beta_m\left(\frac{d}{2}\right) = \int dx\, u_0\left(x - \frac{d}{2}\right) f_m^{(s)}(x), \tag{3.50}$$

Due to the symmetry of the PSF function $u_0(-x) = u_0(x)$, this overlap preserves the symmetry and parity of the measurement modes:

$$\beta_m(-x_0) = (-1)^m \beta_m(x_0). \tag{3.51}$$

The overlaps β_m can be viewed as elements of the normalized vector $u_0\left(x - d/2\right)$

in the orthonormal basis $f_m^{(s)}(x)$. The specific choice of basis (as long as it is symmetric) often does not affect the final results of calculations if the basis is complete

$$\sum_{m=1}^{K} f_m^{(s)}(a) f_m^{(s)}(b) \xrightarrow[K\to\infty]{} \delta(a-b). \tag{3.52}$$

Using the completeness property, we present some expressions that remain invariant across measurement bases and will be used throughout the subsequent sections:

$$A = \sum_{m=0}^{K} \beta_m^2 \xrightarrow[K\to\infty]{} 1 \tag{3.53}$$

$$A' = \frac{\partial A}{\partial d} = 2\sum_{m=0}^{K} \beta_m \frac{\partial \beta_m}{\partial d} \xrightarrow[K\to\infty]{} 0 \tag{3.54}$$

$$\delta = \sum_{m=0}^{K} \beta_m^2 (-1)^m \xrightarrow[K\to\infty]{} \int u_0\left(x-\frac{d}{2}\right) u_0\left(x+\frac{d}{2}\right) dx, \tag{3.55}$$

$$\delta' = \frac{\partial \delta}{\partial d} = 2\sum_{m=0}^{K} (-1)^m \beta_m \frac{\partial \beta_m}{\partial d}, \tag{3.56}$$

$$\Delta k^2 = 4\sum_{m=0}^{K} \left(\frac{\partial \beta_m}{\partial d}\right)^2 \xrightarrow[K\to\infty]{} \int \left(\frac{\partial u_0(x)}{\partial x}\right)^2 dx, \tag{3.57}$$

$$\beta = -4\sum_{m=0}^{K} \left(\frac{\partial \beta_m}{\partial d}\right)^2 (-1)^m \xrightarrow[K\to\infty]{} \int \frac{\partial u_0(x-d/2)}{\partial x} \frac{\partial u_0(x+d/2)}{\partial x} dx. \tag{3.58}$$

Here, A describes the portion of the image intensity captured in the given measurement basis (in the full basis, all the light is being measured), δ stands for the overlap of the two images, Δk^2 describes the norm of the PSF derivative, and β represents the overlap between the derivatives of the two image modes.

These characteristics can be explicitly computed for the Gaussian PSF (3.26), yielding:

$$\delta = \exp\left[-\frac{d^2}{8\sigma^2}\right], \quad \Delta k^2 = \frac{1}{4\sigma^2}, \tag{3.59}$$

$$\beta = \frac{\sigma^2 - (d/2)^2}{4\sigma^4} \exp\left[-\frac{d^2}{8\sigma^2}\right]. \tag{3.60}$$

Most of our subsequent results are independent of the measurement basis. However, in certain instances where numerical computations with a limited number of modes are conducted, we opt for the HG mode basis.

It was shown in [Rehacek 2017], that for the Gaussian PSF (3.26) the optimal choice of modes are the Hermite-Gaussian (HG) modes with width matched to the width of the PSF. The mode profile of HG modes is given by

$$f_m^{HG}(x) = \frac{1}{\sqrt{2^m m!}} H_m\left(\frac{x}{\sqrt{2}\,\sigma}\right) u_0(x), \tag{3.61}$$

where H_n are Hermite polynomials. This measurement basis offers localization of information in the first modes, in case of resolving closely located incoherent sources. Achieving equivalent sensitivity in alternative symmetric mode bases requires measuring a greater number of modes.

The overlap (3.50) of the image mode $u_0(x - x_0)$ with the HG mode $f_m^{HG}(x)$ is given by

$$\beta_m^{HG}(x_0) = \int dx\, u_0(x - x_0) f_m^{HG}(x) = \exp\left[-\frac{1}{2}\left(\frac{x_0}{2\sigma}\right)^2\right] \frac{1}{\sqrt{m!}} \left(\frac{x_0}{2\sigma}\right)^m. \tag{3.62}$$

3.2.4 Background of the SPADE technique

Leveraging the SPADE approach for resolving incoherent sources was suggested in [Tsang 2016]. In this seminal paper, it was demonstrated that the QFI, associated with the estimation of the separation between two faint incoherent sources, in fact, does not depend on the separation itself (see fig. 3.9). At the same time, the FI of direct imaging (DI) vanishes for estimating small separations, which was nicknamed the "Rayleigh's curse". This implies that DI is not an optimal measurement as it does not extract the full information available in the field and has particular difficulties in resolving closely separated sources. An optimal measurement was suggested in the form of intensity measurement in the HG modes, mode-matched to the PSF width and aligned to the centroid between the sources. A simplified SPADE approach called *binary SPADE* was suggested in the same paper [Tsang 2016]. Within this approach, the mode HG_{00} is extracted from the image, while all the remaining light is being detected with a bucket detector. This approach is optimal for resolving closely separated sources $d \ll \sigma$, but provides less information in case of large separation.

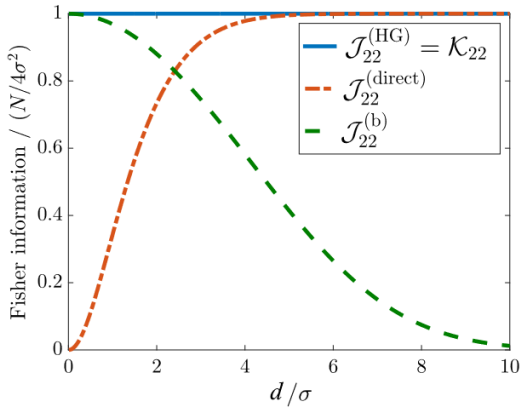

FIGURE 3.9: Comparison of the QFI (blue line), FI of the SPADE in full HG basis (coincide with QFI), FI of binary SPADE in HG mode basis (green dashed line), and FI of the direct imaging (red dot-dashed line). Adapted from [Tsang 2016].

The SPADE approach got a lot of attention and was intensively studied in various generalizations [Tsang 2019b]. Several techniques, somewhat analogous to binary

SPADE were suggested in the literature. SPLICE (super-resolved position localization by inversion of coherence along an edge) manages to extract the first order mode via a π phase shift done by a phase plate in the image plane [Nair 2016; Tham 2017] and SLIVER (super localization via image-inversion interferometry) does the same with an image-inversion interferometer [Bonsma-Fisher 2019; Sajjad 2024]. These techniques approach the quantum limit in perfect conditions for separation estimation in the sub-Rayleigh regime. Similar approaches, based on PSF shaping, also proved to be perspective direction [Paúr 2018; Paúr 2019].

The resolving of intense thermal sources was examined by studying the QFI [Lupo 2016; Nair 2016; Sorelli 2022] and the moment-based sensitivity of the complete SPADE measurement [Nair 2016; Sorelli 2021a]. Directly calculating the FI in this scenario is a challenging task, but one can calculate the sensitivity with the method of moments (MoM). If MoM sensitivity saturates the QFI, it immediately provides the FI matrix through the chain inequality (2.104). These studies showed that the information per photon decreases when resolving intense thermal sources, particularly if their separation is approximately equal to the width σ of the PSF (see Figure 3.10).

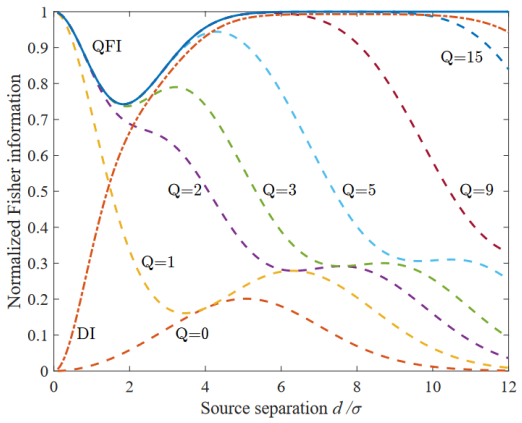

FIGURE 3.10: Comparison of the QFI (blue line), DI (dot-dashed red line), and lower bound of the FI of the SPADE in Q-mode HG basis for resolving thermal sources with $N_S = 1.5$ photons per coherence time. Reprinted from [Nair 2016].

Various imperfections of the SPADE measurement have been considered, including detection noise [Len 2020; Lupo 2020a; Oh 2021], demultiplexing crosstalk [Gessner 2020a; Linowski 2023], measurement basis misalignment from the centroid [Tsang 2016; Almeida 2021], and a combination of all these factors [Sorelli 2021a; Sorelli 2021b]. These studies have shown the resurgence of the Rayleigh curse in the presence of measurement imperfections, although the estimation of large separations was not significantly affected. It has been later demonstrated that resolving large separations is not susceptible to small measurement imperfections of any kind, whereas resolving closely separated sources can be easily corrupted by imperfect measurements [Kurdziałek 2023].

Homodyne and heterodyne detection in HG modes, instead of intensity detection, have been extensively studied in several publications [Yang 2017; Datta 2020;

Datta 2021; Costa Filho 2021; Gosalia 2023]. This method was found to be not particularly efficient for passive imaging, especially with faint sources, as there is no phase reference to the measured field. However, it may prove useful for active coherent imaging where a phase reference is available.

One of the crucial assumptions underlying SPADE models is that the measurement basis aligns with the centroid of the sources. Therefore, if the centroid is unknown, additional preliminary measurements are necessary to determine it. This measurement can be conducted independently of the separation estimation only for a pair of incoherent equally bright sources. Alternatively to a preliminary centroid estimation, a multi-stage algorithm can be employed: alternating SPADE (for estimating the separation) with DI (to refine the centroid position) [Grace 2020]. Another approach involves simultaneously estimating the centroid and separation using Hong-Ou-Mandel interference [Parniak 2018]. The quantum bounds for simultaneous estimation of both sources' positions were studied in the following papers [Ang 2017; Rehacek 2017].

Estimating axial displacements of sources [Koutný 2021; Zhou 2019b] and simultaneously positioning multiple sources in all three dimensions [Tsang 2015; Backlund 2018; Yu 2018; Prasad 2020; Napoli 2019] are other intriguing problems that have gained significant attention in recent years.

The point source resolving problem has also been explored beyond the equally bright sources approximation through multiparameter QFI analysis [Řehaček 2017; Řeháček 2018] and by calculating the FI of the SPADE measurement [Costa Filho 2021; Linowski 2023; Santamaria 2024; Li 2024]. It has been shown that in a somewhat artificial single-parameter setting (where only the separation is estimated, assuming all other parameters to be known), the power imbalance of the sources does not affect the sensitivity of the SPADE measurement but does improve the resolution of DI, eliminating the Rayleigh curse. However, in a more realistic scenario with multiparameter estimation, the Rayleigh curse is present even at the quantum bound, indicating that no measurement with finite statistics can resolve arbitrarily small separations in the multiparameter case. Nonetheless, the SPADE approach still demonstrates considerably higher sensitivity over DI even in this case.

Imaging of multiple incoherent sources was also studied extensively in multiple publications [Bisketzi 2019; Lupo 2020b; Fiderer 2021; Bao 2021; Costa Filho 2021; Liao 2022], bringing the problem closer to realistic imaging setting [Pushkina 2021; Costa Filho 2021]. It has been shown that regardless of the number of sources, the QFI matrix of their one-dimensional positioning retains only two nonzero eigenvalues as the source separations approach zero. This means that one can extract at most two independent combinations of the positions of sources along one axis if the sources are strongly grouped. This finding aligns with existing results on moment estimation [Zhou 2019a; Tsang 2019c; Tsang 2019a; Tsang 2020; Tsang 2021], highlighting the harsh quantum limits to imaging beyond centroid and size estimation.

The SPADE approach has also been investigated in the context of hypothesis testing for "one-versus-two" source determination [Lu 2018; Huang 2021; Huang 2023; Deshler 2024; Wadood 2024; Schlichtholz 2024] and pattern recognition [Grace 2022; Ortolano 2023], demonstrating a clear advantage over approaches base on the DI [Acuna 1997; Shahram 2006] and approaching the quantum bound [Helstrom 1973].

Moreover, it has been combined with other statistical methods such as the Bayesian approach [Tsang 2018; Lee 2022] or biased estimation [Tsang 2016], showcasing precision beyond the CRB.

The straightforward generalization of the SPADE technique to imaging in the time-frequency domain represents another field of active theoretical and experimental studies [Donohue 2018; Ansari 2021; Kumar 2021; De 2021], where unprecedented temporal resolution is demonstrated.

In addition to theoretical studies, numerous experiments have been conducted using SPADE and analogous techniques to demonstrate source resolution beyond the Rayleigh limit [Paúr 2016; Tang 2016; Yang 2017; Tham 2017; Parniak 2018; Zhou 2019a; Boucher 2020; Pushkina 2021; Zanforlin 2022; Greenwood 2023; Santamaria 2023; Tan 2023a; Rouvière 2024; Santamaria 2024]. While a convincing demonstration of the practical utility of this approach is still pending, we firmly believe it is a matter of the near future.

One of the significant recent debates on point source resolving concerns partially coherent sources [Larson 2018; Tsang 2019d; Larson 2019; Hradil 2019; Tsang 2019b; Hradil 2021; Liang 2021; Kurdzialek 2022; Sorelli 2022; Liang 2023b; Wang 2023]. Different studies have reached different conclusions regarding the value of the QFI in this scenario. This discrepancy stems from different representations of the light's state in the image plane. Traditionally, when dealing with incoherent sources within the Poissonian limit, the state of light is often depicted in the single-photon subspace. This approach describes detection statistics by focusing solely on successful detection events, omitting outcomes where no photons are detected. However, due to the interference of partially coherent sources, the overall photon detection frequency may depend on estimated parameters (such as separation between the sources). Hence, it is crucial to incorporate detection probability into the model and accurately account for losses in the imaging system. In other words, the description should include the vacuum component, as discussed in detail in [Tsang 2019d; Kurdzialek 2022; Liang 2023b].

This debate on the effects of coherence has inspired the research presented in this chapter. Despite the scientific community largely converging on a common understanding in this discussion, the practical implications of coherence on point source resolving have not been comprehensively studied to date. In this research, we broaden the scope of this problem beyond the simple Poissonian limit. In section 3.3, we investigate bright mutually coherent sources with arbitrary statistics, ranging from entangled sub-Poissonian to thermally correlated super-Poissonian sources. Additionally, in section 3.4, we explore sources with partial mutual coherence, including cases of separation-dependent mutual coherence and brightness of the sources. The latter case is particularly relevant for resolving interactive emitters and reflective particles under external illumination. We demonstrate that separation-dependent coherence, which emerges in this scenario, can significantly enhance optical resolution.

3.3 Fully coherent sources

First, we study the resolving of a pair of mutually coherent sources, i.e. the case of $|\gamma| = \gamma_0 = 1$. Most of the results in this section, except for section 3.3.8, were presented in [Karuseichyk 2022] in slightly different notations.

3.3.1 Model of the sources

For the mutually coherent sources, the coherency matrix (3.29) is degenerate and its rank is equal to 1. Thus, we can find the principle mode $\hat{s}_0$, defined by the first eigenvector of the coherency matrix, that fully describes the light emitted from this pair of sources. The mode basis $\hat{s}_{1,2}$ is connected with the mode $\hat{s}_0$ with unitary mode transformation. Omitting the global phase, the most general 2×2 unitary transformation can be represented as

$$\begin{pmatrix} \hat{s}_1 \\ \hat{s}_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \hat{s}_0 \\ \hat{v}_0 \end{pmatrix}, \tag{3.63}$$

where $\hat{v}_0$ is the field operator of a vacuum mode. The first matrix sets the mutual phase between the two modes and can be represented as a phase-shifting element ϕ in one of the modes (see fig. 3.11). The second matrix sets the imbalance between the sources $\cos\theta = \sqrt{p}$, and it can be physically represented as a beamsplitter (1.62) (which does not introduce an additional phase to the modes). For concreteness, we assume the source $\hat{s}_1$ to be brighter, i.e. the parameter θ to be within a range $0 \leq \theta \leq \pi/4$, where $\theta = \pi/4$ corresponds to equally bright sources ($p = 1/2$), and $\theta = 0$ to all light in mode $\hat{s}_1$ ($p = 1$).

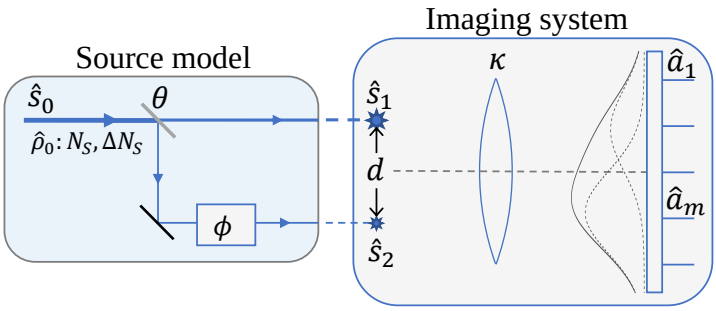

FIGURE 3.11: On the left: conceptual scheme for generating a general two mutually coherent modes based on a beam splitter with transmissivity $T = \cos^2\theta = p$ and a phase shifting element ϕ. On the right: optical scheme for the estimation of the sources' separation.

Since the problem is, in essence, single mode, eq. (3.40) for normally ordered averages can be further simplified using the eq. (3.63) to

$$\langle g_N(\hat{a}_m^\dagger, \hat{a}_n)\rangle = \langle g_N(A_m^* \hat{s}_0^\dagger, A_n \hat{s}_0)\rangle, \tag{3.64}$$

with A_m being an overlap of the principal image mode, associated with $\hat{s}_0$, and an arbitrary real measurement mode $f_m(\vec{r})$:

$$A_m = \sqrt{\kappa} \int d\vec{r}\, f_m(\vec{r}) \left[u_0(\vec{r} - \vec{r}_1) \cos\theta + u_0(\vec{r} - \vec{r}_2) e^{i\phi} \sin\theta \right] \\ = \sqrt{\kappa} \left(\beta_m(\vec{r}_1) \cos\theta + \beta_m(\vec{r}_2) e^{i\phi} \sin\theta \right). \quad (3.65)$$

Since we are interested in intensity measurements in the modes $\hat{a}_m$, we can use the property (3.64) to calculate the mean number of photons in the measurement modes

$$N_m = \langle \hat{N}_m \rangle = \langle \hat{a}_m^\dagger \hat{a}_m \rangle = |A_m|^2 \langle \hat{s}_0^\dagger \hat{s}_0 \rangle = |A_m|^2 N_S. \quad (3.66)$$

To calculate the covariance matrix of the measurements $\hat{N}_m$ we transform it to the normal order

$$\Gamma_{mn} = \langle \hat{N}_m \hat{N}_n \rangle - \langle \hat{N}_m \rangle \langle \hat{N}_n \rangle = \langle \hat{a}_m^\dagger \hat{a}_m \hat{a}_n^\dagger \hat{a}_n \rangle - N_m N_n = \langle \hat{a}_m^\dagger \hat{a}_n^\dagger \hat{a}_m \hat{a}_n \rangle + \delta_{mn} N_m - N_m N_n \quad (3.67)$$

and again use the property (3.64)

$$\langle \hat{a}_m^\dagger \hat{a}_n^\dagger \hat{a}_m \hat{a}_n \rangle = |A_m|^2 |A_n|^2 \langle \hat{s}_0^\dagger \hat{s}_0^\dagger \hat{s}_0 \hat{s}_0 \rangle = |A_m|^2 |A_n|^2 (\langle \hat{s}_0^\dagger \hat{s}_0 \hat{s}_0^\dagger \hat{s}_0 \rangle - N_S) \\ = \frac{N_m}{N_S} \frac{N_n}{N_S} (\Delta^2 N_S + N_S^2 - N_S), \quad (3.68)$$

where

$$\Delta^2 N_S = \langle \hat{s}_0^\dagger \hat{s}_0 \hat{s}_0^\dagger \hat{s}_0 \rangle - \langle \hat{s}_0^\dagger \hat{s}_0 \rangle \langle \hat{s}_0^\dagger \hat{s}_0 \rangle \quad (3.69)$$

is the photon number variance in the principal mode $\hat{s}_0$.

We try to represent further expressions in terms of mean photon numbers N_m, as these quantities are directly measured in experiments and possess a clear physical sense. At this point, we do not use any explicit properties of coefficients A_m (3.65), and only exploit the linearity of the imaging system. Thus, as we will discuss later, our conclusions can be generalized to a wide class of problems.

Combining eqs. (3.67) and (3.68) we find the elements of the photon number covariance matrix

$$\Gamma_{mn} = \delta_{mn} N_m + h N_m N_n, \quad (3.70)$$

where the parameter $h = (\Delta N_S^2 - N_S)/N_S^2 = g^{(2)} - 1$ with $g^{(2)}$ the degree of second-order coherence of the principal mode $\hat{s}_0$ [Scully 1997]. For the Poisson statistics of the source mode $\Delta N_S^2 = N_S$ and $h = 0$ ($g^{(2)} = 1$), for sub-Poissonian statistics $\Delta N_S^2 < N_S$ and $h < 0$ (anti-bunched states), and for super-Poissonian $\Delta N_S^2 > N_S$ and $h > 0$ (bunched states). Further, we will consider these examples in more detail.

In eq. (3.70) we considered only the quantum noise of light, assuming that no extra noise comes from the detectors. This allows us to perform most of the calculations analytically, and study the upper bound of the sensitivity for the practical moment-based approach. We also aim to compare the results of MoM sensitivity computation to the QFI, which does not include any measurement imperfections (as it typically does not consider the measurement process at all). In the general case, the presence of the detection noise requires the following modification of the covariance

matrix:

$$\Gamma^{\mathrm{DC}}_{mn} = \Gamma_{mn} + N_{\mathrm{DC}}\delta_{mn}, \tag{3.71}$$

where N_{DC} is the average number of dark counts per coherence time of the sources τ_c. In eq. (3.71), the dark counts are assumed to have Poisson statistics, being uncorrelated from each other and the measured signal. There are no difficulties in reproducing the results of our further analytical computations with numerical methods, taking into account the modification (3.71). As demonstrated in [Sorelli 2021b; Len 2020; Kurdziałek 2023], typically the detection noise influences only the resolving of very closely located sources and causes the appearance of the Rayleigh curse for SPADE measurement. However, since we are mostly focused on the fundamental properties of different imaging approaches, in most cases we avoid going into detailed consideration of technical effects, such as detection noise and imperfections of the SPADE (including misalignment from the centroid).

3.3.2 The moment matrix

To calculate the moment matrix (2.54) of the given measurements for characterization of the estimation of the parameters $\vec{\theta}$, one needs to inverse the measurement covariance matrix (3.70). Due to its specific structure, one can inverse it analytically with the help of the Sherman-Morrison formula [Sherman 1950]:

$$\left(\mathbf{B} + \vec{u}\vec{v}^T\right)^{-1} = \mathbf{B}^{-1} - \frac{\mathbf{B}^{-1}\vec{u}\vec{v}^T\mathbf{B}^{-1}}{1 + \vec{v}^T\mathbf{B}^{-1}\vec{u}}. \tag{3.72}$$

Using $B_{mn} = \delta_{mn}N_m$, $u_m = hN_m$ and $v_n = N_n$ we immediately obtain

$$\left[\mathbf{\Gamma}^{-1}\right]_{mn} = \delta_{mn}N_m^{-1} - \frac{h}{1 + hN_D}, \tag{3.73}$$

where

$$N_D = \sum_m N_m \tag{3.74}$$

is the total number of detected photons. Note that due to the coherence of the sources, this quantity generally depends on the separation d. The moment matrix (2.54) can be expressed as

$$M_{\alpha\beta} = \sum_{mn}\left[\mathbf{\Gamma}^{-1}\right]_{mn}\frac{\partial N_m}{\partial\theta_\alpha}\frac{\partial N_m}{\partial\theta_\beta} = \sum_m \frac{1}{N_m}\frac{\partial N_m}{\partial\theta_\alpha}\frac{\partial N_m}{\partial\theta_\beta} - \frac{h}{1+hN_D}\sum_{mn}\frac{\partial N_m}{\partial\theta_\alpha}\frac{\partial N_m}{\partial\theta_\beta}. \tag{3.75}$$

This expression takes a more physical meaning by introducing the notion of the *relative measured photon numbers*

$$\varepsilon_m = \frac{N_m}{N_D} = \frac{|A_m|^2}{\sum_m |A_m|^2}. \tag{3.76}$$

These quantities do not depend on the state $\hat{\rho}_0$ of the principle source mode $\hat{s}_0$. By substituting $N_m = N_D\,\varepsilon_m$ to eq. (3.75) we arrive at

$$\boxed{M_{\alpha\beta} = N_D \sum_m \frac{1}{\varepsilon_m}\frac{\partial \varepsilon_m}{\partial \theta_\alpha}\frac{\partial \varepsilon_m}{\partial \theta_\beta} + \frac{1}{\Delta N_D^2}\frac{\partial N_D}{\partial \theta_\alpha}\frac{\partial N_D}{\partial \theta_\beta},} \tag{3.77}$$

with

$$\Delta N_D^2 = \sum_{mn} \Gamma_{mn} = N_D(1 + hN_D) \tag{3.78}$$

being the variance of the total number of detected photons. To understand the meaning of the two terms in this expression let us consider two specific examples.

3.3.3 Single-parameter estimation

If all parameters except for the separation d between the sources are known, i.e. $\vec{\theta} = (d)$, the moment matrix reduces to a single number

$$\boxed{M(d) = N_D \sum_m \frac{1}{\varepsilon_m}\left(\frac{\partial \varepsilon_m}{\partial d}\right)^2 + \frac{1}{\Delta N_D^2}\left(\frac{\partial N_D}{\partial d}\right)^2.} \tag{3.79}$$

The sensitivity S_d (2.56) of the separation estimation in this case coincides with $M(d)$:

$$S_d(d) = \frac{1}{[\boldsymbol{M}^{-1}(d)]_{11}} = M(d) = N_D M_\varepsilon + M_D. \tag{3.80}$$

This expression has two terms, the first term equals $N_D M_\varepsilon$, where

$$M_\varepsilon = \sum_m \frac{1}{\varepsilon_m}\left(\frac{\partial \varepsilon_m}{\partial d}\right)^2 \tag{3.81}$$

does not depend on the state $\hat{\rho}_0$ of the principle mode $\hat{s}_0$ (but still depends on the state of the sources which is also determined by θ and ϕ) but strongly depends on the measurement basis $\{f_m(\vec{r})\}$. M_ε is the sensitivity per detected photon of relative intensity measurements.

The second term

$$M_D = \frac{1}{\Delta N_D^2}\left(\frac{\partial N_D}{\partial d}\right)^2 \tag{3.82}$$

does not depend on the individual signals N_m, but only on the total number of detected photons N_D, thus it stays the same for any measurement basis $\{f_m(\vec{r})\}$ as long as all photons in the image plane are detected. This additional sensitivity M_D occurs due to the interference of mutually coherent sources and the subsequent dependence of the total number of registered photons N_D on the separation d. The variance of the total number of detected photons (3.78) depends on the quantum statistics of the source and grows with source bunching.

Expression (3.82) has a self-consistent structure representing the simple error-propagation formula. Below we show that the sensitivity M_D, coming from the total number of detected photons, vanishes if the sources' brightness N_S is unknown.

3.3.4 Resolving sources with unknown brightness

Now let us consider the two-parameter problem, where both the separation d and the emitted number of photons N_S are unknown and treated as parameters to be estimated. The moment matrix (3.77) for the estimation of the parameters $\vec{\theta}^T = (d, N_S)$ reads

$$\boldsymbol{M}(d, N_S) = \begin{pmatrix} N_D M_\varepsilon + M_D & \frac{1}{\Delta N_D^2}\frac{N_D}{N_S}\frac{\partial N_D}{\partial d} \\ \frac{1}{\Delta N_D^2}\frac{N_D}{N_S}\frac{\partial N_D}{\partial d} & \frac{1}{\Delta N_D^2}\left(\frac{N_D}{N_S}\right)^2 \end{pmatrix}, \tag{3.83}$$

where the linearity of eq. (3.66) is taken into account, which yields $\partial N_D/\partial N_S = N_D/N_S$ and $\partial \varepsilon_m/\partial N_S = 0$. In this case the sensitivity to separation d can be found by inverting the matrix (3.83) yielding:

$$\boxed{S_d(d, N_S) = \frac{1}{[\boldsymbol{M}^{-1}(d, N_S)]_{11}} = N_D M_\varepsilon.} \tag{3.84}$$

This formula shows that without knowing the number of emitted photons N_S one can benefit only from measurements of relative intensities ε_m. Comparing it to eq. (3.80) one can see that knowing the number of emitted photons N_S increases the sensitivity of separation estimation by the sensitivity M_D obtained from the total detected number of photons. This fact fits well with the conclusions of the discussion about the effect of the partial coherence of thermal sources on the resolution [Tsang 2019d; Kurdzialek 2022]. From eq. (3.84) one can conclude that ignorance of the sources' brightness N_S wipes out any possible advantage from non-classical statistics of the sources that are only present in term M_D.

Now, let us consider the bucket detection that corresponds to the detection of all the photons in the image plane. In this particular case, it can be described as a single-mode measurement in the principal mode. Since only one observable is measured, the only relative intensity $\epsilon_0 = 1$ does not depend on the parameter, thus $M_\epsilon = 0$. One can not estimate separation from this measurement without knowing the number of emitted photons N_S. If N_S is known then, as expected, full sensitivity of bucket detection is provided by total photon number detection $M(d) = M_D$.

Note that even though the structure of the sensitivity expression (3.77) is quite intuitive, it was derived specifically for the single mode (fully coherent) case and does not necessarily hold for other cases.

It is worth mentioning that the used property (3.64) is valid for a quite general class of parameter estimation schemes, where the parameters are encoded in an arbitrary number of mutually coherent modes, that are subjected to correlated parameter-dependent linear losses. Therefore, eq. (3.77) for the moment matrix of photon counting is valid for this wider class of systems, since the explicit form of coefficients A_m (3.65) was never used so far. Thus the developed approach can be used for other problems like coherent imaging [Ferraro 2011] or quantum sensing in a continuous variable entangled network in the single-mode regime [Guo 2020].

Moreover, in certain cases, the developed approach can be directly applied even beyond the fully coherent scenario. For instance, it was subsequently used by [Zhang 2023a] to examine the resolving of entangled point sources, with entanglement generated by optical parametric amplification with arbitrary statistics of the pump. In this instance, the field operators of the sources are given by [Agarwal 2013]:

$$\begin{pmatrix} \hat{s}_1 \\ \hat{s}_2^\dagger \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{pmatrix} \begin{pmatrix} \cosh\theta & -\sinh\theta \\ -\sinh\theta & \cosh\theta \end{pmatrix} \begin{pmatrix} \hat{s}_0 \\ \hat{v}_0^\dagger \end{pmatrix}, \tag{3.85}$$

which allows the use of the results of this section with the corresponding modification of coefficients A_m (3.65).

3.3.5 Sensitivity of relative intensity measurements M_ε

Now let us consider separately the two parts of the separation estimation sensitivity (3.80). The sensitivity of the relative intensity measurement M_ε does not depend on the quantum state $\hat{\rho}_0$, but strongly depends on the measurement basis $\{f_m(\vec{r})\}$. As we already mentioned, it equals zero for bucket detection in the image plane. Here, we consider two other measurement bases.

Direct imaging

We start by considering the traditional approach, i.e. direct imaging (DI). Assuming the real-valued PSF function $u_0(\vec{r}) \in \mathbb{R}$, and taking into account $\sqrt{p} = \cos\theta$ and $\gamma = e^{i\phi}$ we rewrite the intensity distribution (3.47) as

$$I(\vec{r}) = \kappa N_S \big[u_0^2(\vec{r}-\vec{r}_1)\cos^2\theta + u_0^2(\vec{r}-\vec{r}_2)\sin^2\theta + u_0(\vec{r}-\vec{r}_1)u_0(\vec{r}-\vec{r}_2)\sin 2\theta\cos\phi\big]. \tag{3.86}$$

Using the continuous limit (3.48), we can find the total number of detected photons:

$$N_D = \int I(\vec{r})d\vec{r} = \kappa N_S(1+\chi\delta), \tag{3.87}$$

where we introduced the parameter

$$\chi = \sin 2\theta\cos\phi, \tag{3.88}$$

which represents the amplitude of the interference term in eq. (3.86). The overlap δ is defined in eq. (3.55). Variable $N_D d$ (3.87) defines the total number of photons, arriving at the image plane (it is independent of the measurement basis).

Again using the continuous limit (3.48) the sensitivity M_ε (3.81) of a relative intensity measurement with DI can be calculated as

$$M_\varepsilon^{DI} = \int \frac{1}{i(\vec{r})}\left(\frac{\partial i(\vec{r})}{\partial d}\right)^2 d\vec{r}, \tag{3.89}$$

with $i(\vec{r}) = I(\vec{r})/N_D$. This integral can be calculated analytically in the cases of in-phase ($\phi = 0$) and anti-phase ($\phi = \pi$) sources, giving

$$M_\varepsilon^{DI}\Big|_{\phi=0,\pi} = \frac{1}{1+\chi\delta}\left(\Delta k^2 - \chi\beta - \frac{(\chi\delta')^2}{1+\chi\delta}\right), \tag{3.90}$$

where coefficients δ, δ', Δk^2 and β are defined in eqs. (3.55) to (3.58). Another example, where M_ε^{DI} can be calculated analytically is the case of fully asymmetric sources $\theta = 0$ (i.e. $p = 1$, which is equivalent to a single source positioning) giving the well-known result [Tsang 2016; Chao 2016]

$$M_\varepsilon^{DI}\Big|_{\theta=0} = \Delta k^2. \tag{3.91}$$

For other values of ϕ and θ the sensitivity M_ε^{DI} is calculated numerically.

The variable M_ϵ corresponds to the sensitivity per detected photon. The sensitivity per emitted photon is instead given by $N_D M_\epsilon / N_S$. We perform numerical computations for a Gaussian PSF (3.26) and additionally normalize the sensitivity value over transmissivity κ and multiply by $4\sigma^2$ to remove dependence on these parameters. A plot of the resulting normalized sensitivity

$$\mathcal{M}_\varepsilon = \frac{4\sigma^2}{\kappa}\frac{N_D M_\varepsilon}{N_S} \tag{3.92}$$

for the case of DI is presented in fig. 3.12.

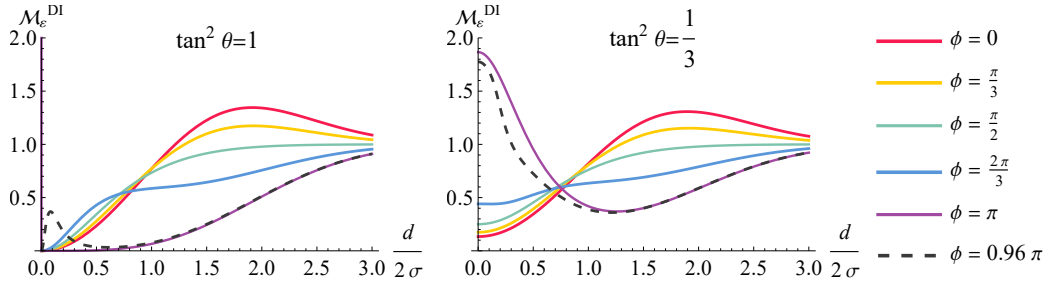

FIGURE 3.12: Normalized sensitivity per emitted photon of relative intensity distribution direct measurement $\mathcal{M}_\varepsilon^{DI}$. The left panel corresponds to the sources with equal intensity and parameter $\chi = \{1, 1/2, 0, -1/2, -1, -0.99\}$, the right panel shows the asymmetric case, and $\chi = \{0.87, 0.43, 0, -0.43, -0.87, -0.86\}$.

On the left panel of fig. 3.12 one can see that direct measurement of the relative intensity distribution in the case of equally bright sources leads to low sensitivity for small separations and, as expected, to "Rayleigh's curse", i.e. vanishingly small sensitivity for infinitesimally small d.

In the case of asymmetric sources the first moment of the intensity distribution, i.e. the center of mass (or barycenter) of the image, does not coincide with the geometrical center between sources. Assuming the position of the geometrical center

between the sources and both sources' brightness to be known, one can recover the separation between the sources from the center of mass of the image, which can be accurately measured with direct imaging. Accordingly, Rayleigh's curse does not occur for DI of asymmetric sources (right of fig. 3.12), although the sensitivity is relatively small for small separations unless the sources are nearly in anti-phase configuration.

Spatial-mode demultiplexing

Now we analyze the sensitivity of the SPADE approach with a symmetric measurement basis $f_m^{(s)}(x)$, which has a parity (3.49). Calculating the overlaps with the image modes we find that the coefficients A_m (3.65) are given by

$$A_m = \sqrt{\kappa}\Big(\cos\theta + (-1)^m e^{i\phi}\sin\theta\Big)\,\beta_m\left(\frac{d}{2}\right). \tag{3.93}$$

This allows us to find the mean photon numbers in the measurement modes (3.66)

$$N_m = \kappa N_S\Big(1 + (-1)^m\chi\Big)\,\beta_m^2, \tag{3.94}$$

where we omit the argument of the overlaps $\beta_m(d/2)$. The measured number of photons in the mode m can be normalized with respect to the total number of detected photons N_D (3.87) to calculate the sensitivity M_ε (3.81) of the separation estimation from the measured relative photon numbers

$$\varepsilon_m = \frac{1 + (-1)^m\chi}{1+\chi\delta}\beta_m^2. \tag{3.95}$$

The derivatives of these values are given by

$$\frac{\partial\varepsilon_m}{\partial d} = \varepsilon_m\left(2\frac{\beta_m'}{\beta_m} - \frac{\chi\delta'}{1+\chi\delta}\right). \tag{3.96}$$

Then, the sensitivity M_ε (3.81) reads

$$M_\varepsilon = \frac{1}{A+\chi\delta}\left(\Delta k^2 - \chi\beta - \frac{(\chi\delta')^2}{A+\chi\delta}\right). \tag{3.97}$$

We do not use any additional indices, further assuming that M_ε stands for the sensitivity SPADE, while M_ε^{DI} corresponds to DI. Note that in this case M_ε depends only on the combination χ of the parameters θ and ϕ.

In fig. 3.13 we plot the normalized sensitivity (3.92) of the relative intensity measurement in the full symmetric basis, calculated for a Gaussian model of the PSF. One can see that Rayleigh's curse is still present for the symmetric in-phase ($\chi = 1$) and anti-phase ($\chi = -1$) cases, and SPADE for these cases is as sensitive as DI. Note, however, that even a small deviation from the symmetric anti-phase case ($\chi = -1$) leads to a significant sensitivity increase for small separations (see dashed line in fig. 3.13).

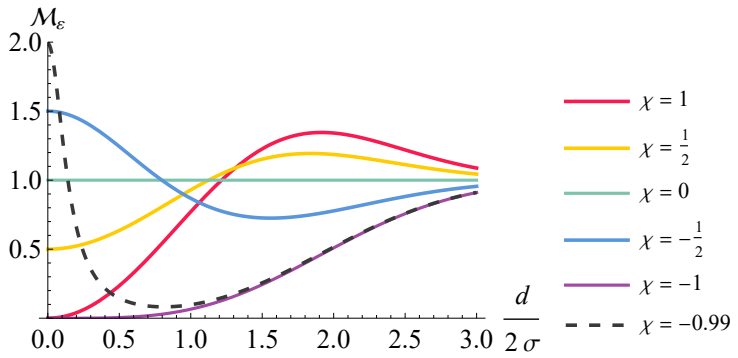

FIGURE 3.13: Normalized sensitivity per emitted photon of relative intensity measurements with SPADE $\mathcal{M}_\varepsilon$.

The sensitivity (3.97) in the case $\chi = 0$ (that corresponds to the mutual phase $\phi = \pi/2$) does not depend on the separation d. It coincides with the sensitivity in the case of weak uncorrelated thermal sources [Sorelli 2021b], which in turn coincides with QFI [Lupo 2016] in this limit. However, for incoherent thermal sources the QFI per emitted photon drops with a growing number of photons, when $\mathcal{M}_\varepsilon$ for correlated sources does not depend on N_S for any photon statistics of the source. We discuss this effect in more detail in section 3.4.

Comparison of the SPADE and DI sensitivities

We observe that the expression for the sensitivity of SPADE M_ε (3.97) coincides with the sensitivity of DI M_ε^{DI} in cases of in-phase, anti-phase (3.90), and fully asymmetric sources (3.91). Here we demonstrate that in the general case

$$M_\varepsilon \geq M_\varepsilon^{DI}, \tag{3.98}$$

i.e. separation estimation sensitivity with SPADE measurements outperforms DI for any pair of mutually coherent sources. The inequality (3.99) is saturated only in cases $\phi = 0$, $\phi = \pi$ or $\theta = 0$.

Proof. To proof the inequality (3.98) we analyze the following difference

$$\Delta M = M_\varepsilon - M_\varepsilon^{DI} = M(d) - M^{DI}(d), \tag{3.99}$$

with the objective to proof that $\Delta M \geq 0$. We use the full sensitivity $M(d)$ in the form given by eq. (3.75):

$$M(d) = \sum_m \frac{1}{N_m} \left(\frac{\partial N_m}{\partial d}\right)^2 - \frac{h}{1 + hN_D} \left(\frac{\partial N_D}{\partial d}\right)^2. \tag{3.100}$$

The second term in the expression is independent of the measurement basis. Then

$$\Delta M = M_0(\theta, \phi) - M_0^{DI}(\theta, \phi), \tag{3.101}$$

where

$$M_0(\theta,\phi) = \sum_m \frac{1}{N_m}\left(\frac{\partial N_m}{\partial d}\right)^2. \tag{3.102}$$

From the equality $M_\varepsilon^{DI}\big|_{\phi=0} = M_\varepsilon\big|_{\phi=0}$ follows that

$$M_0^{DI}(\theta,0) = M_0(\theta,0), \tag{3.103}$$

since the difference between M_ε and M_0 is basis independent. At the same time, M_0 depends only on the parameter combination $\chi = \sin 2\theta\cos\phi$. Then, for any $\phi \le \pi/2$, one can use the following chain of equalities:

$$M_0(\theta,\phi) = M_0(\theta_1,0) = M_0^{DI}(\theta_1,0), \tag{3.104}$$

where $\sin 2\theta_1 = \sin 2\theta\cos\phi$. One can redo all the following analysis for $\phi \ge \pi/2$ using the fact that $M_\varepsilon^{DI}\big|_{\phi=\pi} = M_\varepsilon\big|_{\phi=\pi}$, therefore results are true for any value of ϕ.

Thus the difference in sensitivity can be expressed as

$$\Delta M = M_0^{DI}(\theta_1,0) - M_0^{DI}(\theta,\phi). \tag{3.105}$$

For continuous DI the formula (3.102) takes the form

$$M_0^{DI}(\theta,\phi) = \int \frac{1}{I_{\theta,\phi}(\vec{r})}\left(\frac{\partial I_{\theta,\phi}(\vec{r})}{\partial d}\right)^2 d\vec{r} \tag{3.106}$$

where $I_{\theta,\phi}(\vec{r}) = |E_{\theta,\phi}(\vec{r})|^2$ and

$$E_{\theta,\phi}(\vec{r}) = \sqrt{\kappa}\left(u_0(\vec{r}-\vec{r}_1)\cos\theta + u_0(\vec{r}-\vec{r}_2)e^{i\phi}\sin\theta\right). \tag{3.107}$$

A simple transformation leads to

$$M_0^{DI}(\theta,\phi) = 2\,\mathrm{Re}\left[\int\left(\frac{E^*_{\theta,\phi}(\vec{r})}{E_{\theta,\phi}(\vec{r})}(E'_{\theta,\phi}(\vec{r}))^2 + \left|E'_{\theta,\phi}(\vec{r})\right|^2\right)d\vec{r}\right], \tag{3.108}$$

where $E'_{\theta,\phi}(\vec{r})$ stands for the derivative with respect to the separation d. Using the inequality

$$\mathrm{Re}\left[\int g(x)dx\right] \le \int |g(x)|dx, \tag{3.109}$$

for the first term in (3.108) we find that

$$M_0^{DI}(\theta,\phi) \le 4\int \left|E'_{\theta,\phi}(\vec{r})\right|^2 d\vec{r}. \tag{3.110}$$

Since $E_{\theta_1,0}(\vec{r}) \in \mathbb{R}$, then (3.108) in this special case simplifies to

$$M_0^{DI}(\theta_1,0) = 4\int \left|E'_{\theta_1,0}(\vec{r})\right|^2 d\vec{r}. \tag{3.111}$$

By using the explicit expression for the electric field (3.107), we can calculate the

integral from the right part of (3.110). The result depends only on $\chi = \sin 2\theta \cos\phi$. This means that the right parts of (3.110) and (3.111) are equal, resulting in

$$\frac{\Delta M}{4} \geq \int \left|E'_{\theta_1,0}(\vec{r})\right|^2 d\vec{r} - \int \left|E'_{\theta,\phi}(\vec{r})\right|^2 d\vec{r} = 0. \tag{3.112}$$

The inequality (3.109) is only saturated if $g(x) = |g(x)|$, i.e.

$$\frac{E^*_{\theta,\phi}(\vec{r})}{E_{\theta,\phi}(\vec{r})}(E'_{\theta,\phi}(\vec{r}))^2 = \left|\frac{E^*_{\theta,\phi}(\vec{r})}{E_{\theta,\phi}(\vec{r})}(E'_{\theta,\phi}(\vec{r}))^2\right| \tag{3.113}$$

for any $\vec{r}$. This equality holds only in cases $\phi = 0$, $\phi = \pi$ or $\theta = 0$.

On the fig. 3.14 you can find a comparison of the DI and SPADE sensitivities, built with fixed combinations $\chi = \sin 2\theta \cos\phi$ but with a different ratio between parameters ϕ and θ. These plots illustrate all given relations between M_ε and M_ε^{DI}.

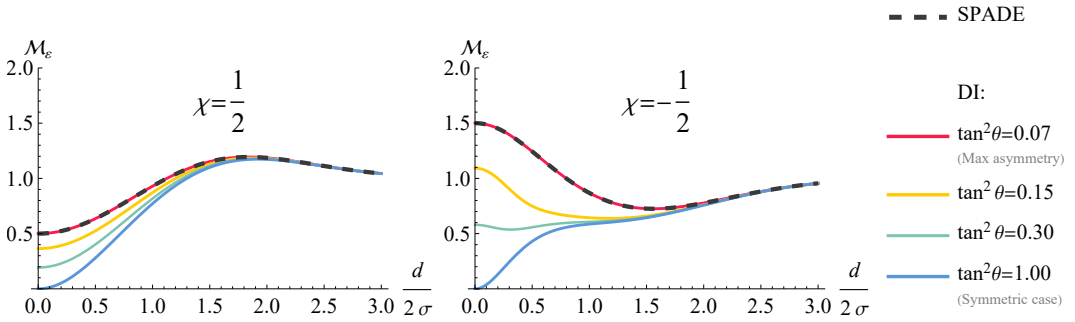

FIGURE 3.14: Sensitivity per emitted photon of relative intensity measurement. Comparison DI and SPADE technique.

3.3.6 Sensitivity of total intensity measurement M_D

Having an expression for the total number of detected photons (3.87) we analytically determine the total photon number sensitivity (3.82)

$$M_D = \kappa N_S \frac{(\chi\delta')^2}{(1+\delta\chi) + h\kappa N_S(1+\delta\chi)^2}. \tag{3.114}$$

Since the total intensity is basis invariant, the sensitivity from measuring it does not depend on the detection basis either. Expression (3.114) also includes all quantum states of the sources (as long as the sources are mutually coherent) via the coefficient $h = g^{(2)} - 1 = (\Delta N_S^2 - N_S)/N_S^2$. From (3.114) it is obvious that anti-bunched states of $\hat{s}_0$ ($h < 0$), leading to entanglement in modes $\hat{s}_{1,2}$, provide a better sensitivity than bunched states ($h > 0$) of $\hat{s}_0$, which corresponds to classical correlations in $\hat{s}_{1,2}$. This is a natural result since a lower photon number variance in $\hat{s}_0$ leads to a smaller variance of N_D and hence a higher sensitivity of the N_D measurement.

Here we consider the sensitivity of separation estimation from a measured total intensity N_D for different quantum statistics of the sources. We are interested in the

normalized sensitivity per emitted photon

$$\mathcal{M}_D = \frac{4\sigma^2}{\kappa}\frac{M_D}{N_S}. \tag{3.115}$$

The characteristics of the source statistics only appear in the combination $h\kappa N_S$. Furthermore, $\mathcal{M}_D$ also depends on χ and the separation d. To explore the impact of the source statistics we study various common initial states.

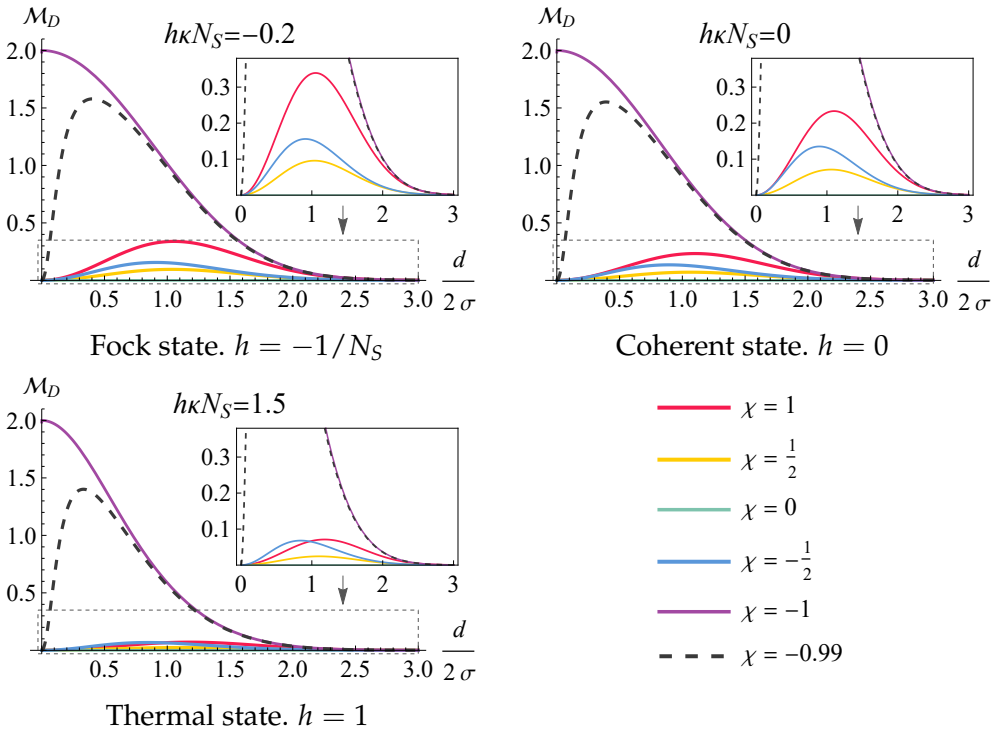

FIGURE 3.15: Normalized sensitivity of total photon number detection $\mathcal{M}_D$ for Fock ($\kappa = 0.2$), coherent and thermal state ($N_S = 1.5/\kappa$) of the mode $\hat{s}_0$.

Fock state. We consider first the most sensitive case, when the mode $\hat{s}_0$ is maximally anti-bunched, i.e. it is in the Fock state, resulting in the entanglement of the modes $\hat{s}_{1,2}$. In this case $h = g^{(2)} - 1 = -1/N_S$, and the combination $h\kappa N_S = -\kappa$. On the first panel of fig. 3.15, we plot the sensitivity $\mathcal{M}_D$ (3.115) with $\kappa = 0.2$ (the used model of linear losses requires $\kappa \ll 1$). Note that, for the Fock state, the sensitivity per emitted photon does not depend on the number of photons N_S.

Coherent state. If the mode $\hat{s}_0$ is in the coherent state, then the states of the modes $\hat{s}_{1,2}$ are uncorrelated. For this case the parameter $h = 0$. Although coherent and Fock states have very different statistical properties, in both cases after propagation through a loss channel, the photon number variance is linear over the initial number of photons, thus the sensitivity per emitted photon does not depend on the source

intensity for both of these cases. The normalized sensitivity for the case of coherent source is plotted on the second panel of fig. 3.15.

Thermal state. Finally, we consider a thermal state of the mode $\hat{s}_0$, that leads to correlated thermal states in modes $\hat{s}_{1,2}$. For thermal statistics $h = 1$. In the small photon number limit ($N_S \to 0$) the sensitivity per emitted photon $\mathcal{M}_D$ coincides with the coherent case, and for high photon number ($N_S \to \infty$) it vanishes ($\mathcal{M}_D \to 0$). We plot the normalized total photon-number sensitivity for correlated thermal sources for $\kappa N_S = 1.5$ on the third panel of fig. 3.15.

For all the considered cases the total photon number sensitivity is high in the case of small separation between symmetric anti-phase sources ($\chi = -1$). This occurs due to destructive interference of mutually coherent anti-phase sources, which leads to zero intensity in the image plane if equally bright sources coincide, and non-zero total intensity in the presence of finite separation between the sources. For any other case, the sensitivity M_D vanishes for zero separation. In the case of $\chi = 0$ the total photon number N_D (3.87) does not depend on the parameter and $M_D = 0$.

3.3.7 Comparison with Quantum Fisher information

Here we analyze the full separation estimation sensitivity $M(d) = N_D M_\varepsilon + M_D$ and its normalized version

$$\mathcal{M} = \frac{4\sigma^2}{\kappa}\frac{M(d)}{N_S} = \mathcal{M}_\varepsilon + \mathcal{M}_D. \tag{3.116}$$

We compare the sensitivity of SPADE, which proved to be always better or equal to that of DI, with the ultimate limit set by the QFI. Note that computing FI in the general case of bright sources is a highly challenging task due to correlations of the detection events in different modes. However, knowing that MoM sensitivity is a lower bound for the FI and that the QFI is its higher bound, we can draw some interesting conclusions by analyzing these bounds.

Fock state. The first example we consider is a Fock state of the mode $\hat{s}_0$. Plots of the SPADE sensitivity for split Fock states are presented on the first panel of fig. 3.16.

Of particular interest are the examples of the Fock state split on a symmetric beam splitter ($\theta = \pi/4$) with added phase $\phi = 0$ or π. Then the states of the sources take the form

$$|\psi\rangle_{s_1 s_2}^{(\pm)} = \frac{1}{\sqrt{2^{N_S}}}\sum_{j=0}^{N_S}\sqrt{\binom{N_S}{j}}(\pm 1)^{N_S-j}|j\rangle_{s_1}|N_S - j\rangle_{s_2}. \tag{3.117}$$

The analytical expression obtained for the SPADE sensitivity $M(d)$ coincides with the QFI for these states [Lupo 2020b]. Note that one of these states corresponds to the maximal QFI of separation estimation (as seen from the fig. 3.16 for small separations state $|\psi\rangle_{s_1 s_2}^{(-)}$ is optimal, for larger $|\psi\rangle_{s_1 s_2}^{(+)}$) [Lupo 2020b]. For other values of the mutual phase ϕ, or asymmetrically split Fock states, the QFI has not been calculated explicitly to the best of our knowledge.

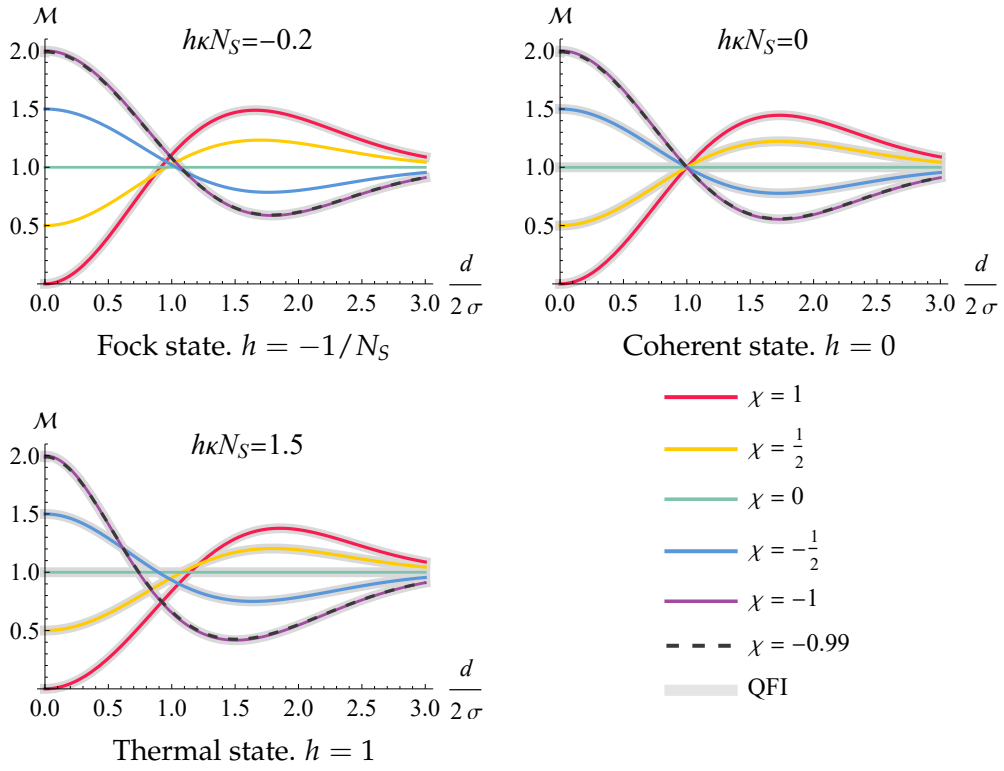

FIGURE 3.16: Normalized full separation estimation sensitivity $\mathcal{M}$ via SPADE in HG basis for Fock ($\kappa = 0.2$), coherent and thermal state ($N_S = 1.5/\kappa$) of the mode $\hat{s}_0$. Cases with known QFI are highlighted with gray, QFI and $M(d)$ coincide for all of them.

Coherent state. Since for Poisson photon number statistics different detection events are independent, we can consider results in the small photon number limit without loosing any generality. The QFI for an arbitrary mutual coherence γ was explicitly calculated in the single-photon subspace [Kurdzialek 2022; Tsang 2019d], and its analytical expression for any $\gamma = e^{i\phi}$ fully coincides with the sensitivity $M(d)$ calculated for Poisson statistics ($h = 0$). The same result was recently obtained for arbitrarily bright coherent sources [Sorelli 2022]. The dependence of the normalized QFI and $\mathcal{M}$ on the separation are presented in the second panel of fig. 3.16, coinciding with each other.

We also plot the sensitivity of DI for symmetric sources with Poisson statistics in fig. 3.17. Comparing this plot to the third panel of fig. 3.16 we clearly see that the choice of symmetric modes (SPADE) results in a significant advantage over DI (comparison of the two plots are presented in fig. 3.18). The special case of symmetric anti-phase sources ($\chi = -1$) leads to the collapse of the DI sensitivity and to Rayleigh's curse once the mutual phase slightly deviates from π.

Thermal state. Another example we consider is that of correlated thermal sources. The QFI for arbitrarily bright correlated thermal sources that are in-phase or anti-phase is calculated in [Lupo 2016]. A more general case with arbitrary Gaussian

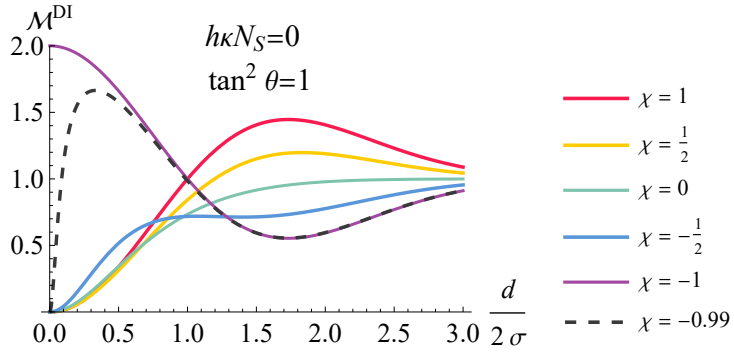

FIGURE 3.17: Normalized full separation estimation sensitivity of DI $\mathcal{M}^{DI}$ for equally bright sources with Poisson statistics.

sources is considered in [Sorelli 2022]. The QFI obtained in these papers coincides with the sensitivity $M(d)$ introduced here. With increasing intensity of the correlated thermal source, the sensitivity per photon $\mathcal{M}$ drops tending to $\mathcal{M}_\varepsilon$ (fig. 3.13).

Fig. 3.16 shows that source statistics do not strongly influence the sensitivity if the sources are mutually coherent. We notice that the resulting sensitivity of SPADE $\mathcal{M}$ is continuous as a function of χ, which is not the case for DI.

3.3.8 Unknown mutual phase and power imbalance

Now, let us consider the scenario where the mutual phase ϕ and the power imbalance $p = \cos^2\theta$ are also unknown parameters.

SPADE

In the SPADE approach, parameters ϕ and θ only appear in the form of the combination $\chi = \sin 2\theta \cos\phi$. On one hand, this implies that they cannot be estimated separately. On the other hand, it means that there is one less parameter to estimate, and the list of estimated parameters can be reduced to $\vec{\theta} = (d, \chi, N_S)$, as we keep the assumption of the known centroid position.

For simplicity, we consider this case in the Poissonian limit, i.e. $h = 0$. Then the moment matrix eq. (3.75) simplifies to

$$M_{\alpha\beta} = \sum_m \frac{1}{N_m} \frac{\partial N_m}{\partial \theta_\alpha} \frac{\partial N_m}{\partial \theta_\beta}, \tag{3.118}$$

with N_m given by eq. (3.94) $N_m = \kappa N_S(1 + (-1)^m\chi)\,\beta_m^2$. The derivatives of N_m with respect to the parameters $\vec{\theta} = (d, \chi, N_S)$ are given by

$$\frac{\partial N_m}{\partial d} = \kappa N_S\Big(1 + (-1)^m\chi\Big)2\beta_m\beta_m', \tag{3.119}$$

$$\frac{\partial N_m}{\partial \chi} = \kappa N_S\,(-1)^m\,\beta_m^2, \tag{3.120}$$

$$\frac{\partial N_m}{\partial N_S} = \kappa\Big(1 + (-1)^m\chi\Big)\beta_m^2. \tag{3.121}$$

Using these expressions most of the elements of the moment matrix (3.118) can be found without difficulties. The only non-trivial case is the calculation of M_{22}, where the summation index in the denominator is not reduced:

$$M_{22} = \sum_m \frac{1}{N_m}\left(\frac{\partial N_m}{\partial \chi}\right)^2 = \sum_m \frac{\kappa N_S\beta_m^2}{1 + (-1)^m\chi} = \kappa N_S\left(\sum_m \frac{\beta_{2m}^2}{1+\chi} + \sum_m \frac{\beta_{2m+1}^2}{1-\chi}\right). \tag{3.122}$$

The even and odd sum of the overlaps can be found as

$$\sum_m \beta_{2m}^2 = \frac{1}{2}\sum_m\left(\beta_m^2 + (-1)^m\beta_m^2\right) = \frac{1}{2}(A + \delta), \tag{3.123}$$

and likewise

$$\sum_m \beta_{2m+1}^2 = \frac{1}{2}\sum_m\left(\beta_m^2 - (-1)^m\beta_m^2\right) = \frac{1}{2}(A - \delta). \tag{3.124}$$

Thus, the M_{22} element of the moment matrix reads

$$M_{22} = \frac{\kappa N_S}{2}\left(\frac{A+\delta}{1+\chi} + \frac{A-\delta}{1-\chi}\right) = \kappa N_S\frac{A - \delta\chi}{1-\chi^2}. \tag{3.125}$$

Note that this expression becomes singular for $|\chi| = 1$. This occurs because the parameter χ is constrained by construction to $|\chi| \leq 1$, meaning that an unbiased estimator for parameter values lying on the boundary of this range ($|\chi| = 1$) does not exist. Further, we avoid this pathological case.

The full moment matrix of the SPADE measurement is then given by

$$\boxed{\boldsymbol{M}(d, \chi, N_S) = \kappa N_S\begin{pmatrix} \Delta k^2 - \chi\beta & \dfrac{\partial\delta}{\partial d} & \dfrac{\chi}{N_S}\dfrac{\partial\delta}{\partial d} \\ \dfrac{\partial\delta}{\partial d} & \dfrac{A - \delta\chi}{1-\chi^2} & \dfrac{1}{N_S}\delta \\ \dfrac{\chi}{N_S}\dfrac{\partial\delta}{\partial d} & \dfrac{1}{N_S}\delta & \dfrac{1}{N_S^2}(A + \delta\chi) \end{pmatrix}}. \tag{3.126}$$

Direct imaging

As before, the sensitivity of the DI is calculated numerically. For this purpose, we use the continuous limit of the expression for the moment matrix (3.118):

$$M^{DI}_{\alpha\beta} = \int d\vec{r} \frac{1}{I(\vec{r})} \frac{\partial I(\vec{r})}{\partial \theta_\alpha} \frac{\partial I(\vec{r})}{\partial \theta_\beta}, \tag{3.127}$$

where the intensity distribution $I(\vec{r})$ is defined in eq. (3.86). Generally, the measured signal in this case depends on four parameters: $\vec{\theta} = (d, p, \phi, N_S)$. In this section, we study approximately balanced sources $p \approx 1/2$ which is typically the most challenging scenario for DI. The results of the numerical computations are presented in the next section.

Sensitivity of the separation estimation S_d in the multiparameter setting

The inverse moment matrix establishes the bound (2.53) for the covariance matrix of the moment-based estimators. We assume our primary interest lies in estimating the separation d, while other parameters are unknown but not of interest to us (such parameters are often referred to as *nuisance parameters* [Suzuki 2020]). The sensitivity to the separation d is given by the following expression

$$S_d(\vec{\theta}) = \frac{1}{\left[\boldsymbol{M}^{-1}(\vec{\theta})\right]_{11}}, \tag{3.128}$$

where the parameters in the list $\vec{\theta}$ are unknown, and all other parameters are assumed to be known exactly.

In this section, we sequentially expand the list of unknown parameters $\vec{\theta}$ to track the change in the separation estimation sensitivity $S_d(\vec{\theta})$ with less and less prior information about the studied system. For building the plots we use the normalized version of the sensitivity

$$\mathcal{S}_d(\vec{\theta}) = \frac{4\sigma^2}{\kappa} \frac{S_d(\vec{\theta})}{N_S}. \tag{3.129}$$

This normalization eliminates dependence on the parameters σ (except for the ratio d/σ), κ, and N_S, simplifying the analysis.

We begin with the simplest case $\vec{\theta} = (d)$, which was already studied above. In this case, we assume that the separation is the only unknown parameter. The separation estimation sensitivity then coincides with the first element of the moment matrix $S_d(d) = M_{11}$ (3.118):

$$S_d(d) = \kappa N_S \left(\Delta k^2 - \chi\beta\right). \tag{3.130}$$

The direct comparison of the normalized sensitivity of SPADE and DI for $\vec{\theta} = (d)$ is shown in fig. 3.18. This figure essentially reproduces the fig. 3.16 and fig. 3.17 together for easier comparison.

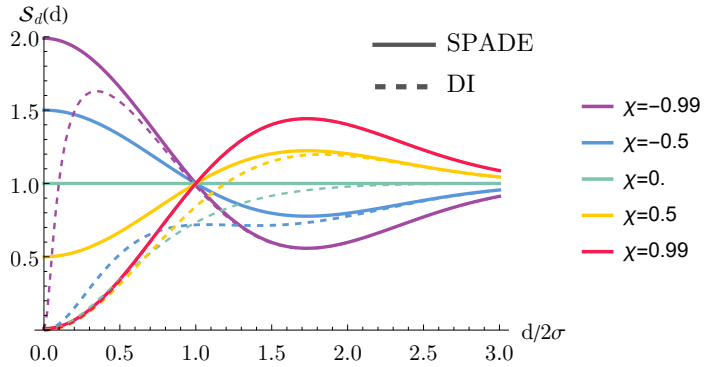

FIGURE 3.18: Normalized separation estimation sensitivity $\mathcal{S}_d(d)$ of SPADE (solid) and DI (dashed).

Another case, that was also analyzed before is the estimation of two parameters $\vec{\theta} = (d, N_S)$. In this case the separation estimation sensitivity of SPADE is given by

$$S_d(d, N_S) = \frac{1}{[\boldsymbol{M}^{-1}(d, N_S)]_{11}} = N_D M_\varepsilon = M(d) - \kappa N_S \, \delta'^2 \frac{\chi^2}{A + \delta\chi}, \tag{3.131}$$

and is plotted in fig. 3.13. The DI sensitivity for this case is presented in fig. 3.12.

In the next example, we consider $\vec{\theta} = (d, \chi)$, where in the context of SPADE, this corresponds to having no prior information about the power imbalance $p = \cos^2\theta$ and/or the relative phase ϕ. The separation estimation sensitivity of SPADE in this case is expressed as follows:

$$S_d(d, \chi) = \frac{1}{[\boldsymbol{M}^{-1}(d, \chi)]_{11}} = M(d) - \kappa N_S \, \delta'^2 \frac{1 - \chi^2}{A - \delta\chi}. \tag{3.132}$$

Figure 3.19 illustrates the decrease in the sensitivity of SPADE caused by the presence of the nuisance parameter χ. It is interesting to note that this drop only occurs for intermediate separations, while for very small ($d \ll \sigma$) and sufficiently large ($d > 5\sigma$) separations, not knowing χ does not complicate the estimation process. This happens because the image overlap δ (3.55) weakly depends on the separation for large and small d, i.e. $\delta'(d) \approx 0$ for $d \ll \sigma$ or $d \gg \sigma$.

Figure 3.20 compares the sensitivity of the SPADE and DI approaches. It is interesting to note that not knowing the relative phase and power ratio of the sources decreases the sensitivity of the DI mostly in the range $2\sigma < d < 6\sigma$. Another interesting property is that knowing the relative power p (but not knowing the phase ϕ) of approximately balanced sources ($p \approx 1/2$) does not provide any useful information for the separation estimation with DI: $S_d^{DI}(d, p, \phi)|_{p=0.5} = S_d^{DI}(d, \phi)$.

Finally, we consider the case where we know neither the total power N_S nor the relative brightness p and phase ϕ of the sources. This corresponds to estimating three parameters with SPADE $\vec{\theta} = (d, \chi, N_S)$. The separation estimation sensitivity

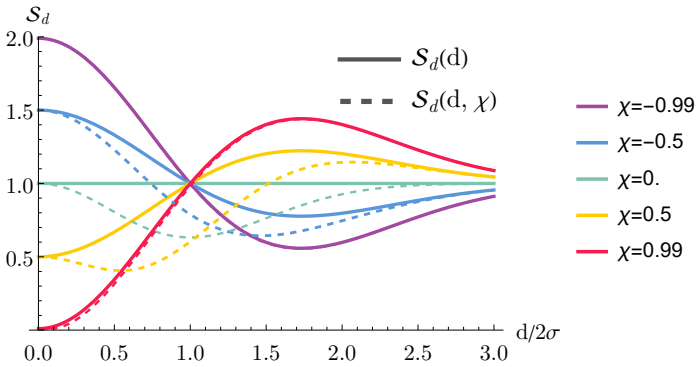

FIGURE 3.19: Normalized separation estimation sensitivity of SPADE $\mathcal{S}_d(d)$ (solid) and $\mathcal{S}_d(d,\chi)$ (dashed).

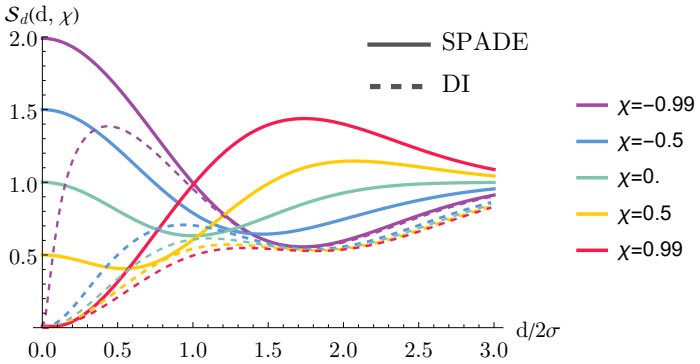

FIGURE 3.20: Normalized separation estimation sensitivity $\mathcal{S}_d(d,\chi)$ of SPADE (solid) and DI (dashed).

of the SPADE in this case is given by:

$$S_d(d,\chi,N_S) = \frac{1}{[\boldsymbol{M}^{-1}(d,\chi,N_S)]_{11}} = M(d) - \kappa N_S\,\delta'^2\frac{A-\delta\chi}{A^2-\delta^2}. \tag{3.133}$$

As one can see in fig. 3.21, this case corresponds to the presence of the Rayleigh curse for SPADE measurements, even though the measurement basis is perfectly aligned to the centroid of the sources. The measurement result of the DI depends on four parameters $\vec{\theta} = (d,p,\phi,N_S)$, however the sensitivity of the separation estimation again does not depend on accessibility of p when sources are approximately balanced $p \approx 1/2$, i.e. $S_d^{DI}(d,p,\phi,N_S)|_{p=0.5} = S_d^{DI}(d,\phi,N_S)|_{p=0.5}$.

It is interesting to note that in this case, SPADE works better for in-phase sources $\phi \approx 0$ (red line) compared to antiphase $\phi \approx \pi$ (purple line), regardless of the separation d. At the same time, the sensitivity of DI very weakly depends on the mutual phase ϕ.

For both techniques, SPADE and DI, the sensitivity goes to zero for small separations. Therefore, we also plot it in a double logarithmic scale on the right of fig. 3.21, to observe its scaling with the separation. The sensitivity of SPADE (3.133) scales like $S_d(d,\chi,N_S) \propto O(d^2)$, while numerical analysis shows the scaling of the DI sensitivity

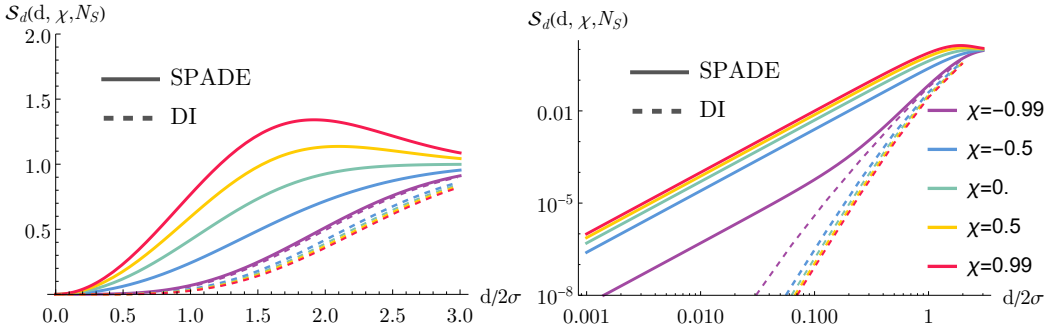

FIGURE 3.21: Normalized separation estimation sensitivity $\mathcal{S}_d(d, \chi, N_S)$ of SPADE (solid) and DI (dashed). On the left — linear scale, on the right — double logarithmic scale.

to be $S_d^{DI}(d, \phi, N_S) \propto O(d^6)$. Thus one needs substantially lower measurement statistics to resolve two mutually coherent sources with unknown brightness and relative phase (but known centroid position) when using SPADE measurement, compared to the DI.

3.3.9 Conclusion

In this section, we presented a general approach to analyze parameter estimation problems based on photon counting in mutually coherent modes. The sensitivity measure based on the method of moments showed to be a very efficient and practical tool for analyzing this class of problems. In contrast to the traditional approach based on Fisher information, moment-based sensitivity allows to consider the sources with arbitrary quantum statistics and provide a simple estimator for the parameters that does not require measurement of high-order moments. These results can be applied to other parameter estimation problems with mutually coherent or single-mode sources, like coherent imaging or distributed quantum sensing with mutually coherent probes.

Specifically, we have considered in detail the problem of separation estimation of two mutually coherent sources. Calculating the moment-based sensitivity, we analytically proved an advantage of SPADE over direct imaging for the considered class of states. Moreover, we showed that the sensitivity of SPADE saturates the QFI for all cases for which the latter is known. This even includes some examples of non-Gaussian entangled states, although they are not fully described by the first two moments, that we used to compute the sensitivity.

We showed that in this case the sensitivity (3.79) consists of two terms that correspond to the measurement of relative photon numbers and total number of detected photons, respectively. The first term only depends on the measurement basis, while the second one depends on the quantum statistics of the sources. Moreover, the second term vanishes in the case of unknown brightness of the sources, wiping out any advantage from non-classical statistics of the sources. The sensitivity from the total photon number measurement is also negligible for intense bunched states, due to the high noise in the total number of photons.

We demonstrate that the presence of other nuisance parameters, such as the relative phase and power ratio of sources, reduces the sensitivity of SPADE and, in some cases, even leads to a resurgence of the Rayleigh curse. However, the scaling of sensitivity with small separation d stays significantly better for SPADE measurement compared to the DI approach.

3.4 Partially coherent sources

In this section, we study the resolving of a pair of equally bright sources ($p = 1/2$) with Gaussian statistics and arbitrary mutual coherence $\gamma = \gamma_0 e^{i\phi}$. Most of the text of this section is adapted from [Karuseichyk 2024] with some modification of notations. In the considered case the coherency matrix (3.29) takes the form

$$\Gamma^{(1)\hat{s}}_{jk} = \langle \hat{s}^\dagger_j \hat{s}_k \rangle = \frac{1}{2} \begin{pmatrix} N_S & \gamma N_S \\ \gamma^* N_S & N_S \end{pmatrix}. \tag{3.134}$$

In this section, we also consider cases where the brightness N_S and mutual coherence γ of the sources depend on the separation d.

In this part, we only describe the SPADE measurement in the symmetric basis $f^{(s)}_m(x)$ with parity (3.49), and do not consider the DI approach. Using the property (3.40) we find the average photon numbers in the measurement modes to be

$$N_m = \langle \hat{N}_m \rangle = \langle \hat{a}^\dagger_m \hat{a}_m \rangle = \kappa N_S \beta^2_m \left(1 + (-1)^m \gamma_0 \cos\phi\right) = \xi_m + \zeta_m \gamma_0 \cos\phi, \tag{3.135}$$

where

$$\xi_m = \kappa N_S \beta^2_m, \quad \zeta_m = \kappa N_S \beta^2_m \, (-1)^m. \tag{3.136}$$

The second term in (3.135) accounts for light interference arising from partial coherence. The reader may notice that the average detected photon numbers depend solely on the real part of the degree of mutual coherence, $\mathrm{Re}\,\gamma = \gamma_0 \cos\phi$. However, as we show in the following, the imaginary part $\mathrm{Im}\,\gamma$ does influence the photon counting statistics and, consequently, changes the sensitivity of the parameter estimation.

The total number of photons detected in K measurement modes equals

$$N_D = \sum_{m=0}^{K} N_m = \kappa N_S (A + \delta \gamma_0 \cos\phi), \tag{3.137}$$

where measurement basis completeness A and overlap δ are defined in eqs. (3.53) and (3.55) respectively. The total number of detected photons N_D (3.137) depends on the separation d through the overlap δ, and possibly through $N_S(d)$ and $\gamma(d)$. One can leverage these dependencies to achieve a more accurate estimation of the separation d, only provided that the losses are correctly accounted for in the model [Tsang 2019d; Kurdzialek 2022; Liang 2023b], thus lossless models tend to underestimate the sensitivity in this case [Larson 2018; Larson 2019; Liang 2021; Hradil 2021; Hradil 2019; Wang 2023].

To find the photon number covariance matrix

$$\Gamma_{mn} = \langle \hat{N}_m \hat{N}_n \rangle - \langle \hat{N}_m \rangle \langle \hat{N}_n \rangle = \langle \hat{a}^\dagger_m \hat{a}_m \hat{a}^\dagger_n \hat{a}_n \rangle - N_m N_n \tag{3.138}$$

we use the following property of non-displaced Gaussian states [Phillips 2019][5]:

$$\langle \hat{a}^\dagger_m \hat{a}_m \hat{a}^\dagger_n \hat{a}_n \rangle = \langle \hat{a}^\dagger_m \hat{a}_m \rangle \langle \hat{a}^\dagger_n \hat{a}_n \rangle + \langle \hat{a}^\dagger_m \hat{a}^\dagger_n \rangle \langle \hat{a}_m \hat{a}_n \rangle + \langle \hat{a}^\dagger_m \hat{a}_n \rangle \langle \hat{a}_m \hat{a}^\dagger_n \rangle. \tag{3.139}$$

[5] This property is an analog of Wick's theorem [Wick 1950] for quantum optics.

Taking into account global phase averaging of the emitted state (3.30), the second term in this expression vanishes and we find that

$$\Gamma_{mn} = \langle \hat{a}_m^\dagger \hat{a}_n \rangle \langle \hat{a}_m \hat{a}_n^\dagger \rangle = \Gamma_{mn}^{(1)\hat{a}} \left(\delta_{mn} + \Gamma_{mn}^{(1)\hat{a}} \right) \tag{3.140}$$

with

$$\Gamma_{mn}^{(1)\hat{a}} = \langle \hat{a}_m^\dagger \hat{a}_n \rangle \tag{3.141}$$

being the coherency matrix calculated in the measurement mode basis. Since it is the normal ordered average, we compute this matrix using the property (3.40), yielding

$$\Gamma_{mn}^{(1)\hat{a}} = \frac{\kappa N_S}{2} \beta_m \beta_n \left(1 + (-1)^{m+n} + \gamma (-1)^n + \gamma^* (-1)^m \right). \tag{3.142}$$

One can recover the photon numbers (3.135) from the diagonal part of this matrix as $N_m = \Gamma_{mm}^{(1)\hat{a}}$.

Substituting eq. (3.142) to the eq. (3.140), after some mathematical transformations we arrive at the following expression for the photon number covariance matrix:

$$\mathbf{\Gamma} = \operatorname{diag}\left(\vec{Z}_1 + \vec{Z}_2 \right) + \vec{Z}_1 \vec{Z}_1^T + \vec{Z}_2 \vec{Z}_2^T + \frac{1 - \gamma_0^2}{2} \sin^2 \phi \ \vec{\zeta} \vec{\zeta}^T, \tag{3.143}$$

where $\operatorname{diag}(\vec{x})$ is a diagonal matrix with diagonal elements equals x_n and

$$\vec{Z}_{1,2} = \frac{1 \pm \gamma_0}{2} \left(\vec{\xi} \pm \vec{\zeta} \cos \phi \right), \tag{3.144}$$

with $\vec{\xi}$ and $\vec{\zeta}$ defined in eq. (3.136).

In the limit of faint sources, $\kappa N_S \ll 1$, one can neglect $(\kappa N_S)^2$ and the covariance matrix (3.143) becomes diagonal $\Gamma_{nm} = \delta_{nm} N_m$, that corresponds to the Poissonian limit, discussed in section 2.2.5. In this case, the MoM sensitivity M (2.46) coincides with the FI and it only depends on the real part of the mutual coherence $\operatorname{Re} \gamma = \gamma_0 \cos \phi$. In the more general scenario, determining the sensitivity M requires inverting the non-diagonal photon number covariance matrix $\mathbf{\Gamma}$ (3.143). It is possible to do this analytically by applying the Sherman-Morrison formula (3.72) three times, but the resulting general expressions are very bulky. Therefore, in the following sections, we focus on several special cases for which the inversion simplifies.

3.4.1 Constant mutual coherence

Real-valued coherence ($\gamma = \gamma_0$)

First, we consider separation-independent real-valued coherence $\gamma = \gamma_0 \in \mathbb{R}$, i.e. $\phi = 0$. For this case, the photon number covariance matrix reduces to:

$$\mathbf{\Gamma} = \operatorname{diag}\left(\vec{Z}_1 + \vec{Z}_2 \right) + \vec{Z}_1 \vec{Z}_1^T + \vec{Z}_2 \vec{Z}_2^T. \tag{3.145}$$

The inversion of this matrix gives

$$\Gamma_{mn}^{-1} = \delta_{mn} (\xi_n + \gamma_0 \zeta_n)^{-1} - t_{mn}, \tag{3.146}$$

where

$$t_{mn} = \begin{cases} 2/h_1, & \text{if } m \text{ and } n \text{ are both even} \\ 2/h_2, & \text{if } m \text{ and } n \text{ are both odd} \\ 0, & \text{otherwise,} \end{cases} \tag{3.147}$$

with

$$h_{1,2} = 2 + \kappa N_S(1 \pm \gamma_0)(A \pm \delta), \tag{3.148}$$

and A and δ are defined in Eqs. (3.53) and (3.55).

If the mutual coherence γ and the brightness N_S are independent of the separation, i.e. $\partial\gamma_0/\partial d = 0$, and $\partial N_S/\partial d = 0$, the derivative of the measured signal (3.135) is

$$\frac{\partial N_m}{\partial d} = 2\frac{N_m}{\beta_m}\frac{\partial \beta_m}{\partial d}. \tag{3.149}$$

Then, using the expression (2.46), the normalized sensitivity of the separation estimation can be found as

$$\frac{M_{\mathrm{Re}}}{\kappa N_S} = \Delta k^2 - \gamma_0\beta - \kappa N_S(A' + \delta')^2\left(\frac{(1+\gamma_0)^2}{2h_1} + \frac{(1-\gamma_0)^2}{2h_2}\right), \tag{3.150}$$

where parameters $A, \delta, \Delta k^2, \beta$ are defined in eqs. (3.53) to (3.58). The sensitivity (3.150) coincides with the quantum Fisher information [Lupo 2016; Sorelli 2022], which implies the possibility of constructing an estimator $\tilde{d}(\{\overline{N}_m^{(\mu)}\})$ based solely on the observed sample mean numbers of counts $\overline{N}_m^{(\mu)}$ in any basis with the parity defined in eq. (3.49) and achieving the ultimate resolution limit dictated by the quantum Cramér-Rao bound.

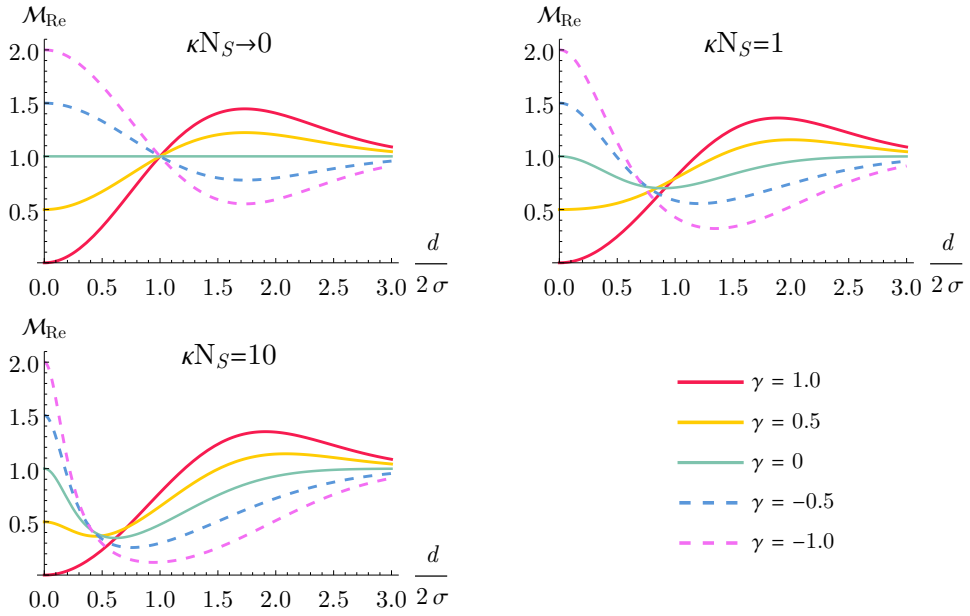

FIGURE 3.22: The normalized sensitivity $\mathcal{M}_{\mathrm{Re}} = M_{\mathrm{Re}} 4\sigma^2/\kappa N_S$ for different real values of the mutual coherence γ and numbers of emitted photons N_S vs. separation d.

The sensitivity calculated with eq. (3.150) for a Gaussian PSF (3.26) and full measurement basis ($K \to \infty$) is shown in fig. 3.22. The first plot corresponds to a small number of detected photons and matches other results obtained in the Poissonian limit [Tsang 2019d; Kurdzialek 2022]. With an increasing number of photons, there is a reduction in the sensitivity for the separations $d \approx \sigma$. This reduction is also present in the QFI and occurs due to the quadratic term of the noise in the thermal statistics [Lupo 2016]. However, as we show in the next section, this drop does not occur even for bright thermal states, if their mutual coherence degree γ is an imaginary number.

Imaginary coherence ($\gamma = i\gamma_0$)

Another interesting example that we consider is that of a purely imaginary degree of coherence $\gamma = i\gamma_0 \in \mathbb{I}$, which corresponds to a relative phase $\phi = \pi/2$ between the sources. In this case, the measured mean photon numbers N_m (3.135) do not contain an interference term, i.e. it is the same as for a pair of incoherent sources. However, the presence of coherence does influence the covariance matrix

$$\mathbf{\Gamma} = \operatorname{diag}\left(\vec{\xi}\right) + \frac{1+\gamma_0^2}{2}\,\vec{\xi}\vec{\xi}^T + \frac{1-\gamma_0^2}{2}\,\vec{\zeta}\vec{\zeta}^T, \tag{3.151}$$

which only depends on γ_0^2 and thus is not affected by its sign. The inverse of this matrix with Sherman-Morrison formula (3.72) gives

$$\Gamma^{-1}_{mn} = \delta_{mn}\xi_n - \frac{(-1)^{m+n}\eta_1 + \eta_2 - \kappa N_S\delta((-1)^n + (-1)^m)}{\eta_1\eta_2 - (\kappa N_S\delta)^2}, \tag{3.152}$$

where $\eta_{1,2} = \kappa N_S A + 2(1 \pm \gamma_0^2)^{-1}$. Calculating the sensitivity (2.46) with the derivative vector (3.149), we obtain

$$\frac{M_{\mathrm{Im}}}{\kappa N_S} = \Delta k^2 - \kappa N_S \frac{\eta_1(\delta')^2 + \eta_2(A')^2 - \kappa N_S\delta A'\delta'}{\eta_1\eta_2 - (\kappa N_S\delta)^2}. \tag{3.153}$$

This sensitivity calculated for Gausian PSF (3.26) is plotted in fig. 3.23 for different values of the mutual coherence and sources' intensities.

One can see that in the limit of faint sources (Poissonian limit) the sensitivity

$$\frac{M_{\mathrm{Im}}}{\kappa N_S} \xrightarrow{\kappa N_S \to 0} \Delta k^2 \tag{3.154}$$

is the same for any imaginary degree of mutual coherence, including the case of incoherent sources.

However, for bright incoherent thermal sources the sensitivity per photon drops significantly in the sub-Rayleigh region $d < \sigma$, while for sources with imaginary mutual coherence, this drop is smaller or doesn't occur at all for perfect coherence $\gamma = i$. Generally, if imaginary coherence is close to the complex unity then the

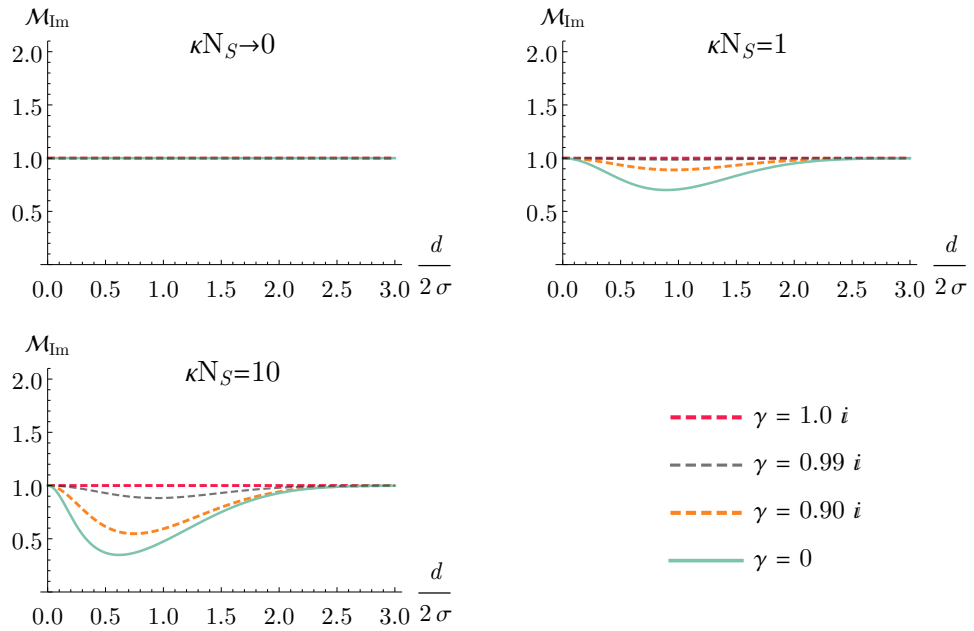

FIGURE 3.23: The normalized sensitivity $\mathcal{M}_{\mathrm{Im}} = M_{\mathrm{Im}} 4\sigma^2/\kappa N_S$ for different imaginary values of the mutual coherence γ and numbers of emitted photons N_S vs. separation d.

sensitivity can be expressed as

$$\frac{M_{\mathrm{Im}}}{\kappa N_S} = \Delta k^2 - \kappa N_S (\delta')^2 \, (1 - \gamma_0) + O((1 - \gamma_0)^2). \tag{3.155}$$

This means that the presence of correlations between sources can improve the sensitivity of the separation estimation even in cases when no interference is visible in the mean values N_m. One can use a simple intuition to interpret this observation: two incoherent thermal sources do not share any correlations, while mutually coherent sources are correlated in the number of photons, even if they do not interfere due to the mutual phase $\phi = \pi/2$, since the interference term is zero in this case. Utilizing these correlations allows to cancel part of the intensity noise in the measurement modes, resulting in a more precise estimation of the separation.

3.4.2 Parameter-dependent coherence

Now let us consider the situation when the mutual coherence γ depends on the separation d of the sources. Below we will discuss in detail physical examples of such systems, but first, we analyze the general case of real-valued separation-dependent mutual coherence $\gamma(d) = \gamma_0(d) \in \mathbb{R}$.

This dependence leads to an additional term in the derivative of the mean measured numbers of photons (3.135):

$$\frac{\partial N_m}{\partial d} = 2 \frac{N_m}{\beta_m} \frac{\partial \beta_m}{\partial d} + \frac{\partial \gamma_0}{\partial d} \zeta_m. \tag{3.156}$$

Then, an extra term also appears in the sensitivity of the separation estimation

$$M_{\gamma_0(d)} = M_{\mathrm{Re}} + \Delta M_{\gamma_0(d)}. \tag{3.157}$$

where

$$\frac{\Delta M_{\gamma_0(d)}}{\kappa N_S} = 2\gamma_0' \left[\delta' \left(\frac{1}{h_1} + \frac{1}{h_2} \right) + A' \left(\frac{1}{h_1} - \frac{1}{h_2} \right) \right] + (\gamma_0')^2 \left[\frac{A - \delta\gamma_0}{1 - \gamma_0^2} - \kappa N_S \left(\frac{(A+\delta)^2}{2h_1} + \frac{(A-\delta)^2}{2h_2} \right) \right], \tag{3.158}$$

with $\gamma_0' = \partial\gamma_0/\partial d$ and M_{Re} is defined in eq. (3.150).

The additional sensitivity $\Delta M_{\gamma_0(d)}$ originates from the fact that in the case of separation-dependent coherence, the measurement results change faster with changing separation, and one can do estimation with higher precision. The extra sensitivity $\Delta M_{\gamma_0(d)}$ typically has maxima around the maxima of γ_0', i.e. in the regions of fast-changing mutual coherence.

In the limit of faint sources, the additional sensitivity takes the form

$$\frac{\Delta M_{\gamma_0(d)}}{\kappa N_S} \xrightarrow{\kappa N_S \to 0} 2\gamma_0'\delta' + (\gamma_0')^2 \frac{A - \gamma_0\delta}{1 - \gamma_0^2}. \tag{3.159}$$

This expression is valid for an arbitrary complex value of the mutual coherence γ if one replaces γ_0 with $\mathrm{Re}\,\gamma$.

3.4.3 Finite coherence width of the illumination

In the first example, we investigate the estimation of separation between two reflective objects. When illuminated with light of finite coherence width, the mutual coherence of the reflected light becomes dependent on the separation between the reflectors (see fig. 3.24).

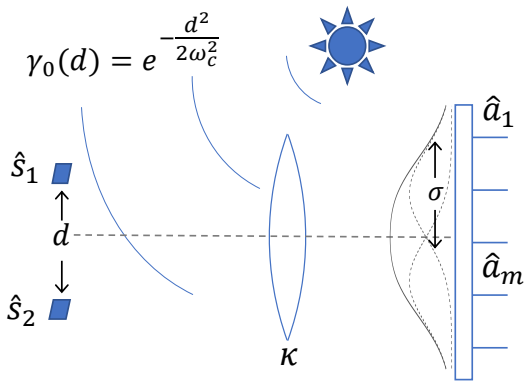

FIGURE 3.24: Objects reflect the light from the common thermal illumination source with finite coherence width ω_c.

In many cases, the transverse coherence of the illuminating light can be effectively approximated by a Gaussian function. In such scenarios, the field reflected by two small reflectors exhibits mutual coherence

$$\gamma_0(d) = \exp\left[-\frac{d^2}{2\omega_c^2}\right], \tag{3.160}$$

where the coherence width ω_c can be adjusted by changing the optical parameters of the illumination system.

Using the Gaussian PSF (3.26), we calculate the full sensitivity (3.157) of the reflectors' separation estimation (fig. 3.25). The red lines on the plots correspond to the spatially coherent illumination. One can see that separation-dependent coherence, compared to the fully coherent case ($\omega_c = \infty$), results in significantly higher sensitivity for the separations close to the coherence width of the illumination source $d \approx \omega_c$, thus it is desirable to use illumination with coherence width of the order of the measured separations. A similar increase in the resolution, when the coherence width of the source matches the size of the features of the studied object, was observed earlier experimentally and numerically for the case of quantum imaging with pseudo-thermal light [Mikhalychev 2019].

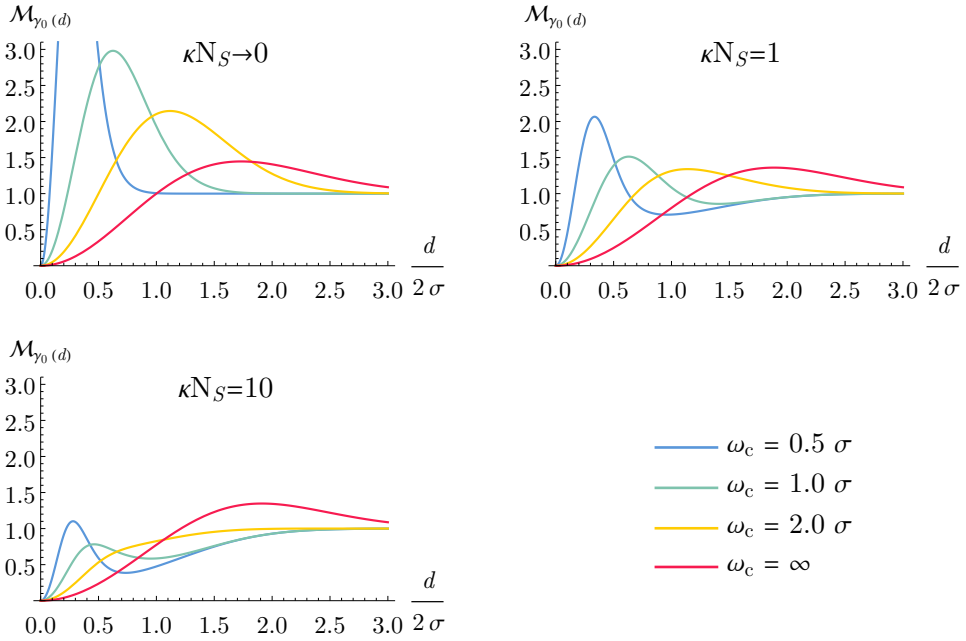

FIGURE 3.25: Normalized sensitivity $\mathcal{M}_{\gamma_0(d)} = M_{\gamma_0(d)} 4\sigma^2/\kappa N_S$ in case of the separation-dependent mutual coherence with Gaussian profile (3.160).

The coherence width of the light expands as it passes through a diffractive imaging system. Thus, if the same optical system (with PSF width σ) is used for the imaging and the illumination of the object, the minimal achievable coherence width of the illumination is $\omega_c \geq 2\sigma$. This case was considered in the independent work [Wang 2023] in the faint source limit, where an analogous increase of the QFI was

demonstrated for the separations around the Rayleigh limit (in comparison to spatially coherent illumination). However, as we show in fig. 3.25, for brighter sources this increase is much less noticeable. In the extreme case of very bright sources, additional sensitivity per photon (3.158) vanishes

$$\frac{\Delta M_{\gamma_0(d)}}{\kappa N_S} \xrightarrow[N_S\to\infty]{} 0. \tag{3.161}$$

This is important to take into account for practical imaging with pseudo-thermal sources since this scenario often does not fit into the faint sources approximation. Therefore, it is crucial to aim for the utilization of an illumination source with a short coherence time to ensure that the number of photons N_S reflected by each imaged object per coherence time is small.

One could achieve better sensitivity in the sub-Rayleigh regime via a smaller coherence width, which is achievable in two possible scenarios: using an independent illumination scheme with narrower PSF (for example, if the illumination source is located closer to objects than the detection apparatus) or in the case of sources' mutual coherence originating from their interaction, which is considered in the next section.

3.4.4 Interacting emitters

Another possible scenario where the mutual coherence of the emitted light depends on the distance between the sources is the case of interacting emitters (see fig. 3.26). As an example, we consider two identical, dipole-dipole interacting two-level systems prepared initially in their excited states. In this scenario, each dipole emits precisely 1 photon during the decay of the excited state (sources decay freely, without additional pump during the emission process). For simplicity, we assume that the dipole moments are parallel to each other.

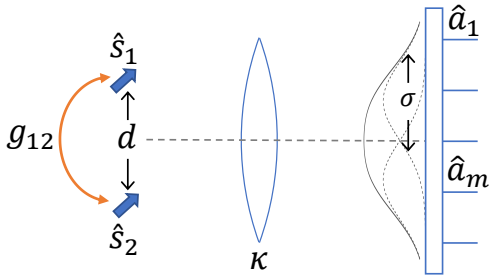

FIGURE 3.26: Dipoles emit partially coherent light due to the interaction.

The time evolution of the quantum state of the two coupled dipoles in the interaction picture can be described by the following master equation [Lehmberg 1970;

Agarwal 1974]

$$\frac{d}{dt}\hat{\rho} = -if_{12}[\hat{\sigma}_1^+\hat{\sigma}_2^- + h.c., \hat{\rho}] + \frac{1}{2}\sum_{j,l=1}^{2} g_{jl}\left(2\hat{\sigma}_j^-\hat{\rho}\hat{\sigma}_l^+ - \hat{\sigma}_j^+\hat{\sigma}_l^-\hat{\rho} - \hat{\rho}\hat{\sigma}_j^+\hat{\sigma}_l^-\right) + \sum_{j=1}^{2}\eta_j\left(2\hat{\sigma}_j^+\hat{\sigma}_j^-\hat{\rho}\hat{\sigma}_j^+\hat{\sigma}_j^- - \hat{\sigma}_j^+\hat{\sigma}_j^-\hat{\rho} - \hat{\rho}\hat{\sigma}_j^+\hat{\sigma}_j^-\right), \quad (3.162)$$

where $\hat{\sigma}_j^{\pm}$ are the dipoles' transition operators. The first term in eq. (3.162) is proportional to the collective frequency shift f_{12} of the energy levels [Agarwal 1974; James 1993]. The second term describes the spontaneous emission of the individual emitters ($j = l$) and their collective radiative decay ($j \neq l$). In the case of identical dipoles,

$$g_{11} = g_{22} = 1/\tau, \quad (3.163)$$

$$g_{12} = \frac{3}{2}\frac{1}{\tau}\left(\frac{\sin z}{z} + \frac{\cos z}{z^2} - \frac{\sin z}{z^3}\right), \quad (3.164)$$

where τ is the natural lifetime of the excited state of an individual dipole, and $z = 2\pi d/\lambda$, with λ the wavelength of the dipole transition.

The final term in the master equation accounts for the dephasing of the dipoles with the rates η_i ($i = 1, 2$), due to their potentially different local environments.

Coherency matrix of dipole emission

Now let us analyze the property of the light, emitted by the pair of interacting dipoles. In the far field of the dipole, the positive-frequency part of the emitted field operator in point $\vec{r}$ reads [Scully 1997; Mikhalychev 2018]

$$\hat{\vec{E}}(\vec{r}, t) \propto \frac{\vec{r} \times [\vec{r} \times \vec{p}]}{|\vec{r}|^3}\hat{\sigma}^-(t), \quad (3.165)$$

where $\vec{p}$ is a dipole moment. This emission mode does not have spherical symmetry, however, in the far field, it is locally indistinguishable from a spherical wave. Thus in the paraxial approximation, dipole emitters can be considered as point sources and the model developed in the previous sections can be applied to the problem of resolving dipoles.

In contrast to our previous models, the state of the light emitted by the dipoles is time-dependent. We describe it in the transient temporal modes $\hat{s}_{1,2}(t)$, of duration $\Delta t \ll \tau$, centered around time t. Due to the proportionality (3.165) between the emitted field operator $\hat{\vec{E}}$ and the dipole transition operators $\hat{\sigma}^-$, the time-dependent coherency matrix of the emitted field can be expressed through the atomic dipole correlators as

$$\Gamma_{jk}^{(1)\hat{s}(t)} = \langle\hat{\sigma}_j^+\hat{\sigma}_k^-\rangle\frac{\Delta t}{\tau}, \quad (3.166)$$

where all the constant geometrical factors, are omitted since they will later be included in the transmissivity factor κ. In this case, κ represents the coupling between the emission modes and image modes of individual sources. It implicitly accounts not only for losses of the imaging system but also for the directional profile of the emission.

The diagonal part of the coherency matrix $\Gamma_{jk}^{(1)\hat{s}(t)}$ defines the mean number of photons emitted by each dipole during the time interval Δt. Thus the emission rate is given by

$$\dot{N} = \tau\, \Gamma_{11}^{(1)\hat{s}(t)}/\Delta t = \langle \hat{\sigma}_1^+ \hat{\sigma}_1^- \rangle = \langle \hat{\sigma}_2^+ \hat{\sigma}_2^- \rangle, \tag{3.167}$$

where we multiplied it by τ to make it unitless. The degree of mutual coherence γ as defined by eq. (1.53) is given by

$$\gamma = \frac{\Gamma_{12}^{(1)\hat{s}}}{\sqrt{\Gamma_{11}^{(1)\hat{s}}\Gamma_{22}^{(1)\hat{s}}}} = \frac{\langle \hat{\sigma}_1^+ \hat{\sigma}_2^- \rangle}{\dot{N}}. \tag{3.168}$$

Both the unitless emission rate $\dot{N}$ and the mutual coherence γ depend on both time and dipoles' separation, however, we will not explicitly indicate it in our notations.

Model without dephasing ($\eta_1 = \eta_2 = 0$)

First, we examine the evolution of the dipoles' state without taking into account the dephasing process ($\eta_i = 0$). By solving analytically the master equation (3.162), we find that

$$\dot{N} = \frac{e^{-g_{11}t}}{2\left(g_{11}^2 - g_{12}^2\right)}\Big(\left(g_{11}+g_{12}\right)^2 e^{-g_{12}t} + \left(g_{11}-g_{12}\right)^2 e^{g_{12}t} - 4g_{12}^2 e^{-g_{11}t}\Big), \tag{3.169}$$

$$\dot{N}\gamma = \frac{e^{-g_{11}t}}{2\left(g_{11}^2 - g_{12}^2\right)}\Big(\left(g_{11}+g_{12}\right)^2 e^{-g_{12}t} - \left(g_{11}-g_{12}\right)^2 e^{g_{12}t} - 4g_{11}g_{12} e^{-g_{11}t}\Big), \tag{3.170}$$

where g_{jl} is defined in eqs. (3.163) and (3.164). One can see, that in the non-interacting limit $g_{12} \to 0$ the emission rate decay exponentially $\dot{N} = e^{-t/\tau}$ and the mutual coherence does not appear $\gamma = 0$. Note that the unitary component of the master equation (3.162), featuring a collective frequency shift f_{12}, does not impact the coherency matrix of the excited modes. Nevertheless, the spectra of the excited modes are influenced by f_{12}. Consequently, in instances of frequency-resolving measurement, the dependence $f_{12}(d)$ introduces an additional mechanism for parameter encoding and, in general, enhances sensitivity.

The associated emission rate $\dot{N}$ (3.169) and the degree of mutual coherence γ (3.170) for different time instances t are depicted in fig. 3.27 as functions of the separation d normalized by the wavelength λ.

One can see that at an early stage of the emission ($t = 0.02\tau$) the dipoles fluoresce almost independently, with the emission rate close to that of an individual dipole, $\dot{N} \approx e^{-t/\tau}$, and with almost no coherence, $\gamma \approx 0$. However, after some time, the dipole-dipole interaction creates correlations and the difference with the

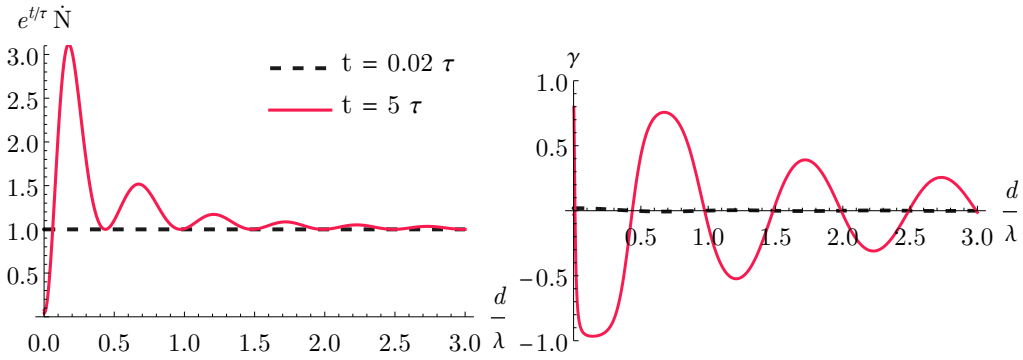

FIGURE 3.27: Properties of the dipoles radiation: the normalized emission rate $\dot{N}$ and the degree of mutual coherence γ vs. separation d. The dashed line corresponds to the time $t = 0.02\tau$, solid line $t = 5\tau$. No dephasing ($\eta = 0$).

individual dipole emission becomes evident in both the emission rate [DeVoe 1996] and the mutual coherence. Crucially, this difference depends on the separation between the dipoles, which allows for a measurement of the separation with higher accuracy. Without accounting for the decoherence process, the mutual coherence exhibits a particularly strong dependence on separation during the late emission stages ($t > \tau$) for separations $d \approx n\frac{\lambda}{2}, n \in \mathbb{N}$. Therefore, we can expect heightened sensitivity in estimating separation around these points. Note that due to the geometrical symmetry of the problem the mutual coherence γ stays real at any moment, i.e. $\gamma(t) \in \mathbb{R}\ \forall t$. However, all the conclusions below can be generalized to the case of the complex degree of mutual coherence by replacing γ_0 with Re γ.

Thus, the mean number of photons N_m, detected in the symmetric mode $\hat{a}_m(t)$ of a temporal duration Δt is given by eq. (3.135) with

$$N_S = 2\dot{N}\Delta t/\tau. \tag{3.171}$$

Considering short time intervals ($\Delta t \ll \tau$) and high losses of the paraxial optical system ($\kappa \ll 1$), we obtain that $N_S \ll 1$. Therefore, irrespective of the sources' statistics, the detection statistics can be described by a Poisson distribution which also approximates that of faint thermal sources discussed earlier. Note that in this case, we assume that the detectors have high temporal resolution, and can distinguish photo-counts in different time intervals Δt.

Now we need to calculate the derivatives of the detected signals $\partial N_m/\partial d$. The only modification required to eq. (3.156) is the addition of an extra term, associated with the separation-dependent emission rate:

$$\frac{\partial N_m}{\partial d} = 2\frac{N_m}{\beta_m}\frac{\partial \beta_m}{\partial d} + \frac{\partial \gamma_0}{\partial d}\zeta_m \cos\phi + \frac{N_m}{N_S}\frac{\partial N_S}{\partial d}. \tag{3.172}$$

This extra-dependence results in an additional term $\Delta M_{N_S(d)}$ in the sensitivity. In the limit $\kappa N_S = 2\kappa \dot{N} \Delta t/\tau \ll 1$ it takes the form

$$\Delta M_{N_S(d)} = \frac{\kappa N_S'}{N_S}\Big[N_S'(A+\gamma\delta) + 2N_S(A+\gamma\delta)'\Big], \tag{3.173}$$

where $X' = \partial X/\partial d$. Combining all the contributions we get the normalized sensitivity per unit of time (or sensitivity rate) in the form

$$\dot{\mathcal{M}} = \frac{4\sigma^2}{\kappa}\frac{\tau}{\Delta t}\left(M_{\mathrm{Re}} + \Delta M_{\gamma_0(d)} + \Delta M_{N_S(d)}\right). \tag{3.174}$$

This quantity characterizes how much information about the separation one obtains per unit of time. It is plotted in fig. 3.28 and 3.29.

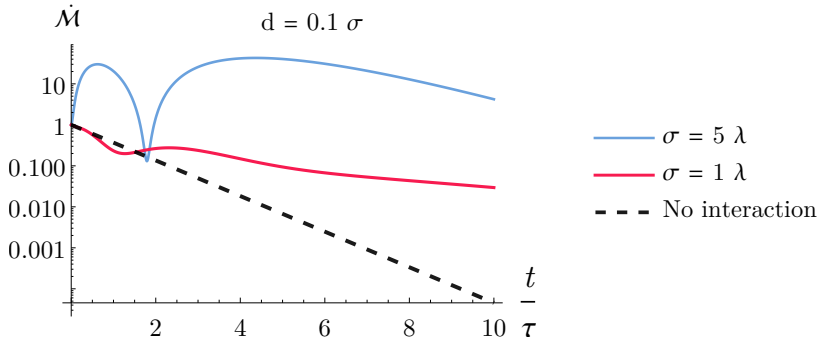

FIGURE 3.28: Normalized sensitivity rate $\dot{\mathcal{M}}$ (3.174) of two dipoles' separation estimation vs. time t. The light blue line corresponds to $\sigma = 5\lambda$, the red line to $\sigma = \lambda$, and the dashed line to non-interacting dipoles.

As one can expect, at an early stage of emission the problem resembles the resolving of two incoherent sources (i.e. non-interacting emitters). In the context of independent emitters, the sensitivity per unit of time is directly proportional to the emission rate, exhibiting an exponential decrease, as indicated by $\dot{N} \approx e^{-t/\tau}$ (depicted by the dashed black lines in fig. 3.28 and 3.29). On the contrary, in the case of interacting emitters, the sensitivity rate may even increase over time due to the cumulative effects of interaction. The specific values of the sensitivity depend on the ratio between the separation d, PSF width σ, and the wavelength λ. We present plots for the cases of $\sigma = \lambda$ and $\sigma = 5\lambda$. In the paraxial approximation, the PSF width of the imaging system typically significantly exceeds the wavelength, i.e. $\sigma > \lambda$. Therefore, the case $\sigma = \lambda$ stretches the boundaries of our model. Nevertheless, we investigate this case to explore the model's limits and offer an illustrative example that is easy to analyze.

In fig. 3.29, one can observe a pronounced enhancement in sensitivity within regions where the emission characteristics (the mutual coherence and the emission rate) exhibit rapid changes with a change in separation. Meanwhile, the sensitivity remains low around the extrema of the emission characteristics. In typical cases, the sensitivity for interacting dipoles is significantly larger and decaying significantly

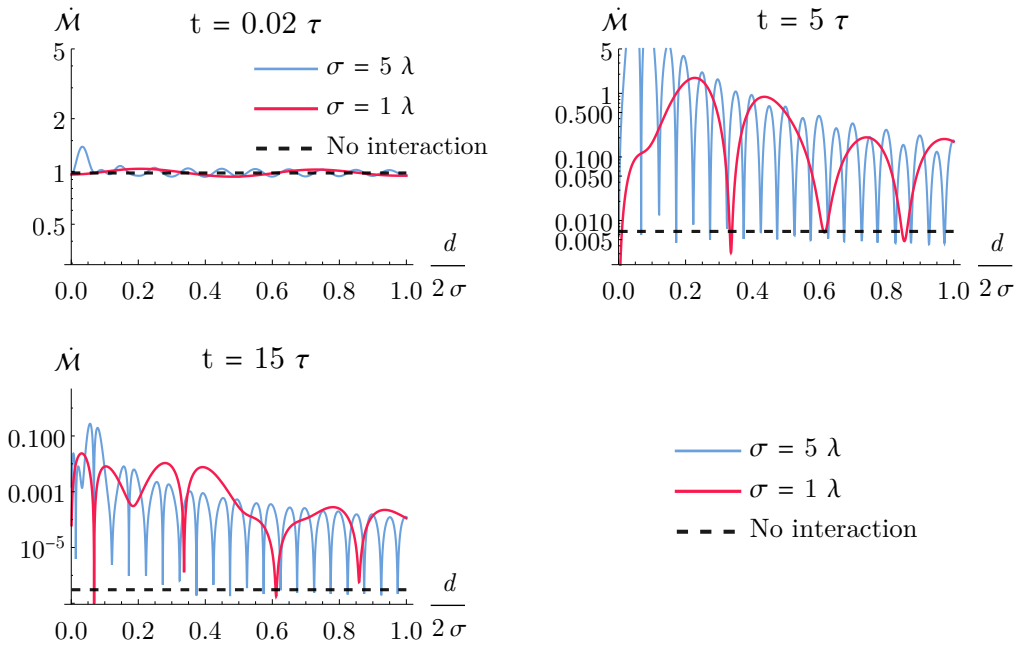

FIGURE 3.29: Normalized sensitivity rate $\dot{\mathcal{M}}$ (3.174) of two dipoles' separation estimation vs. separation d. The light blue line corresponds to $\sigma = 5\lambda$, the red line to $\sigma = \lambda$, and the dashed line to non-interacting dipoles. No dephasing ($\eta = 0$).

slower compared to non-interacting dipoles. Even after more than 15 lifetimes of the excited state, when the probability of photon detection is very low, the average sensitivity rate remains notably high due to the substantial amount of information in late detection events.

It is important to note that the moment-based sensitivity provides a limit for a local estimation strategy. The presence of multiple narrow peaks in the sensitivity plot signals the potential degeneracy of the estimator. Therefore, when dealing with a low-resolution optical system ($\sigma \gg \lambda$), one may require an increased number of measurement repetitions to reach the saturation of the bound (2.53).

In general, the interaction between the emitters induces an entanglement between them [Tanaś 2004] and creates temporal correlations of the emitted light [Peshko 2019]. However, if the losses in the imaging system are large $\kappa \ll 1$, the detection events for different time intervals can be considered independent. The sensitivity of independent detection events is additive thus the total sensitivity can be calculated as

$$\mathcal{M}_{Tot} = \frac{1}{\tau}\int_0^\infty \dot{\mathcal{M}}\, dt, \tag{3.175}$$

where $\dot{\mathcal{M}}$ is defined in (3.174). The result of this calculation is shown in fig. 3.30. One can see that the interaction effects can increase the total sensitivity of the separation estimation by several orders of magnitude around the points $d \approx n\frac{\lambda}{2}, n \in \mathbb{N}$, where the mutual coherence of the emission strongly depends on the separation, if the dephasing effect is not present. In between these points, for the separations around $d \approx \left(n - \frac{1}{2}\right)\frac{\lambda}{2}, n \in \mathbb{N}$, both the emission rate and mutual coherence exhibit extrema.

Consequently, the sensitivity is not heightened by the separation-dependent emission characteristics around these specific points.

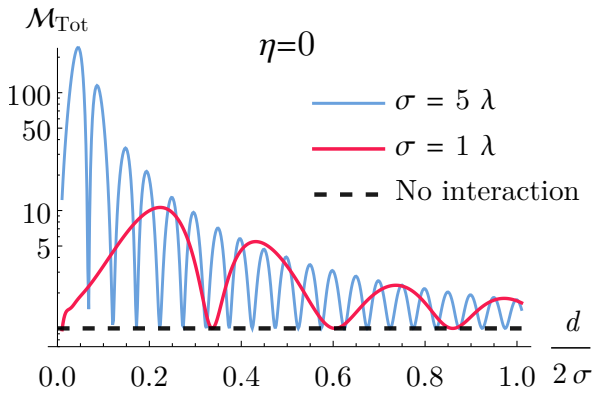

FIGURE 3.30: Normalized total sensitivity $\mathcal{M}_{Tot}$ (3.175) of two dipoles' separation estimation vs. separation d. The light blue line — $\sigma = 5\lambda$, the red line — $\sigma = \lambda$, the dashed line — non-interacting dipoles. No dephasing ($\eta = 0$).

Model with dephasing ($\eta_1,\ \eta_2 \neq 0$)

In realistic scenarios, the process of building up coherence through the dipole-dipole interactions competes with the coherence loss caused by dipoles' individual dephasing. The master equation (3.162) with dephasing can be solved analytically with computer algebra packages, however, the solution is too long and cumbersome to be presented here. Nevertheless, it demonstrates that the coherency matrix for our system only depends on the average dephasing rate $\eta = (\eta_1 + \eta_2)/2$ of the dipoles. In the following, we discuss the impact of dephasing for different values of η. Figure 3.31 illustrates the emission characteristics (the emission rate and the mutual coherence) at the time instance $t = 5\tau$. In the weak dephasing regime ($\eta = 0.2/\tau$), where the dephasing time is substantially longer than the natural lifetime of the excited state τ, the effect of dephasing has a limited influence on the emission process. On the other hand, when the dephasing process is faster than the emission ($\eta = 2.0/\tau$), one can observe a significant decrease in the mutual coherence γ. As the dephasing rate η increases, the emission characteristics approach those of non-interacting emitters.

As expected, the reduction in emission coherence due to the dephasing effect leads to a weaker sensitivity boost. In fig. 3.32 we plot the total sensitivity of the separation estimation defined in eq. (3.175). The left plot highlights the robustness of the considered scheme to weak dephasing, where the sensitivity is only minimally affected, compared to the dephasing-free case, shown in fig. 3.30. Conversely, the right plot in fig. 3.32 reveals that strong dephasing significantly diminishes the sensitivity boost arising from dipole interaction. However, even in this scenario, the sensitivity can be several times higher compared to non-interacting dipole emission.

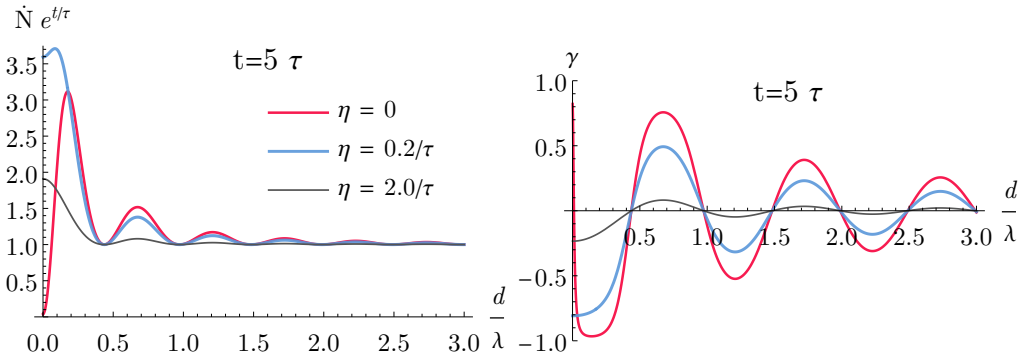

FIGURE 3.31: Emission characteristics: normalized emission rate $\dot{N}$ and degree of mutual coherence γ at the time moment $t = 5\tau$. The red line corresponds to the dephasing-free model $\eta = 0$, the light blue line – weak dephasing $\eta = 0.2/\tau$, the grey line – strong dephasing $\eta = 2.0/\tau$.

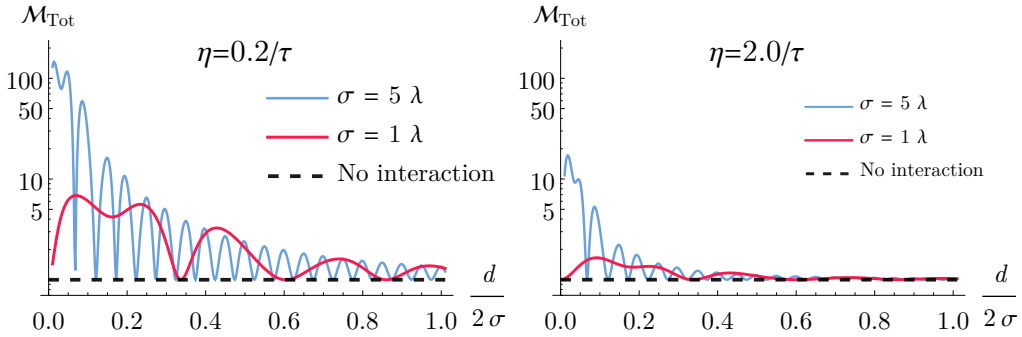

FIGURE 3.32: Normalized total sensitivity $\mathcal{M}_{Tot}$ (3.175) for the model with dephasing. Right inset – weak dephasing $\eta = 0.2/\tau$, left inset – strong dephasing $\eta = 2.0/\tau$. The light blue line — $\sigma = 5\lambda$, the red line — $\sigma = \lambda$, the dashed line — non-interacting dipoles.

Model with detection noise

Another factor that can potentially significantly reduce sensitivity is the presence of detection noise. If the detection system produces dark counts or detects background light, then late rare detection events may be dominated by these false counts. To address this effect, we modify the detection covariance matrix (3.143) as suggested by eq. (3.71)

$$\Gamma^{\text{DC}}_{mn} = \Gamma_{mn} + N_{\text{DC}}\delta_{mn}. \tag{3.176}$$

Here, the dark counts are assumed to have Poisson statistics, being uncorrelated from each other and the measured signal.

Using the modified covariance matrix of noisy detection (3.176), we numerically calculate the sensitivity (2.46), taking into account the separation-dependent emission rate (3.169) and mutual coherence (3.170) of dipoles. We consider detection in the first four Hermite-Gauss modes and express the rate of dark counts relative to

the rate of real counts at the early stages $\dot{N}_0 = 2\kappa$. Note that the rate of real detection events exponentially decreases with time, while the rate of dark counts remains constant.

The sensitivity rate for noisy detection is presented in fig. 3.33. Here we consider strong detection noise, with the dark count rate being 50% of the initial rate of real counts $\dot{N}_{DC} = 0.5\dot{N}_0$. Comparing this plot to the ideal case (fig. 3.28), one can observe a significant decrease in the sensitivity rate for late detection events. After $t = 5\tau$ one obtains almost no information about the separation of dipoles. This decrease occurs because the dark count rate is much higher than the rate of real counts for these late stages. As a result, in this case, most of the useful information comes from the early and intermediate radiation stages ($t < 5\tau$).

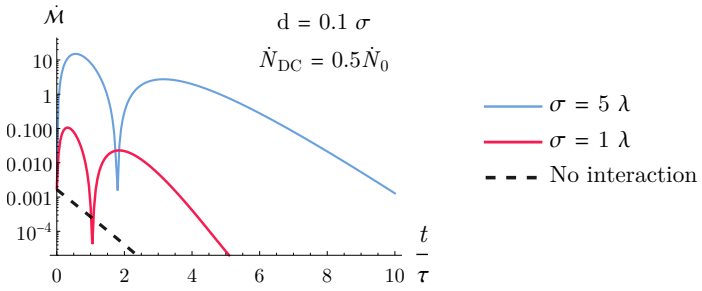

FIGURE 3.33: Normalized sensitivity rate $\dot{\mathcal{M}}$ of two dipoles' separation estimation vs. time t in case of noisy detection $\dot{N}_{DC} = 0.5\dot{N}_0$.

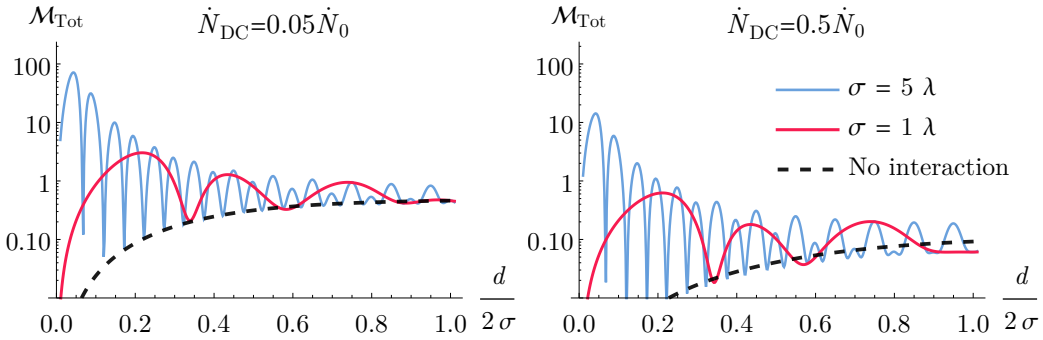

FIGURE 3.34: Normalized total sensitivity $\mathcal{M}_{Tot}$ vs. separation d for the model with detection noise. Left inset — low rate of the dark counts, right inset — intense noise.

Integrating the sensitivity rate over time we obtain the total sensitivity for noisy detection and present it in fig. 3.34. As expected, compared to the ideal case (fig. 3.30), the sensitivity is suppressed in the presence of the detection noise. However, this is also the case for non-interacting dipoles. Thus, one can still see a significant increase in the sensitivity (around three orders of magnitude for small separations) coming from the dipole interaction, even for relatively intense detection noise $\dot{N}_{DC} = 0.5\dot{N}_0$.

Curiously, in all the considered cases the sensitivity boost is more pronounced for smaller separations, where the interaction between the emitters is stronger. As

a result, in the context of single-parameter estimation, one can achieve higher precision in estimating smaller separations compared to larger ones.

Finally, we note that in the case of frequency-resolved measurements, the collective frequency shift f_{12} may come into play, such that one may be able to locally estimate the dipoles' separation via the emission spectrum. Therefore, by integrating spectral and spatial approaches, the sensitivity could be further enhanced by introducing frequency-resolving detection into our scheme.

3.4.5 Conclusion

In this section, we have analyzed the effects of coherence on the problem of resolving point sources with SPADE measurement. We employed the method of moments, which allowed us to make no assumptions about the brightness of the sources. We have found analytical expressions for the sensitivity including the cases of separation-dependent mutual coherence and emission rate. We studied two specific examples of separation-dependent coherence: a reflection of light coming from a finite-coherence-width illumination source and creating mutual coherence due to the interaction of the emitters. In both cases, we demonstrate the possibility of a significant boost in the separation estimation sensitivity due to the additional mechanism of the parameter encoding to the problem. Our analysis shows that for efficient resolving of the reflecting objects one needs to use an illumination source with a narrow coherence width (on the order of the reflectors' separation) and a short coherence time. This ensures the full advantage of the separation-dependent coherence of the reflected light. Examining the interacting emitters, we demonstrate that the sensitivity of separation estimation can be increased by several orders of magnitude compared to independent emitters. This enhancement arises from the separation-dependent mutual coherence and emission rates of the interacting dipoles. The effect remains robust in the presence of weak dephasing of dipoles. Although strong dephasing decreases the observed resolution boost, it does not eliminate it entirely. We have also demonstrated that the presence of detection noise reduces sensitivity but preserves the relative boost caused by the interaction of dipoles.

Chapter 4

Gaussian states characterization

Contents

In this chapter, we employ the method of moments for the characterization of quantum states. Specifically, we focus on a simple, but commonly encountered practical scenario —- the characterization of single-mode Gaussian states. As discussed in section 1.3.3, states in this class are described by a pair of mean quadratures and three parameters of the quadrature covariance matrix. Thus, following the approach introduced in section 2.4.4, we undertake the characterization of such states using parameter estimation techniques, specifically employing the method of moments.

Within the scope of this chapter, we introduce a straightforward moment-based strategy that efficiently utilizes homodyne (and heterodyne) detection data to estimate the parameters of a Gaussian state. This approach does not necessitate complex optimization operations or heavy computations, relying instead on simple algebraic transformations. Consequently, it can be applied for the dynamic characterization of sources with drifting parameters, where computational complexity in data processing is critical and only a limited number of measurements are available.

4.1 Motivation

We have previously discussed the complexity and significance of the quantum state characterization problem in section 2.4. In this chapter, we focus on characterizing the state from a specific class — Gaussian squeezed states (as introduced in section 1.3.1). These states are arguably the most readily accessible non-classical states, having been utilized in numerous quantum optics laboratories for over 35 years [Andersen 2016]. They found their application in various branches of quantum metrology [Dodonov 2002; Schnabel 2017; Lawrie 2019] ranging from phase estimation [Bondurant 1984; Combes 2004; Pezze 2014; Yu 2020], including gravitational wave detection [Grote 2013; Barsotti 2018; Tse 2019; Virgo Collaboration 2019; McCuller 2020], to spatial [Kolobov 1993; Soh 2023] and temporal [Patera 2019] imaging, displacement measurement [Treps 2003; Treps 2004], clock synchronization [Giovannetti 2001], etc. They are also known to be valuable for quantum communication [Gottesman 2003; Usenko 2018; Laudenbach 2018; Derkach 2020; Pirandola 2020] quantum teleportation [Furusawa 1998; Bowen 2003] and other quantum information protocols [Braunstein 2005; Yonezawa 2010; Weedbrook 2012].

For certain applications, knowing the precise characteristics of states generated by a particular source can be crucial. These characteristics can often change over time due to fluctuations in temperature and other properties of the source. Therefore, it is essential for the characterization procedure to be fast enough to be repeated frequently (potentially to adjust the parameters of the source via a feedback loop). We will show in sections 4.4.1 and 4.4.2, that traditional methods for characterizing squeezed states use the measured data in an inefficient way. As a result, more measurements are required to accurately characterize the state, leading to longer measurement time (often on the order of seconds [Fainsin 2023; Kouadou 2023; Roman-Rodriguez 2024]). One approach to address this is to frame the problem as parameter estimation and use universal optimal estimators (like MLE) to exploit the full potential of the measured data. However, the data processing complexity can be significant in this case, and longer computation times may offset any gains from reduced acquisition time. Thus, in this chapter, we develop a straightforward moment-based estimator that effectively utilizes homodyne detection data to estimate the parameters of the single-mode Gaussian state. Our fast estimator, based on simple algebraic transformation, has the potential to dynamically characterize sources with varying parameters, enabling feedback control for stabilization of such sources.

4.2 Measurement statistics of the squeezed Gaussian states

Any Gaussian state is fully characterized by the mean-field vector $\vec{Q}$ and the quadrature covariance matrix $\mathbf{\Gamma}_Q$ (1.108). In the case of a single-mode state, the decomposition (1.110) of the quadrature covariance matrix takes the form

$$\mathbf{\Gamma}_Q = \begin{pmatrix} \cos\phi_s & \sin\phi_s \\ -\sin\phi_s & \cos\phi_s \end{pmatrix} \begin{pmatrix} \kappa & 0 \\ 0 & \kappa \end{pmatrix} \begin{pmatrix} s & 0 \\ 0 & s^{-1} \end{pmatrix} \begin{pmatrix} \cos\phi_s & -\sin\phi_s \\ \sin\phi_s & \cos\phi_s \end{pmatrix}. \tag{4.1}$$

The physical sense of the parameters $\vec{\theta} = (s, \kappa, \phi_s)^T$, defining this matrix, is discussed below.

In this case, the probability density of obtaining measurement result q_ψ when measuring the generalized quadrature $\hat{q}_\psi$ (1.84) (i.e. the phase of the LO equals ψ), given the mean-field vector $\vec{Q} = (\langle q \rangle, \langle p \rangle)^T$ and quadrature covariance matrix defined by parameters $\vec{\theta} = (s, \kappa, \phi_s)^T$, follows the Gaussian distribution

$$P(q_\psi | \psi, \vec{Q}, \vec{\theta}) = \mathcal{N}\Big(Q_\psi, X(\psi, \vec{\theta})\Big), \tag{4.2}$$

where the explicit probability density is defined in eq. (2.66). The mean value Q_ψ is given by

$$Q_\psi = (\cos\psi, \sin\psi) \cdot \vec{Q} \tag{4.3}$$

and the variance

$$\begin{aligned} X(\psi, \vec{\theta}) = \Delta^2 q_\psi &= \kappa \left(s \cos^2[\psi - \phi_s] + \frac{\sin^2[\psi - \phi_s]}{s} \right) \\ &= \frac{\kappa}{2s}\Big(1 + s^2 - (1 - s^2)\ \cos[2(\psi - \phi_s)]\Big). \end{aligned} \tag{4.4}$$

Here, $\kappa \geq 1$ defines the thermal part of the state (the purity of the state is $P = 1/\kappa$), ϕ_s defines the squeezed quadrature, and s is responsible for the amount of squeezing. To remove the degeneracy, we set the parameters in such a way that $0 < s \leq 1$ and $0 \leq \phi_s < \pi$. Consequently, the quadrature with the smallest variance

$$\min_\psi \Delta^2 q_\psi = \kappa s \tag{4.5}$$

corresponds to $\psi = \phi_s$. We term it the *squeezed quadrature*, although its noise can in general exceed the vacuum noise. The amount of squeezing below the vacuum noise is often expressed in Decibel (dB) as

$$L_s = -10 \log_{10}[\kappa s] \text{ dB}, \tag{4.6}$$

then 50% of vacuum noise corresponds to 3 dB of squeezing, and 10% — to 10 dB. States with levels of squeezing up to 15 dB have already been demonstrated experimentally [Vahlbruch 2016; Andersen 2016]. The antisqueezed quadrature $\psi = \phi_s + \pi/2$ has a variance of κ/s, which is always not less than 1.

One can sample the quadrature $\hat{q}_\psi$ (1.84) using homodyne detection with the phase of the LO ψ. In section 4.6 we discuss the requirements for the measurement apparatus.

In what follows, we assume that the studied state does not contain any mean-field

$$\vec{Q} = 0, \tag{4.7}$$

meaning we consider non-displaced squeezed thermal states. Thus, the parameters of the state to be estimated are the parameters of the quadrature covariance matrix $\vec{\theta} = (s, \kappa, \phi_s)^T$. If the mean-field is present, it can be readily estimated from the first moments of the measured quadratures and subtracted from the measured data

during post-processing, effectively displacing the state to a zero mean-field configuration. This process, however, influences the estimation precision of the covariance matrix parameters, as demonstrated in [Řeháček 2015; Teo 2017]. Nevertheless, we leave the problem of the mean-field beyond the scope of this research and instead concentrate on constructing efficient estimators for the parameters $\vec{\theta} = (s, \kappa, \phi_s)^T$ of the covariance matrix in zero mean-field configuration.

4.3 Sensitivity of state characterization

Now, let us examine the precision bounds for estimating the parameters $\vec{\theta} = (s, \kappa, \phi_s)^T$. We begin by exploring the sensitivity of the moment-based approach.

4.3.1 The method of moments

The smallest non-trivial statistical moment of the quadrature operator $\hat{q}_\psi$ is the second moment $\hat{X}_\psi = \hat{q}_\psi^2$, thus we construct the MoM based on it. The expected value of the operator $\hat{X}_\psi$ is

$$\langle \hat{X}_\psi \rangle_{\vec{\theta}} = \langle \hat{q}_\psi^2 \rangle_{\vec{\theta}} = X(\psi, \vec{\theta}) \tag{4.8}$$

and its variance is

$$\Delta^2 X_\psi = \langle \hat{q}_\psi^4 \rangle_{\vec{\theta}} - \langle \hat{q}_\psi^2 \rangle_{\vec{\theta}}^2 = 2X^2(\psi, \vec{\theta}), \tag{4.9}$$

where the final equality follows from the Gaussian statistics of $\hat{q}_\psi$ (4.2), and $X(\psi, \vec{\theta})$ is given by eq. (4.4).

The moment matrix (2.54) of the measurement of a single quadrature $\hat{q}_\psi$ is given by

$$M_{\alpha\beta}(\psi, \vec{\theta}) = \frac{1}{2X^2(\psi, \vec{\theta})} \frac{\partial X(\psi, \vec{\theta})}{\partial \theta_\alpha} \frac{\partial X(\psi, \vec{\theta})}{\partial \theta_\beta}. \tag{4.10}$$

This matrix is a direct product of two vectors, i.e. it has rank 1, meaning one cannot estimate all three parameters $\vec{\theta} = (s, \kappa, \phi_s)^T$ by measuring only one quadrature $\hat{q}_\psi$. Thus, we need to change the phase of the LO and measure various quadratures. For this purpose, we choose the set of LO phases $\{\psi_j\}, j = 1...N_\psi$.

The minimal set of phases required to estimate all three parameters of the quadrature covariance matrix is $N_\psi = 3$. Typically, quadratures with phases $\psi = 0, \pi/4$, and $\pi/2$ are measured for this purpose [DAuria 2009]. Subsequently, the parameters $\vec{\theta} = (s, \kappa, \phi_s)^T$ are determined from the variances of these three quadratures through simple algebraic transformations. However, this approach has a drawback: it necessitates precise control over the phase of the LO. It must remain constant during the measurement of each quadrature and then be rapidly adjusted to a new value, which can be hard to achieve in practice.

In practical scenarios, the phase of the LO is frequently continuously "scanned". Then each new measurement corresponds to a slightly different quadrature. Assuming a large number of measurements N_ψ, we can describe the set of LO phases continuously with a density function $p(\psi)$. As quadrature measurements with different ψ are performed over distinct probes, they are independent, and the moment

matrix is additive. In other words, the moment matrix of all the measurements can be expressed as:

$$M_{\alpha\beta} = \sum_{j=1}^{N_\psi} M_{\alpha\beta}(\psi_j, \vec{\theta}) = N_\psi \int_0^{2\pi} p(\psi) M_{\alpha\beta}(\psi, \vec{\theta})\, d\psi. \tag{4.11}$$

Having some preliminary information about the angle of squeezing ϕ_s one can optimize the choice of the measured quadratures $p(\psi)$ to increase the sensitivity. However, here we consider the most general and the most simple case, where we do not have (or do not use) any preliminary information about the parameters while making the measurement. Thus, conveniently, the measurement procedure will be always the same. In this case, the best choice of the LO phases has a uniform distribution

$$p(\psi) = \frac{1}{n\pi},\ 0 \le \psi < n\pi,\ n \in \mathbb{N}. \tag{4.12}$$

Consequently, the moment matrix (4.11) for estimating the parameters $\vec{\theta} = (s, \kappa, \phi_s)^T$ reads

$$\boldsymbol{M} = N_\psi \begin{pmatrix} \frac{1+s^2}{2s^2(1+s)^2} & \frac{-(1-s)}{2\kappa s(1+s)} & 0 \\ \frac{-(1-s)}{2\kappa s(1+s)} & \frac{1}{2\kappa^2} & 0 \\ 0 & 0 & \frac{(1-s)^2}{s} \end{pmatrix}. \tag{4.13}$$

One can notice, that this matrix is independent of ϕ_s. This could be expected since many quadratures $\hat{q}_\psi$, $0 \le \psi < n\pi$ were measured, and the absolute direction of squeezing ϕ_s only shifts the results of measurements along ψ. Note that the matrix (4.13) corresponds to the measurement of N_ψ quadratures, to calculate the average moment matrix per measurement one can divide $\boldsymbol{M}$ by N_ψ.

The inverse of the matrix (4.13) sets the bound for the estimator covariance (2.53), and can be found to be

$$\boldsymbol{M}^{-1} = \frac{1}{N_\psi} \begin{pmatrix} s(1+s)^2 & \kappa(1-s^2) & 0 \\ \kappa(1-s^2) & \kappa^2 \frac{1+s^2}{s} & 0 \\ 0 & 0 & \frac{s}{(1-s)^2} \end{pmatrix}. \tag{4.14}$$

The sensitivities S_α to the parameter θ_α are defined by the inverse diagonal elements of the matrix $\boldsymbol{M}^{-1}$, as given by eq. (2.56):

$$\Delta^2 \tilde{\theta}_\alpha \ge \frac{1}{\mu S_\alpha} = \frac{1}{\mu} \left[\boldsymbol{M}^{-1}\right]_{\alpha\alpha}, \tag{4.15}$$

where μ denotes the number of repeated measurements of each quadrature $\hat{q}_{\psi_j}$.

4.3.2 Fisher information

To compare the efficiency of the moment-based approach to a more general parameter estimation strategies, such as the maximum likelihood estimation, we also calculate the FI of the quadrature measurement. The FI matrix in the case of Gaussian measurement statistics (4.2) is given by eq. (2.67). If only one observable $\hat{q}_\psi$ is measured, while $\partial\langle\hat{q}_\psi\rangle_{\vec{\theta}}/\partial\theta_\alpha = 0$ and $\Delta^2 q_\psi = X(\psi,\vec{\theta})$, the FI matrix takes the form

$$\mathcal{F}_{\alpha\beta}(\psi,\vec{\theta}) = \frac{1}{2}X^{-1}(\psi,\vec{\theta})\frac{\partial X(\psi,\vec{\theta})}{\partial\theta_\alpha}X^{-1}(\psi,\vec{\theta})\frac{\partial X(\psi,\vec{\theta})}{\partial\theta_\beta} = M_{\alpha\beta}(\psi,\vec{\theta}). \tag{4.16}$$

Thus, in this case, the FI matrix equals the moment matrix (4.10). This could be expected, due to the fact, that the first two moments of Gaussian distribution contain all the information about its parameters. This equality also holds for a set of measurements of different quadratures $\hat{q}_{\psi_j}$ since both the FI and moment matrices are additive. Therefore, the CRB (2.22) for estimation of the parameters $\vec{\theta}$ totally coincides with MoM bound (2.53), i.e. the second moments of quadratures contain all the information about the parameters (s,κ,ϕ_s). Thus the estimator based solely on the second moment is optimal, it does not lose any relevant information. At the same time, the computational complexity of the moment-based approach is significantly smaller, than the complexity of the MLE, which is usually used to saturate the CRB.

4.3.3 Quantum Fisher information

To benchmark the measurement scheme used we calculate the QFI matrix for estimating the parameters $\vec{\theta} = (s,\kappa,\phi_s)^T$. Since we consider single-mode Gaussian states, the QFI can be calculated with eq. (2.110) giving the diagonal QFI matrix

$$\boldsymbol{\mathcal{F}}^{\mathrm{Q}} = \mathrm{diag}\left(\frac{1}{s^2}\frac{\kappa^2}{\kappa^2+1}, \frac{1}{\kappa^2-1}, \frac{(1-s^2)^2}{s^2}\frac{\kappa^2}{\kappa^2+1}\right). \tag{4.17}$$

As discussed in section 2.3.4 the diagonality of the QFI matrix indicates that the QCRB (2.103) is saturable, i.e. measurements that saturate the inequality are compatible with each other.

The inverse of diagonal QFI matrix (4.17) is easy to find as

$$\left(\boldsymbol{\mathcal{F}}^{\mathrm{Q}}\right)^{-1} = \mathrm{diag}\left(s^2\frac{\kappa^2+1}{\kappa^2}, \kappa^2-1, \frac{s^2}{(1-s^2)^2}\frac{\kappa^2+1}{\kappa^2}\right). \tag{4.18}$$

Comparing it to MoM sensitivity eq. (4.14) normalized per measurement, it is straightforward to show that

$$\left[\left(\boldsymbol{\mathcal{F}}^{\mathrm{Q}}\right)^{-1}\right]_{\alpha\alpha} < N_\psi\left[\boldsymbol{M}^{-1}\right]_{\alpha\alpha} = N_\psi\left[\boldsymbol{\mathcal{F}}^{-1}\right]_{\alpha\alpha}, \tag{4.19}$$

for any $\kappa \geq 1$ and $0 < s \leq 1$. This implies that measuring a uniform set of quadratures is never the quantum optimal approach to the characterization of Gaussian states. However, it is important to note that optimal measurements are generally

challenging to implement experimentally, and they are only effective within a limited range of parameter values. In contrast, the measurement of the uniform set of quadratures is straightforward to implement in experiments and remains universally applicable across all parameter values. By choosing the uniform distribution for the LO phase ψ we unavoidably measure some quadratures, that are more sensitive to the parameters, and some that are less so. Thus, we pay the price of a reduction in the average sensitivity and, in return, get a universally applicable method.

4.4 Estimators

The next step involves studying different estimators for the parameters $\vec{\theta} = (s, \kappa, \phi_s)^T$. We consider the estimators based on a large set of measured quadratures with phases $\{\psi_j\}$, $j = 1, ..., N_\psi$ evenly distributed within a 4π range[1], and $N_\psi \gtrsim 10^2$. In the general case, each quadrature $\hat{q}_j = \hat{q}_{\psi_j}$ is measured μ times, leading to a total number of measured samples $N = \mu N_\psi$. However, since we consider the continuous scan of the LO phase, we assume obtaining only one measurement sample per quadrature, i.e., $\mu = 1$ and $N = N_\psi$. We denote the outcome of measuring $\hat{q}_j$ as q_j, and samples of the observables $\hat{X}_j = \hat{q}_j^2$ are, accordingly, $x_j = q_j^2$.

Typical homodyne data for a squeezed thermal state are illustrated in fig. 4.1 (obtained through numerical simulation). Each point represents a single sample of the quadrature, with the left plot depicting $N = 10^4$ samples, and the right plot showing $N = 10^6$ samples. The parameter values used for the simulations are $(s, \kappa, \phi_s) = (0.3, 1.2, 2.5)$. These plots demonstrate variations in measurement noise across different phases of the LO. In the upcoming sections, we will discuss the strategies for inferring specific values of the parameters $\vec{\theta}$ from this data.

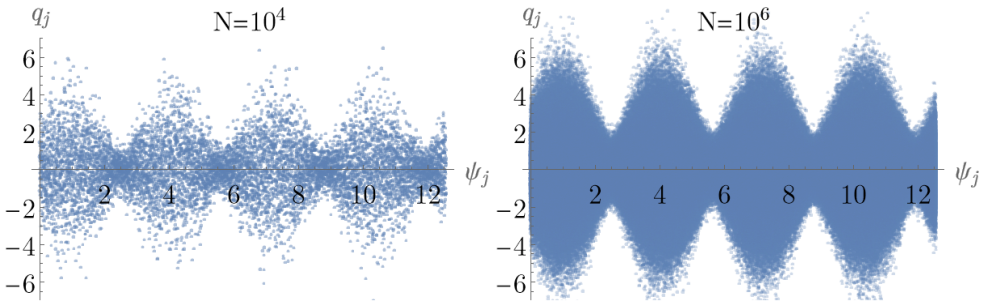

FIGURE 4.1: Simulated results of homodyne measurement q_j with $N = 10^4$ samples and $N = 10^6$ samples. True values of the parameters $(s, \kappa, \phi_s) = (0.3,\ 1.2,\ 2.5)$.

4.4.1 Minimal and maximal variance

The standard practical approach for estimating the squeezing of the state involves dividing the measured data into bins, corresponding to $\delta\psi$ range of the phase and

[1]The range of 4π is chosen for aesthetic reasons in the plots; it could be any integer multiple of π.

calculating the sample variances of these bins

$$\overline{\Delta^2 q_\psi} = \mathrm{Var}\{q_{\psi_j}\}, \text{with } \psi_j \in [\psi - \delta\psi/2, \psi + \delta\psi/2]. \tag{4.20}$$

The sample variances calculated for the data from fig. 4.1 are shown in fig. 4.2. Subsequently, the minimum value of the sample variance corresponds to the squeezed quadrature κs, while the maximum value corresponds to the antisqueezed quadrature κ/s. The positions of the minima and maxima define the angle of squeezing ϕ_s. Most often in practice, the plots 4.2 are directly experimentally obtained with the spectrum analyzer, which automatically bins the data and calculates the variance [Sansavini 2023; Kaiser 2016; Fainsin 2023; Roman-Rodriguez 2024].

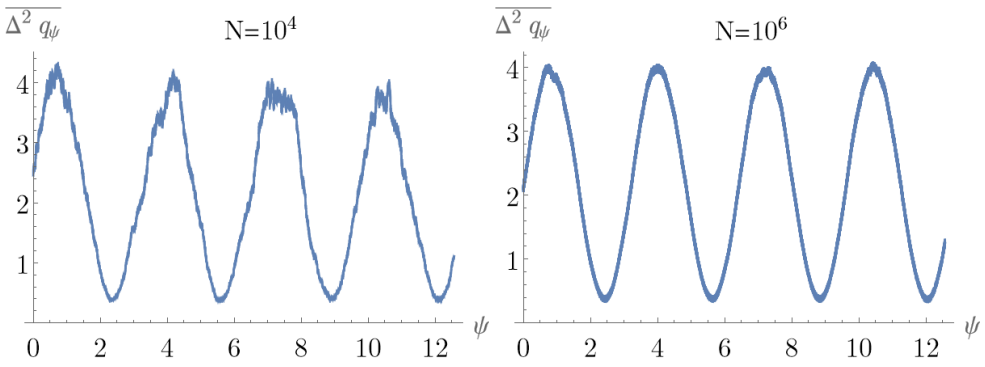

FIGURE 4.2: Sample variance of the homodyne measurement, obtained by binning the data from fig. 4.1. 10^3 (left plot) and 10^5 (right plot) samples are divided into bins $\delta\psi = 4\pi/50$.

This approach gives a simple and useful estimator, that does not rely on precise control[2] of the LO phase ψ for estimating the parameters κ and s. However, this estimator clearly underutilizes the information available in the measurement results. In this scenario, the estimators rely solely on the data bins, corresponding to minimal and maximal variances, while disregarding the remaining data. Consequently, this estimation approach yields noisy results if the data bins contain an insufficient amount of samples. Simultaneously, enlarging the size of the bins $\delta\psi$ results in increased bias of the estimator, since larger bins are averaging the variance of different quadratures and plots 4.2 become smoother.

4.4.2 Least squares method and Fourier estimator

A more developed approach involves fitting the plots 4.2 by the function $X(\psi, \vec{\theta})$ (4.4). Furthermore, to decrease the bias of the estimation it makes sense to abandon binning and fit the squared quadrature samples $x_j = q_j^2$ directly by their expected values function $X(\psi, \vec{\theta})$ as given by eq. (4.8). This fitting can be done using the classical *least squares* (LS) method, which minimizes the square difference between

[2]It is only necessary to change the phase ψ of the LO within the range of at least π, but precise control of it is not required.

data samples and the expected values of the statistical model [Kay 1998]:

$$\vec{\tilde{\theta}}^{LS} = \arg \min_{\vec{\theta}} \sum_{j=1}^{N_\psi} (q_j^2 - X(\psi_j, \vec{\theta}))^2. \tag{4.21}$$

To perform the minimization analytically we choose another parametrization of the function $X(\psi, \vec{\theta})$ (4.4):

$$X(\psi, \vec{C}) = C_0 + C_2 e^{2i\psi} + C_2^* e^{-2i\psi}, \tag{4.22}$$

where

$$C_0 = \kappa \frac{1+s^2}{2s}, C_2 = -\kappa \frac{1-s^2}{4s} e^{-2i\phi_s}. \tag{4.23}$$

Representation (4.22) corresponds to the complex Fourier series of $X(\psi_j, \vec{C})$. It is straightforward to perform analytical minimization (4.21) for the parameters $\vec{C}$:

$$\frac{\partial}{\partial C_0} \sum_{j=1}^{N_\psi} (q_j^2 - X(\psi_j, \vec{\theta}))^2 = -\sum_{j=1}^{N_\psi} (q_j^2 - X(\psi_j, \vec{\theta})) = 0 \tag{4.24}$$

which yields

$$\tilde{C}_0^{LS} = \frac{1}{N_j} \sum_{j=1}^{N_\psi} q_j^2, \tag{4.25}$$

i.e. LS estimator of C_0 is given by zero Fourier component of the dataset $\{q_j^2\}$. Likewise, one can show that the same is true for the second Fourier component [Kay 1998]:

$$\tilde{C}_2^{LS} = \frac{1}{N_j} \sum_{j=1}^{N_\psi} q_j^2 \, e^{-2i\psi_j}. \tag{4.26}$$

One can construct an estimator for parameters θ by inverting the parameter transformation (4.23):

$$\tilde{s}^F = \sqrt{\frac{\tilde{C}_0 - 2|\tilde{C}_2|}{\tilde{C}_0 + 2|\tilde{C}_2|}}, \quad \tilde{\kappa}^F = \sqrt{\frac{\tilde{C}_0 - 2|\tilde{C}_2|}{(\tilde{C}_0 + 2|\tilde{C}_2|)^{-1}}}, \quad \tilde{\phi}_s^F = -\frac{1}{2} \operatorname{Arg} \tilde{C}_2. \tag{4.27}$$

This intuitive estimator can be developed without using the LS approach, instead using straightforward Fourier analysis of the plots 4.2. That is why we use the shorter notation $\vec{\tilde{\theta}}^F = \vec{\tilde{\theta}}^{LS}$.

Note that due to the statistical noise, the expressions under the square roots in eq. (4.27) can occasionally yield negative values $\tilde{C}_0 - 2|\tilde{C}_2| < 0$. This indicates a negative estimation of squared parameters s^2 and κ^2, signifying that the estimator fails to produce physically valid parameter values. The frequency of these outcomes is plotted in the fig. 4.3 for datasets of $N = 10^4$ and $N = 10^6$ samples.

In the event of the LS estimator failure, one can choose to either discard the corresponding measurement dataset and repeat the measurement or utilize the minimum values of the parameters s and κ, estimated from certain physical characteristics of

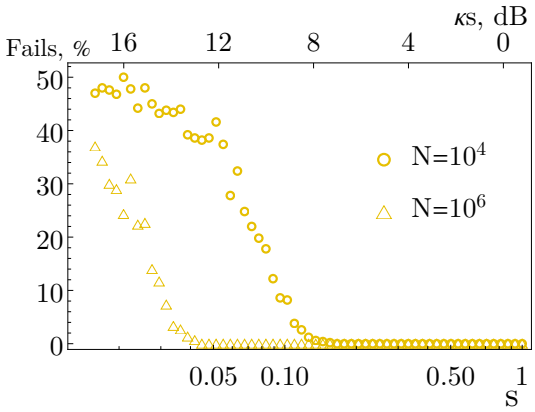

FIGURE 4.3: Rate of LS approach failures. Each point is obtained with 500 simulated datasets of $N = 10^4$ (circles) and $N = 10^6$ (triangles) samples with $\kappa = 1.2$.

the source, such as losses within the system (we adopt $s \geq 10^{-3}$ and $\kappa \geq 1 + 10^{-4}$). However, both approaches introduce bias into the estimators, and a high failure rate indicates inadequate statistical data to perform an estimation with the selected estimator. We present the variance and the bias of the LS estimator $\vec{\tilde{\theta}}^F$ in fig. 4.4.

From the plots in fig. 4.4 one can see, that LS estimator is close to the CRB for $s > 0.5$ (i.e. squeezing below 3 dB) for all of the parameters. For stronger squeezing (smaller s) the variance of this estimator does not saturate the CRB. With further increase in squeezing, variance continues growing and estimators $\tilde{s}^F$ and $\tilde{\kappa}^f$ become increasingly biased. As long as bias (defined in eq. (2.1)) stays significantly lower than the variance of the estimator, i.e. $B(\theta_\alpha)^2 \ll \Delta^2\theta_\alpha$, the confidence interval for the estimator contains the ground truth value of the parameter θ_α with high probability. Contrary, if $B(\theta_\alpha)^2 \sim \Delta^2\theta_\alpha$, the estimator is significantly biased, and the estimation interval may not contain the true parameter value. One can see that this is the case for LS estimation of parameters s and κ if the squeezing is strong enough, especially for the lower measurement statistics. At the same time, the estimator $\tilde{\phi}^f$ stays unbiased, and its variance does not grow with increasing squeezing, exceeding the CRB by no more than 1-2 orders of magnitude. Further, we use the value of $\tilde{\phi}^f$ estimator as a good zero approximation for the moment-based approach.

It is interesting to note that failures of the LS estimator are not the only sources of bias. For instance, the estimation of a 12 dB squeezed state with $N = 10^6$ samples does not yield imaginary values of the estimator, as shown in fig. 4.3. However, the right plots in fig. 4.4 reveal noticeable bias in the estimators $\tilde{s}^F$ and $\tilde{\kappa}^f$ in this scenario. Hence, increasing the statistics N does not necessarily result in the full elimination of estimation bias, though the bias is significantly decreased for larger N and typically stays below the level of noise.

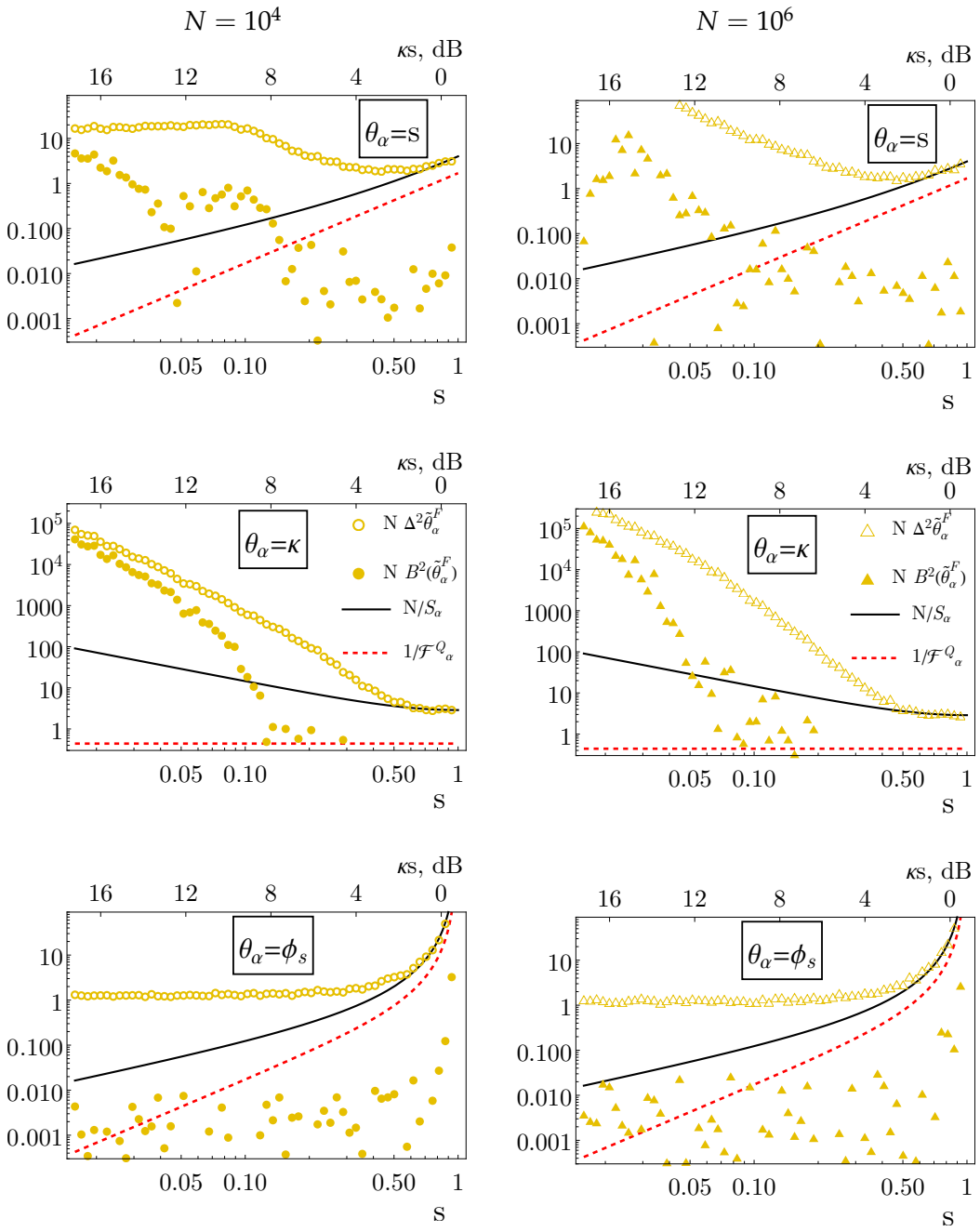

FIGURE 4.4: Performance of the LS (Fourier) estimator. Hollow markers correspond to the variance of the estimators $\Delta^2\tilde{\theta}_\alpha$ multiplied by N, filled markers — to the squared bias $B^2(\tilde{\theta}_\alpha)$, given by eq. (2.1), multiplied by N. Each point is obtained with 500 simulated datasets of $N = 10^4$ (left) and $N = 10^6$ (right) samples with $\kappa = 1.2$. Black lines correspond to the MoM limit per one measurement (4.15) (coincide with the CRB), dashed red line represents the QCRB (4.18).

In general, the LS estimator is optimal and unbiased when data is normally distributed around the mean with constant variance [Kay 1998]. However, neither of these conditions is met in the considered case, which is why optimality and unbiasedness of the LS estimator are not guaranteed, even in the asymptotic limit. A more rigorous and precise approach would involve accounting for different variances corresponding to different quadratures in eq. (4.21), a concept often referred to as *errors-in-variables models*. However, in this scenario, optimization would necessitate numerical methods, which compromises the simplicity of the approach.

Here, we present a couple of reconstruction results, beginning with the scenario depicted in fig. 4.1:

$$\vec{\theta} = \begin{pmatrix} 0.300 \\ 1.200 \\ 0.900 \end{pmatrix}, \quad \tilde{\vec{\theta}}^{F}_{10^4} = \begin{pmatrix} 0.299 \pm 0.015 \\ 1.19 \pm 0.05 \\ 0.900 \pm 0.013 \end{pmatrix}, \quad \sqrt{\frac{1}{S_\alpha}} = \begin{pmatrix} 0.007 \\ 0.023 \\ 0.008 \end{pmatrix}, \tag{4.28}$$

i.e. for the given parameter values, the LS estimator exhibits nearly unbiased behavior, although its standard deviation is approximately twice that of the CRB. With stronger squeezing (9.2 dB), the non-optimality of the LS becomes more pronounced:

$$\vec{\theta} = \begin{pmatrix} 0.100 \\ 1.200 \\ 0.900 \end{pmatrix}, \quad \tilde{\vec{\theta}}^{F}_{10^4} = \begin{pmatrix} 0.09 \pm 0.04 \\ 1.26 \pm 0.27 \\ 0.899 \pm 0.011 \end{pmatrix}, \quad \sqrt{\frac{1}{S_\alpha}} = \begin{pmatrix} 0.004 \\ 0.038 \\ 0.004 \end{pmatrix}, \tag{4.29}$$

providing little to no information about the parameter κ and wide and biased intervals for s. Despite this, LS estimation can still be used for characterizing this state with larger statistics (e.g., $N = 10^6$), although the estimation interval remains 10 times wider than the CRB. For characterizing even stronger squeezing (over 10 dB), the statistics required for LS estimation quickly become impractical.

In conclusion, we note that the LS estimator is not optimal for squeezing stronger than 3 dB, and its variance increases rapidly with the squeezing. Besides, the LS method provides biased estimations for squeezing levels above 7-8 dB, with a growing rate of non-physical estimation outcomes, which requires increased measurement statistics. Therefore, this approach may not be the most suitable choice for characterizing highly squeezed states.

4.4.3 Moment-based estimator

Finally, we consider the moment-based estimator of the parameters $\vec{\theta} = (s, \kappa, \phi_s)^T$, following the approach introduced in section 2.2.2. The estimator $\tilde{\vec{\theta}}$ is based on the following linear combination of the measurement results ($x_j = q_j^2$):

$$y_\alpha(\vec{\theta}_0) = \frac{1}{N_\psi} \sum_{j=1}^{N_\psi} c_\alpha(\psi_j, \vec{\theta}_0)\, x_j, \tag{4.30}$$

with the coefficients (2.58)

$$c_\alpha(\psi_j,\vec{\theta}_0)=\left(\frac{1}{2X^2(\psi_j,\vec{\theta})}\frac{\partial X(\psi_j,\vec{\theta})}{\partial\theta_\alpha}\right)\Big|_{\vec{\theta}=\vec{\theta}_0}, \tag{4.31}$$

where we took into account, that the measurements performed over separate probes are not correlated, and $\vec{\theta}_0$ being the initial guess about the parameters.

It is important to note that since each quadrature $\hat{q}_j$ is measured only once, and observables $\hat{X}_j=\hat{q}_j^2$ do not follow Gaussian statistics, our typical assumptions for the validity of the MoM are not met. However, one can use the Lyapunov condition for the central limit theorem [Ash 2000] to demonstrate that $y_\alpha(\vec{\theta}_0)$ is distributed normally for sufficiently large N_ψ. The Lyapunov condition is satisfied if for some $\delta>0$

$$\lim_{N_\psi\to\infty}\frac{1}{\left(\sum_{j=1}^{N_\psi}2c_\alpha^2(\psi_j,\vec{\theta}_0)X^2(\psi_j,\vec{\theta})\right)^{1+\delta/2}} \times\sum_{j=1}^{N_\psi}\int_q dq\;P(q|\psi_j,\vec{\theta})\;\left|c_\alpha(\psi_j,\vec{\theta}_0)(q^2-X(\psi_j,\vec{\theta}))\right|^{2+\delta}=0, \tag{4.32}$$

where the statistics of quadratures $P(q|\psi_j,\vec{\theta})$ is defined in eq. (4.2). The validity of this limit is straightforward to demonstrate for $\delta=2$, as both sums in eq. (4.32) are proportional to N_ψ, resulting in the overall expression being proportional to N_ψ^{-1}. This is a direct consequence of the coefficients $c_\alpha(\psi_j,\vec{\theta}_0)$ and the moments of the distribution $P(q|\psi_j,\vec{\theta})$ being finite for all j and all physical values of parameters $\vec{\theta}$.

Therefore, if the number of measured quadratures N_ψ is sufficiently large we can apply the MoM to this problem, even in cases where each quadrature is measured only once ($\mu=1$). The MoM constructs the estimators $\vec{\tilde{\theta}}$ as a solution of the system of 3 equations

$$y_\alpha(\vec{\theta}_0)=Y_\alpha(\vec{\theta}_0,\vec{\tilde{\theta}}),\;\alpha=1,2,3, \tag{4.33}$$

where

$$Y_\alpha(\vec{\theta}_0,\vec{\tilde{\theta}})=\frac{1}{N_\psi}\sum_{j=1}^{N_\psi}c_\alpha(\psi_j,\vec{\theta}_0)\;X(\psi_j,\vec{\tilde{\theta}}). \tag{4.34}$$

In the limit of large N_ψ this sum can be replaced with an integral and calculated analytically. Again assuming uniform sampling of the LO phase we obtain

$$\vec{Y}(\vec{\theta}_0,\tilde{\theta})=-\frac{\tilde{\kappa}(1-\tilde{s}^2)}{2\kappa_0(1+s_0)^2\tilde{s}}\begin{pmatrix}\cos 2(\phi_{s_0}-\tilde{\phi}_s)\\ \frac{(1+s_0)^2}{2\kappa_0}\frac{(1+\tilde{s}^2)}{(1-\tilde{s}^2)}-\frac{(1-s_0^2)}{2\kappa_0}\cos 2(\phi_{s_0}-\tilde{\phi}_s)\\ (1-s_0^2)\sin 2(\phi_{s_0}-\tilde{\phi}_s)\end{pmatrix}. \tag{4.35}$$

Now the system (4.33) can be solved analytically, but the solution is quite bulky. Therefore we start with the initial approximation for the phase ψ_s given by the

Fourier analysis $\phi_{s0} = \tilde{\phi}_s^F$ (4.27). Since $\tilde{\phi}_s^F$ is an unbiased estimator, the difference $\phi_{s0} - \tilde{\phi}_s$ is typically small, that can be used for simplifying eq. (4.35). The solution of the simplified system (4.33) reads

$$\begin{pmatrix} \tilde{s} \\ \tilde{\kappa} \\ \tilde{\phi}_s \end{pmatrix} = \begin{pmatrix} \sqrt{\left| \dfrac{y_1(\vec{\theta}_0)s_0(1+s_0) + y_2(\vec{\theta}_0)\kappa_0}{y_1(\vec{\theta}_0)(1+s_0) - y_2(\vec{\theta}_0)\kappa_0} \right|} \\ 2\kappa_0 \sqrt{\left| \dfrac{y_1(\vec{\theta}_0)s_0(1+s_0) + y_2(\vec{\theta}_0)\kappa_0}{(y_1(\vec{\theta}_0)(1+s_0) - y_2(\vec{\theta}_0)\kappa_0)^{-1}} \right|} \\ \phi_{s0} - \dfrac{1}{2}\dfrac{y_3(\vec{\theta}_0)}{y_1(\vec{\theta}_0)(1-s_0^2)} \end{pmatrix}. \tag{4.36}$$

These moment-based estimators are expected to saturate the bound (4.15) if the initial guess is close to the true values of the parameters $\vec{\theta}_0 \approx \vec{\theta}$, i.e. to build an optimal estimator one needs to have a good initial guess $\vec{\theta}_0$. However, this limitation can be easily overcome using an iterative approach: i.e. starting from random initial guess $\vec{\theta}_0 = (s_0, \kappa_0, \phi_{s0} = \tilde{\phi}_{s_F})$, one should use estimated values of the parameters as a new initial guess, to estimate parameters more precisely. The procedure can be repeated a couple of times, however, in most cases, it converges from any random initial guess $\vec{\theta}_0$ within a very few iterations[3].

The variance and bias of the moment-based estimator (4.36) are illustrated in fig. 4.5. Each data point on the plot is derived from 500 simulated datasets of $N = 10^4$ samples with $\kappa = 1.2$. It is apparent that the moment-based estimator is practically unbiased, and its variance saturates the CRB, indicating that it is an optimal estimator.

Repeating the same simulations for dataset sizes ranging from $N = 10^3$ to $N = 10^7$ samples produces analogous results. Even though we observe a minor difference in the bias, the latter consistently remains negligible.

When reconstructing the same state as described in eq. (4.29), we find that the moment-based estimator effectively handles this task:

$$\vec{\theta} = \begin{pmatrix} 0.100 \\ 1.200 \\ 0.900 \end{pmatrix}, \quad \tilde{\vec{\theta}}_{10^4} = \begin{pmatrix} 0.100 \pm 0.004 \\ 1.20 \pm 0.04 \\ 0.900 \pm 0.004 \end{pmatrix}, \quad \sqrt{\frac{1}{S_\alpha}} = \begin{pmatrix} 0.004 \\ 0.038 \\ 0.004 \end{pmatrix}, \tag{4.37}$$

While the computational complexity of the moment-based estimator, and the required number of iterations for the convergence of the scheme are to be more rigorously studied in the future, we want to point out that this estimator is built on simple algebraic transformation without any optimization involved. For large datasets the longest part of the computation is a calculation of the sum (4.30) and the coefficients $c_\alpha(\psi_j, \vec{\theta}_0)$ (4.31). The complexity of this part scales at worst like $O(N_\psi)$ (without proper optimization and caching). With further optimization of the method, like

[3] When using initial values $\vec{\theta}_0$ provided by the Fourier estimator $\tilde{\vec{\theta}}^F$ (4.27), the scheme typically converges within 1-2 iterations.

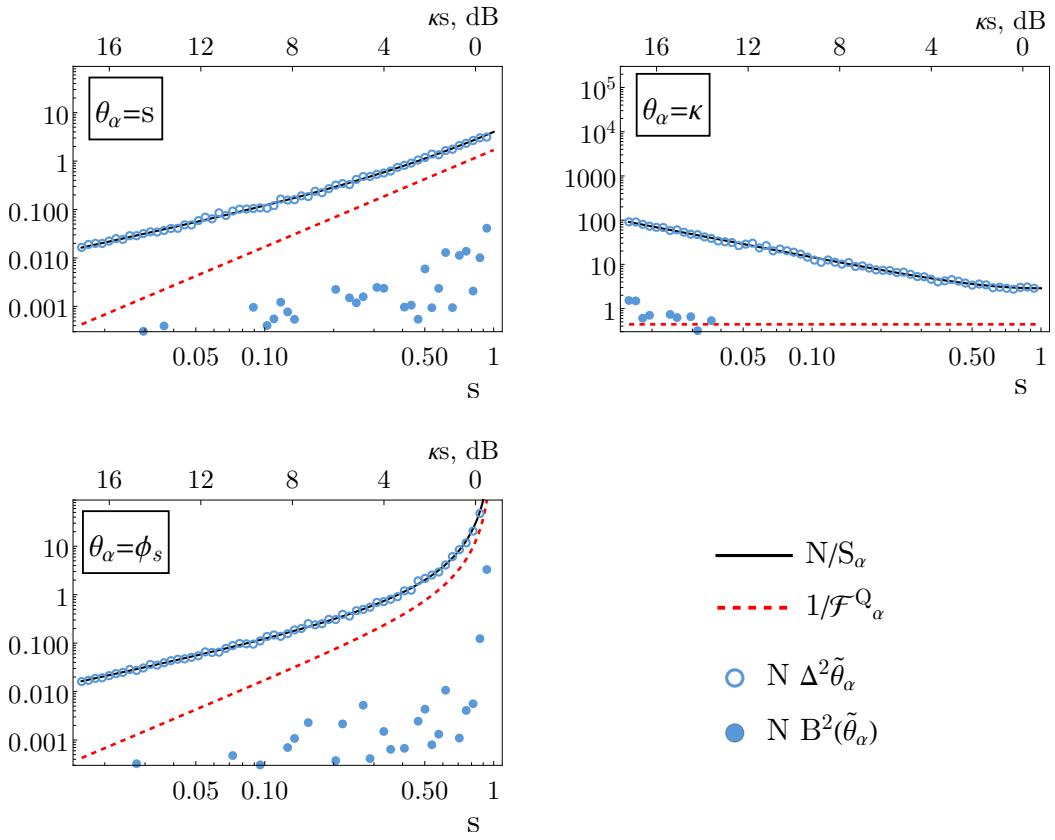

FIGURE 4.5: Performance of the moment-based estimators (4.36). Hollow markers correspond to the variance of the estimators $\Delta^2\tilde{\theta}_\alpha$ multiplied by N, filled markers — to the squared bias $B^2(\tilde{\theta}_\alpha)$ multiplied by N. Each point is obtained with 500 simulated datasets of $N = 10^4$ samples with $\kappa = 1.2$. Black lines correspond to the MoM limit per one measurement (4.15) (coincide with the CRB), dashed red line represents the QCRB (4.18).

data partition, the computational complexity of the proposed approach can scale better than $O(N_\psi)$.

4.4.4 Other estimators

As mentioned earlier, the CRB can always be asymptotically saturated with the MLE (2.25). The challenge with this approach lies in its computational complexity: it involves a 3-parameter optimization of the sum of N terms, which typically necessitates numerous iterations of computing the probability distribution for each quadrature on different values of parameters.

Another approach to characterizing Gaussian states is based on machine learning [Hsieh 2022b; Hsieh 2022a]. While this method can provide rapid and efficient estimators, it also has several disadvantages, as discussed in section 2.4.3. The main drawback is the inability to assess the quality of the results, as neural networks produce estimators without providing error bars on them. The unbiasedness of neural

network estimation is also not guaranteed, and training it with large data samples ($N_\psi \gtrsim 10^4$) can also be challenging.

4.5 Fidelity of the reconstructed state

Using the estimated values of the parameters $\vec{\tilde{\theta}}$, one can construct the covariance matrix $\tilde{\mathbf{\Gamma}}_Q$ (1.110) and the corresponding quantum state. To assess how closely this estimated state matches the real one, we calculate the fidelity as defined in (2.106), computed for a Gaussian state using (2.109). The results of these calculations for simulated data are presented in fig. 4.6, where we plot the infidelity $1 - F$ multiplied by number of measurements N.

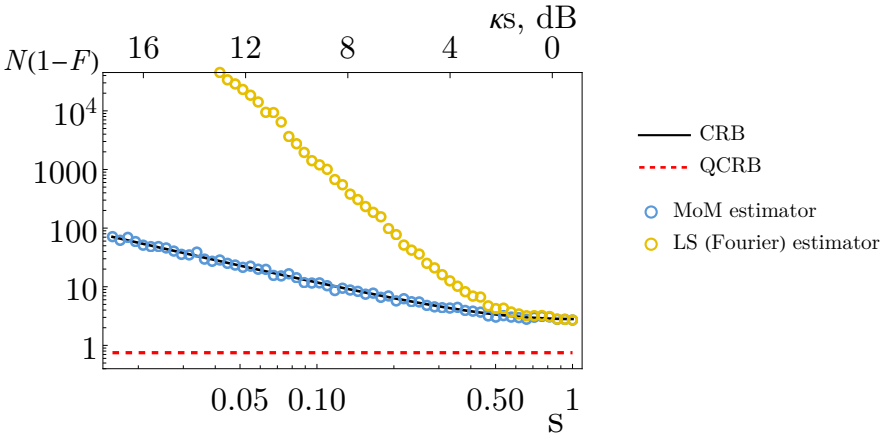

FIGURE 4.6: Infidelity of the reconstructed state. Blue circles correspond to the moment-based estimator (4.36), yellow — to the LS (4.27) (averaged from 500 simulated datasets with $N = 10^6$, $\kappa = 1.2$). The solid black line corresponds to the CRB (4.40), and the dashed red line corresponds to the quantum limit (4.39).

When the fidelity is calculated on two close sets of parameters $\vec{\theta}$ and $\vec{\tilde{\theta}}$, with the latter one being an unbiased estimator for $\vec{\theta}$, the expected fidelity can be decomposed into the Taylor series around $\vec{\tilde{\theta}} = \vec{\theta}$, resulting in

$$F \approx 1 + \frac{1}{2}\sum_{\alpha,\beta} \left.\frac{\partial^2 F(\hat{\rho}_{\vec{\theta}}, \hat{\rho}_{\vec{\theta}+\delta\vec{\theta}})}{\partial(\delta\theta_\alpha)\partial(\delta\theta_l)}\right|_{\delta\vec{\theta}=0} \mathrm{cov}(\tilde{\theta}_\alpha, \tilde{\theta}_\beta). \tag{4.38}$$

One can mention that the derivatives in this equation are proportional to the elements of the QFI matrix (2.108). Therefore the quantum bound for fidelity is given by

$$F^Q = 1 - \frac{1}{\mu}\frac{3}{4}, \tag{4.39}$$

where we took into account that QCRB is saturable due to the diagonality of the QFI matrix (4.17). This bound is depicted by a dashed red line in fig. 4.6.

For the estimator based on the homodyne measurement results, its covariance is limited by the inverse moment matrix (4.14), and minimal fidelity takes the form

$$F = 1 - \frac{1}{N}\frac{\kappa^4(3+4s+3s^2) - \kappa^2(1+4s+s^2)}{4s(\kappa^4-1)}, \tag{4.40}$$

that is plotted in black in fig. 4.6. To achieve the state reconstruction fidelity of 99 % it is enough to use $10^3 - 10^4$ homodyne samples, combined with the moment-based estimator. While using the LS estimator may require $3-4$ orders of magnitude larger statistics to obtain the same reconstruction fidelity.

4.6 Requirements for experimental implementation

In our theoretical description, we modeled homodyne detection without considering losses or additional noise from the electronics or the LO. In real experiments, these effects are inevitably present. However, even with these imperfections, the measurement statistics remain Gaussian with modified parameters κ and s. Therefore, one can always estimate the parameters of the state assuming the detection system to be ideal, and any detection imperfections can be addressed later by adjusting the estimated parameters accordingly.

Additionally, our analysis was conducted in a single-mode setting. Therefore, to attain the calculated sensitivity (4.15) in practice, the homodyne detection should access individual temporal modes of the source. This implies that the integration time of the homodyne detector τ_{int} should match the coherence time τ_c or pulse rate[4] of the source. In this scenario, the detection system directly samples the quadratures of the field without averaging over several temporal modes.

In cases where homodyne detection fails to resolve individual temporal modes of the sources ($\tau_{\text{int}} > \tau_c$), the measurement statistics (4.2) can be modified by replacing the measurement variance $X(\psi,\vec{\theta})$ with $\frac{\tau_c}{\tau_{\text{int}}}X(\psi,\vec{\theta})$, accounting for the averaging property of Gaussian random variables. The constant factor $\frac{\tau_c}{\tau_{\text{int}}}$ does not alter the moment matrix (4.10), which defines the sensitivity of a single measurement. Thus, eqs. (4.13) and (4.14), and the parameter estimation sensitivity $\vec{S}$ (4.15) remain unchanged. It's important to note that with a slower detector, one requires $\frac{\tau_{\text{int}}}{\tau_c}N_\psi$ samples of the states to accumulate N_ψ measurement samples. Consequently, the total measurement time needed to achieve the same sensitivity with slow homodyne detection is $\frac{\tau_{\text{int}}}{\tau_c}$ times longer than with a fast detector ($\tau_{\text{int}} = \tau_c$).

Another crucial requirement for the measurement scheme is precise control of the LO phase[5]. Since the developed moment-based algorithm uses the LO phase ψ_j to calculate the weights $c_\alpha(\psi_j,\vec{\theta}_0)$ for each data sample x_j, accurate knowledge of the phase ψ_j for each homodyne sample is necessary. Furthermore, the direction of squeezing ϕ_s should remain constant throughout the data acquisition process. Therefore, it is essential that the source and the LO are either phase-locked or the measurement is conducted fast enough to prevent significant drift in the phase ϕ_s. The latter seems feasible, as obtaining $10^3 - 10^5$ homodyne samples, required for

[4] If the source of the states is pulsed.

[5] The same requirement applies to LS and MLE estimators.

high-quality estimation of the state parameters, can typically be achieved within $10^{-5} - 10^{-3}$ seconds (for a 100 MHz pulse repetition rate), falling within the stability time of most sources of squeezed light [Andersen 2016].

4.7 Heterodyne measurement

We also briefly examine another measurement commonly used for state characterization — a heterodyne (or double homodyne) measurement introduced in section 1.3.4. This measurement corresponds to the detection of two observables: $\hat{q}'_1 = \hat{q} + \hat{q}_v$ and $\hat{p}'_2 = \hat{p} - \hat{p}_v$, where $\hat{q}_v$ and $\hat{p}_v$ are quadrature operators of the vacuum mode. The first moments of these observables are zero, so the method of moments is built based on the second moments, i.e.

$$\vec{\hat{X}} = \left(\hat{q}'^2_1,\ \hat{p}'^2_2,\ \hat{q}'_1\hat{p}'_2\right)^T \tag{4.41}$$

and

$$\langle\vec{\hat{X}}\rangle_{\vec{\theta}} = \begin{pmatrix} \langle\hat{q}'^2_1\rangle_{\vec{\theta}} \\ \langle\hat{p}'^2_2\rangle_{\vec{\theta}} \\ \langle\hat{q}'_1\hat{p}'_2\rangle_{\vec{\theta}} \end{pmatrix} = \begin{pmatrix} [\mathbf{\Gamma}_Q]_{11} + 1 \\ [\mathbf{\Gamma}_Q]_{22} + 1 \\ [\mathbf{\Gamma}_Q]_{12} \end{pmatrix}, \tag{4.42}$$

whith quadrature covariance matrix $\mathbf{\Gamma}_Q$ defined in eq. (4.1). Using the properties of Gaussian states one can show that the covariance matrix between the second moments is given by

$$\mathbf{\Gamma} = \begin{pmatrix} 2\langle\hat{X}_1\rangle^2_{\vec{\theta}} & 2\langle\hat{X}_3\rangle^2_{\vec{\theta}} & 2\langle\hat{X}_1\rangle_{\vec{\theta}}\langle\hat{X}_3\rangle_{\vec{\theta}} \\ 2\langle\hat{X}_3\rangle^2_{\vec{\theta}} & 2\langle\hat{X}_2\rangle^2_{\vec{\theta}} & 2\langle\hat{X}_2\rangle_{\vec{\theta}}\langle\hat{X}_3\rangle_{\vec{\theta}} \\ 2\langle\hat{X}_1\rangle_{\vec{\theta}}\langle\hat{X}_3\rangle_{\vec{\theta}} & 2\langle\hat{X}_2\rangle_{\vec{\theta}}\langle\hat{X}_3\rangle_{\vec{\theta}} & \langle\hat{X}_1\rangle_{\vec{\theta}}\langle\hat{X}_2\rangle_{\vec{\theta}} + \langle\hat{X}_3\rangle^2_{\vec{\theta}} \end{pmatrix}. \tag{4.43}$$

Then the moment matrix of the heterodyne detection can be calculated as:

$$\begin{aligned} M^{\text{Het}}_{\alpha\beta} &= \frac{\partial\langle\vec{\hat{X}}^T\rangle_{\vec{\theta}}}{\partial\theta_\alpha}\mathbf{\Gamma}^{-1}\frac{\partial\langle\vec{\hat{X}}\rangle_{\vec{\theta}}}{\partial\theta_\beta} \\ &= \frac{1}{2}\begin{pmatrix} \frac{\kappa^2\left(s^4+2\kappa^2s^2+2\kappa\left(1+s^2\right)s+1\right)}{s^2(\kappa+s)^2(1+\kappa s)^2} & -\frac{\kappa\left(1-s^2\right)\left(s^2+2\kappa s+1\right)}{s(\kappa+s)^2(1+\kappa s)^2} & 0 \\ -\frac{\kappa\left(1-s^2\right)\left(s^2+2\kappa s+1\right)}{s(\kappa+s)^2(1+\kappa s)^2} & \frac{s^2}{(1+\kappa s)^2}+\frac{1}{(\kappa+s)^2} & 0 \\ 0 & 0 & \frac{2\kappa^2\left(1-s^2\right)^2}{s(\kappa+s)(1+\kappa s)} \end{pmatrix}_{\alpha\beta}. \end{aligned} \tag{4.44}$$

As one might expect the moment matrix of the heterodyne detection does not depend on the direction of squeezing ϕ_s and has rank 3. Thus all the parameters can be estimated from the series of the measurements of the same observables $\vec{\hat{X}}$ (4.41). Note, however, that the direction of squeezing ϕ_s should still remain constant during the measurement process. Therefore, the source should either be phase-locked with the LO, or the acquisition time should be kept short.

The covariance matrix of the moment-based estimators is bounded by the inverse moment matrix

$$\left(M^{\mathrm{Het}}\right)^{-1} = \begin{pmatrix} \frac{s^4+2\kappa s^3+2\kappa^2 s^2+2\kappa s+1}{2\kappa^2} & \frac{\left(1-s^2\right)\left(s^2+2\kappa s+1\right)}{2\kappa s} & 0 \\ \frac{\left(1-s^2\right)\left(s^2+2\kappa s+1\right)}{2\kappa s} & \kappa^2+\frac{s^2}{2}+\frac{1}{2s^2}+\kappa s+\frac{\kappa}{s} & 0 \\ 0 & 0 & \frac{s(\kappa+s)(1+\kappa s)}{\kappa^2\left(1-s^2\right)^2} \end{pmatrix}. \tag{4.45}$$

The statistics of the heterodyne detection is given by the two-dimensional Gaussian distribution with the mean vector $\vec{m} = 0$ and the covariance matrix $\mathbf{C} = \mathbf{\Gamma}_Q + \mathbf{I}$, as given by eq. (1.123). Thus, the FI matrix of this measurement can be determined using eq. (2.67), resulting in a matrix identical to the moment matrix eq. (4.44). This signifies that the moment-based estimator optimally uses all the information contained in the heterodyne detection results and saturates CRB. This result is anticipated given the properties of Gaussian statistics of quadratures.

In fig. 4.7, we compare the CRB for homodyne and heterodyne detection. One can see, that for states with high purity (low κ, depicted in the left column of plots), heterodyne detection is the preferable choice as it offers lower noise in parameter estimation. Conversely, one should opt for heterodyne detection to analyze states with a stronger thermal contribution (higher κ, shown in the right column of plots 4.7), except in scenarios where squeezing is high. This observation aligns with the well-known phenomenon in purity estimation of thermal states, where homodyne detection is preferred for low-energy thermal states, while heterodyne detection yields better results for higher-energy states [Cenni 2022].

An analogous advantage of homodyne detection for estimating the covariance matrix of squeezed single-mode Gaussian states with high purity has been demonstrated in [Kumar 2022; Teo 2017; Řeháček 2015] through FI analysis. This behavior arises from the presence of an additional vacuum mode introduced into the system with heterodyne detection, resulting in additional noise. However, if the purity of the initial state is low, the noise in its quadratures is already high, and the additional vacuum noise does not significantly impact the measurement statistics. In such cases, measuring two quadratures (with heterodyne detection) instead of one (with homodyne detection) becomes advantageous. Besides, heterodyne detection often demonstrates an advantage in estimating the mean field [Teo 2017; Kumar 2022].

In [Kumar 2022] another measurement scheme is analyzed, wherein the studied state interacts with two squeezed vacuum probes. This measurement scheme, often referred to as the Arthurs-Kelly measurement scheme [Arthurs 1965], is shown to offer more comprehensive information about the studied state (especially about the mean-field [Kumar 2022]). However, its practical implementation is challenging.

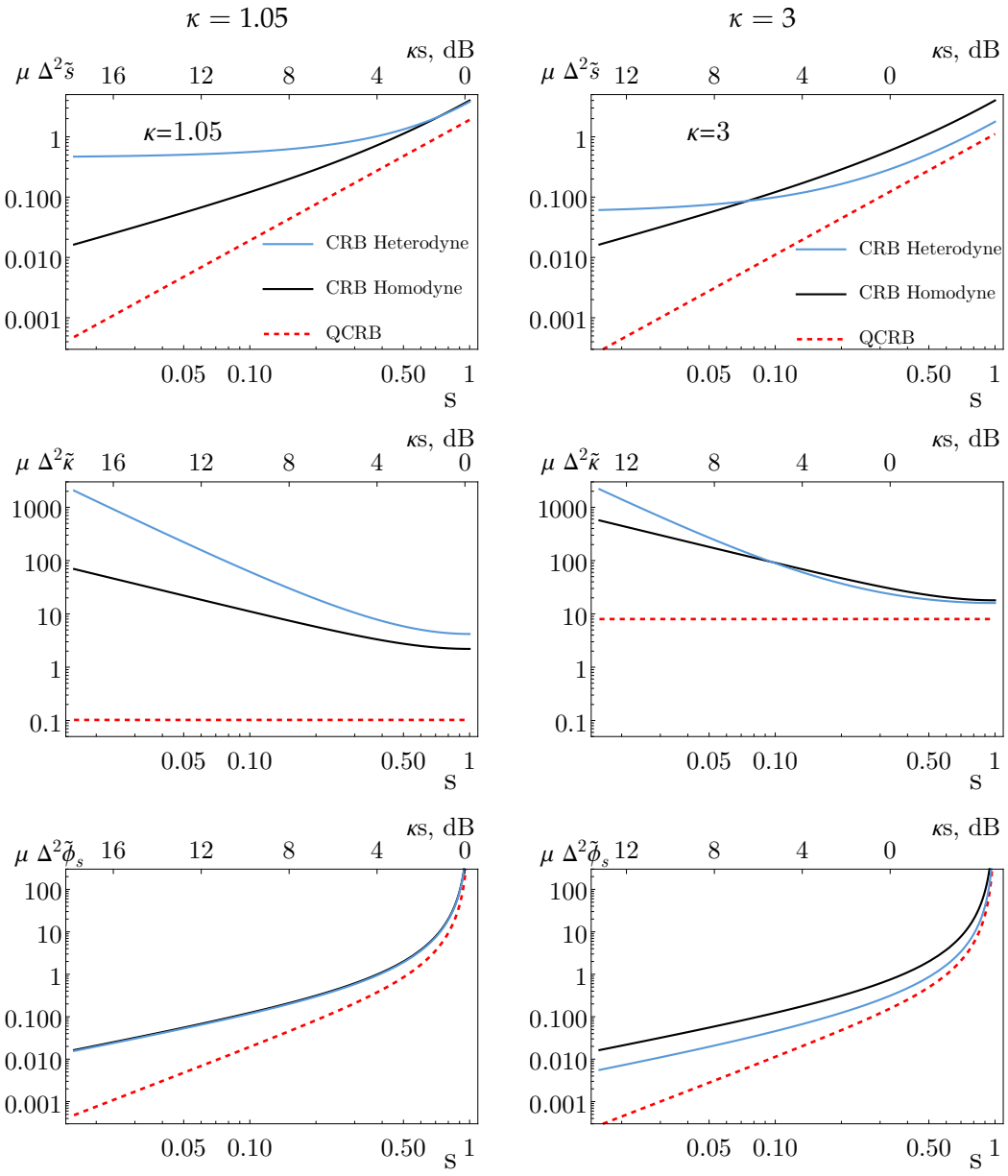

FIGURE 4.7: Parameter estimation sensitivity bounds for homodyne (black lines) and heterodyne (blue lines) detection. The ultimate bound, given by the QFI, is represented by the dashed red line. The left column of plots corresponds to the states with high purity ($\kappa = 1.05$, $P = 0.95$), the right one — to the states with low purity ($\kappa = 3$, $P = 0.33$).

Moment-based estimator

Constructing the estimator from the heterodyne measurement is a straight forward task. Since we build the estimator for three parameters, using the three moments (4.41), there is no need to construct linear combinations of the measurement results.

The estimator can be found as a solution of the three equations

$$\langle \hat{X}_\alpha \rangle_{\vec{\theta}} = \overline{x}_\alpha^{(\mu)}, \quad \alpha = 1,2,3, \tag{4.46}$$

where $\overline{x}_\alpha^{(\mu)}$ are sample means of the observables $\hat{X}_\alpha$ (4.41). The solution of this system reads:

$$\begin{pmatrix} \tilde{s} \\ \tilde{\kappa} \\ \tilde{\phi}_s \end{pmatrix} = \begin{pmatrix} \sqrt{\frac{1+2\overline{x}_1^{(\mu)}(\overline{x}_1^{(\mu)}-1)+2\overline{x}_2^{(\mu)}(\overline{x}_2^{(\mu)}-1)+4\overline{x}_3^{(\mu)2}-2(\overline{x}_1^{(\mu)}+\overline{x}_2^{(\mu)}-1)\sqrt{(\overline{x}_1^{(\mu)}-\overline{x}_2^{(\mu)})^2+4\overline{x}_3^{(\mu)2}}}{1-2(\overline{x}_1^{(\mu)}+\overline{x}_2^{(\mu)}-2\overline{x}_1^{(\mu)}\overline{x}_2^{(\mu)})-4\overline{x}_3^{(\mu)2}}} \\ \left((\overline{x}_1^{(\mu)}+\overline{x}_2^{(\mu)}-1)+\sqrt{(\overline{x}_1^{(\mu)}-\overline{x}_2^{(\mu)})^2+4\overline{x}_3^{(\mu)2}}\right)\tilde{s} \\ \arctan[\overline{x}_2^{(\mu)}-\overline{x}_1^{(\mu)}+\sqrt{(\overline{x}_1^{(\mu)}-\overline{x}_2^{(\mu)})^2+4\overline{x}_3^{(\mu)2}},\ -2\overline{x}_3^{(\mu)}], \end{pmatrix} \tag{4.47}$$

For this measurement scheme one does not need to have an initial guess about the parameters or process the data iteratively, it can be done in one step.

This procedure arises from the fact that the covariance matrix $\mathbf{\Gamma}(\hat{q}_1', \hat{p}_2')$ of the heterodyne observables $\hat{q}_1'$ and $\hat{p}_2'$ is linked to the quadrature covariance $\mathbf{\Gamma}_Q$ by a simple relation (1.123): $\mathbf{\Gamma}(\hat{q}_1', \hat{p}_2') = \mathbf{\Gamma}_Q + \mathbf{I}$. Consequently, the estimators for the elements of the covariance matrix are given by

$$\left([\tilde{\mathbf{\Gamma}}_Q]_{11}, [\tilde{\mathbf{\Gamma}}_Q]_{22}, [\tilde{\mathbf{\Gamma}}_Q]_{12}\right) = (\overline{x}_1^{(\mu)}-1, \overline{x}_2^{(\mu)}-1, \overline{x}_3^{(\mu)}). \tag{4.48}$$

And eq. (4.47) simply transforms the set of parameters $([\mathbf{\Gamma}_Q]_{11}, [\mathbf{\Gamma}_Q]_{22}, [\mathbf{\Gamma}_Q]_{12})$ to (s, κ, ϕ_s).

4.8 Conclusion

In this chapter, we introduced a moment-based estimator for the parameters of the quadrature covariance matrix of a single-mode Gaussian state. We constructed the estimator for the outcomes of scanning-homodyne (and heterodyne) measurements, demonstrating its optimality and unbiasedness. Concurrently, we illustrated that the intuitive Fourier estimator (which coincides with the LS estimator, fitting the variance curve) falls short in providing satisfactory estimation results for states with significant squeezing.

The developed estimator relies on straightforward algebraic transformations of the measured data and avoids the need for computationally intensive optimization methods like MLE. Low computational complexity and efficient use of the measured data make it promising for dynamic estimation of state parameters, in cases when the properties of the squeezed states source drift with time. This suggests potential applications in implementing real-time feedback control of such sources. For more efficient practical use in the dynamic setting, this approach may be further combined with the Bayesian parameter estimation, to account for the previous estimation results on each new step.

Another potential extension of this study involves generalizing the results to multi-mode Gaussian states. This presents a significant challenge for analytical investigation, as even the two-mode covariance matrix depends on 8 parameters. Furthermore, one might wish to analyze the 4 parameters (in the case of two modes) related to the mean field. Developing a moment-based estimator in this scenario will require numerical inversions of relatively large matrices and solving systems of nonlinear algebraic equations. However, the computational complexity of MLE in this case is also very high, and the reduction in complexity with the MoM can potentially be even greater than in the single-mode case.

Conclusion and perspectives

We started this thesis by presenting the theoretical concepts behind optical metrology, introducing the classical and quantum descriptions of the multimode electromagnetic field. We overviewed some existing approaches to the problem of parameter estimation and focused our attention on the approach based on the statistical moments of the measured data. We demonstrated that estimators based on the first moments of the observables often turned out to be optimal while being simple to calculate (along with the corresponding error bars) and to use in practice.

Using this approach, we have analyzed the problem of resolving point sources, expanding its traditional scope to include bright sources, sources with arbitrary statistics, and sources with separation-dependent mutual coherence and emission rate. We have studied the sensitivity of moment-based spatial mode demultiplexing, demonstrating not only its advantage over direct imaging but also its quantum optimality in many scenarios. For interactive dipoles and reflective particles under external illumination, we have shown that separation-dependent coherence, which arises in such scenarios, can significantly enhance the resolution capabilities.

We also considered the cases of simultaneous estimation of the separation, brightness, and relative phase of the sources. While in some instances we observed the resurgence of the Rayleigh curse in multiparameter scenarios, we demonstrated that SPADE sensitivity scales better with the separation compared to direct imaging.

Another application of the moment-based approach, that we analyzed in this thesis, is the problem of quantum state characterization. We constructed the moment-based estimator for the parameters of a single-mode Gaussian state based on the outcomes of scanning-homodyne (and heterodyne) measurements, demonstrating the optimality and unbiasedness of this approach.

This research only scratches the surface of the dynamically developing approach to the imaging problem, framed as parameter estimation with optimization of the measurement basis. In the context of microscopy, one can anticipate the extension of the model used in this thesis to optical systems with high numerical apertures, characterized by complex PSF and various aberrations. Addressing the challenge of realistic imaging necessitates studying multiple sources and objects with complex shapes, where coherence effects may yield even more unexpected results compared to incoherent imaging, which is typically considered at present. We acknowledge the pioneering steps towards the application of the SPADE approach, combined with neural network-based data processing, to the imaging of complex objects [Pushkina 2021; Costa Filho 2021] and foresee active development of both theoretical and experimental studies in this direction.

Developing adaptive approaches, based on the optimization of the measurement basis for imaging of the specific scene, naturally requires a Bayesian approach to the

description, especially in the dynamic setting, when the imaging object is not static. Thus, combining the moment-based approach with Bayesian analysis may prove practical for this and many other tasks.

The method of moments, employed in this research, holds promise to significantly expand its application scope. One of the interesting and practically useful directions of its development is the construction of moment-based estimators in the presence of nuisance parameters. Specifically in the context of realistic imaging, the total number of unknown parameters is usually very large, however, the number of parameters of interest is often limited. In this scenario, the moment-based analysis may help to find both the sensitive measurements and practical estimators, independent of values of nuisance parameters.

Another interesting statistical problem in this context is to develop some simple criteria for the optimality of the moment-based estimation. This may help to better understand the observation of this thesis, that the first-moment-based estimators often saturate the CRB in practical situations. Certainly, a more rigorous comparative study of the moment-based estimator with other standard estimation techniques, especially from the point of view of computational complexity, can further disseminate this approach, and it will find its application far beyond the scope of this thesis.

The tiny step we took towards moment-based state characterization may lead further research to numerous avenues in the world of quantum tomography. Our immediate next step goes undoubtedly in the direction of experimental testing of the developed approach. Subsequent generalizations may involve efficient characterization of multimode states, extension of the class of considered states and measurements, partial characterization of the state via the nuisance parameters approach, and addressing many other emerging challenges that may not even exist yet.

Bibliography

[Aaronson 2014] Scott Aaronson and Alex Arkhipov. "The Computational Complexity of Linear Optics". *Research in Optical Sciences (2014), paper QTh1A.2*. Optica Publishing Group, Mar. 2014, QTh1A.2. DOI: 10.1364/QIM.2014.QTh1A.2 (cit. on p. 55).

[Abbe 1873] E. Abbe. "Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung". *Archiv für Mikroskopische Anatomie* 9.1 (Dec. 1873), pp. 413–468. DOI: 10.1007/BF02956173 (cit. on p. 71).

[Acharya 2019] Anirudh Acharya, Theodore Kypraios, and Mădălin Guţă. "A comparative study of estimation methods in quantum tomography". *Journal of Physics A: Mathematical and Theoretical* 52.23 (May 2019), p. 234001. DOI: 10.1088/1751-8121/ab1958 (cit. on p. 63).

[Acuna 1997] Carmen O. Acuna and Joseph Horowitz. "A statistical approach to the resolution of point sources". *Journal of Applied Statistics* 24.4 (Aug. 1997), pp. 421–436. DOI: 10.1080/02664769723620 (cit. on p. 86).

[Adesso 2009] G. Adesso et al. "Optimal estimation of losses at the ultimate quantum limit with non-Gaussian states". *Physical Review A* 79.4 (Apr. 2009), p. 040305. DOI: 10.1103/PhysRevA.79.040305 (cit. on pp. 2, 48).

[Agarwal 1974] G. S. Agarwal. "Quantum statistical theories of spontaneous emission and their relation to other approaches". *Quantum Optics*. Ed. by G. Höhler. Springer Tracts in Modern Physics. Berlin, Heidelberg: Springer, 1974, pp. 1–128. DOI: 10.1007/BFb0042382 (cit. on p. 117).

[Agarwal 2013] Girish S. Agarwal. *Quantum Optics*. Cambridge University Press, 2013 (cit. on p. 93).

[Almeida 2021] J. O. de Almeida et al. "Discrimination and estimation of incoherent sources under misalignment". *Physical Review A* 103.2 (Feb. 2021), p. 022406. DOI: 10.1103/PhysRevA.103.022406 (cit. on p. 85).

[Altmann 2018] Yoann Altmann et al. "Quantum-inspired computational imaging". *Science* 361.6403 (Aug. 2018), eaat2298. DOI: 10.1126/science.aat2298 (cit. on p. 3).

[Andersen 2016] Ulrik L. Andersen et al. "30 years of squeezed light generation". *Physica Scripta* 91.5 (Apr. 2016), p. 053001. DOI: 10.1088/0031-8949/91/5/053001 (cit. on pp. 24, 128, 129, 144).

[Ang 2017] Shan Zheng Ang, Ranjith Nair, and Mankei Tsang. "Quantum limit for two-dimensional resolution of two incoherent optical point sources". *Physical Review A* 95.6 (June 2017), p. 063847. DOI: 10.1103/PhysRevA.95.063847 (cit. on p. 86).

[Ansari 2021] Vahid Ansari et al. "Achieving the Ultimate Quantum Timing Resolution". *PRX Quantum* 2.1 (Jan. 2021), p. 010301. DOI: 10.1103/PRXQuantum.2.010301 (cit. on p. 87).

[Arthurs 1965] E. Arthurs and J. L. Kelly. "B.S.T.J. briefs: On the simultaneous measurement of a pair of conjugate observables". *The Bell System Technical Journal* 44.4 (Apr. 1965), pp. 725–729. DOI: 10.1002/j.1538-7305.1965.tb01684.x (cit. on p. 145).

[Artiles 2005] L. M. Artiles, R. D. Gill, and M. I. Guţă. "An Invitation to Quantum Tomography". *Journal of the Royal Statistical Society Series B: Statistical Methodology* 67.1 (Feb. 2005), pp. 109–134. DOI: 10.1111/j.1467-9868.2005.00491.x (cit. on p. 60).

[Arvind 1995] Arvind et al. "The real symplectic groups in quantum mechanics and optics". *Pramana* 45.6 (Dec. 1995), pp. 471–497. DOI: 10.1007/BF02848172 (cit. on p. 27).

[Ash 2000] Robert B. Ash and Catherine A. Doleans-Dade. *Probability and Measure Theory*. Academic Press, 2000 (cit. on p. 139).

[Backlund 2018] Mikael P. Backlund, Yoav Shechtman, and Ronald L. Walsworth. "Fundamental Precision Bounds for Three-Dimensional Optical Localization Microscopy with Poisson Statistics". *Physical Review Letters* 121.2 (July 2018), p. 023904. DOI: 10.1103/PhysRevLett.121.023904 (cit. on p. 86).

[Bade 2018] Satyanarayana Bade et al. "Fabrication and Characterization of a Mode-selective 45-Mode Spatial Multiplexer based on Multi-Plane Light Conversion". *Optical Fiber Communication Conference Postdeadline Papers (2018), paper Th4B.3*. Optica Publishing Group, Mar. 2018, Th4B.3. DOI: 10.1364/OFC.2018.Th4B.3 (cit. on p. 81).

[Bao 2021] Fanglin Bao et al. "Quantum-accelerated imaging of N stars". *Optics Letters* 46.13 (July 2021), pp. 3045–3048. DOI: 10.1364/OL.430404 (cit. on p. 86).

[Barbastathis 2019] George Barbastathis, Aydogan Ozcan, and Guohai Situ. "On the use of deep learning for computational imaging". *Optica* 6.8 (Aug. 2019), pp. 921–943. DOI: 10.1364/OPTICA.6.000921 (cit. on p. 3).

[Barbieri 2022] Marco Barbieri. "Optical Quantum Metrology". *PRX Quantum* 3.1 (Jan. 2022), p. 010202. DOI: 10.1103/PRXQuantum.3.010202 (cit. on p. 1).

[Barrett 2004] Harrison H. Barrett and Kyle J. Myers. *Foundations of Image Science*. Wiley, 2004 (cit. on pp. 66–68).

[Barsotti 2018] Lisa Barsotti, Jan Harms, and Roman Schnabel. "Squeezed vacuum states of light for gravitational wave detectors". *Reports on Progress in Physics* 82.1 (Dec. 2018), p. 016905. DOI: 10.1088/1361-6633/aab906 (cit. on p. 128).

[Belthangady 2019] Chinmay Belthangady and Loic A. Royer. "Applications, promises, and pitfalls of deep learning for fluorescence image reconstruction". *Nature Methods* 16.12 (Dec. 2019), pp. 1215–1225. DOI: 10.1038/s41592-019-0458-z (cit. on p. 3).

[Berchera 2019] I. Ruo Berchera and I. P. Degiovanni. "Quantum imaging with sub-Poissonian light: challenges and perspectives in optical metrology". *Metrologia* 56.2 (Jan. 2019), p. 024001. DOI: 10.1088/1681-7575/aaf7b2 (cit. on p. 3).

[Beskrovnyy 2005] Vladislav N. Beskrovnyy and Mikhail I. Kolobov. "Quantum limits of super-resolution in reconstruction of optical objects". *Physical Review A* 71.4 (Apr. 2005), p. 043802. DOI: 10.1103/PhysRevA.71.043802 (cit. on p. 4).

[Betzig 2006] Eric Betzig et al. "Imaging Intracellular Fluorescent Proteins at Nanometer Resolution". *Science* 313.5793 (Sept. 2006), pp. 1642–1645. DOI: 10.1126/science.1127344 (cit. on p. 2).

[Birks 2015] T. A. Birks et al. "The photonic lantern". *Advances in Optics and Photonics* 7.2 (June 2015), pp. 107–167. DOI: 10.1364/AOP.7.000107 (cit. on p. 80).

[Bisketzi 2019] Evangelia Bisketzi, Dominic Branford, and Animesh Datta. "Quantum limits of localisation microscopy". *New Journal of Physics* 21.12 (Dec. 2019), p. 123032. DOI: 10.1088/1367-2630/ab58a0 (cit. on p. 86).

[Bloch 1962] Claude Bloch and Albert Messiah. "The canonical form of an antisymmetric tensor and its application to the theory of superconductivity". *Nuclear Physics* 39 (Dec. 1962), pp. 95–106. DOI: 10.1016/0029-5582(62)90377-2 (cit. on p. 27).

[Blume-Kohout 2010] Robin Blume-Kohout. "Optimal, reliable estimation of quantum states". *New Journal of Physics* 12.4 (Apr. 2010), p. 043034. DOI: 10.1088/1367-2630/12/4/043034 (cit. on p. 63).

[Bondurant 1984] Roy S. Bondurant and Jeffrey H. Shapiro. "Squeezed states in phase-sensing interferometers". *Physical Review D* 30.12 (Dec. 1984), pp. 2548–2556. DOI: 10.1103/PhysRevD.30.2548 (cit. on p. 128).

[Bonsma-Fisher 2019] Kent A. G. Bonsma-Fisher et al. "Realistic sub-Rayleigh imaging with phase-sensitive measurements". *New Journal of Physics* 21.9 (Sept. 2019), p. 093010. DOI: 10.1088/1367-2630/ab3d97 (cit. on p. 85).

[Born 1999] Max Born, Emil Wolf, and A. B. Bhatia. *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge University Press, Oct. 1999 (cit. on p. 10).

[Boucher 2020] Pauline Boucher et al. "Spatial optical mode demultiplexing as a practical tool for optimal transverse distance estimation". *Optica* 7.11 (Nov. 2020), pp. 1621–1626. DOI: 10.1364/OPTICA.404746 (cit. on pp. 81, 87).

[Bowen 2003] Warwick P. Bowen et al. "Experimental investigation of continuous-variable quantum teleportation". *Physical Review A* 67.3 (Mar. 2003), p. 032302. DOI: 10.1103/PhysRevA.67.032302 (cit. on p. 128).

[Braunstein 1994] Samuel L. Braunstein and Carlton M. Caves. "Statistical distance and the geometry of quantum states". *Physical Review Letters* 72.22 (May 1994), pp. 3439–3443. DOI: 10.1103/PhysRevLett.72.3439 (cit. on p. 34).

[Braunstein 2005] Samuel L. Braunstein and Peter van Loock. "Quantum information with continuous variables". *Reviews of Modern Physics* 77.2 (June 2005), pp. 513–577. DOI: 10.1103/RevModPhys.77.513 (cit. on pp. 27, 128).

[Brida 2010] G. Brida, M. Genovese, and I. Ruo Berchera. "Experimental realization of sub-shot-noise quantum imaging". *Nature Photonics* 4.4 (Apr. 2010), pp. 227–230. DOI: 10.1038/nphoton.2010.29 (cit. on p. 3).

[Brown 1954] R. Hanbury Brown and R.Q. Twiss. "LXXIV. A new type of interferometer for use in radio astronomy". *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 45.366 (July 1954), pp. 663–682. DOI: 10.1080/14786440708520475 (cit. on p. 4).

[Brown 2013] Robert Hanbury Brown and Richard Q. Twiss. "2. A Test of a New Type of Stellar Interferometer on Sirius". *2. A Test of a New Type of Stellar Interferometer on Sirius*. Harvard University Press, Oct. 2013, pp. 8–12. DOI: 10.4159/harvard.9780674366688.c4 (cit. on p. 4).

[Butucea 2007] Cristina Butucea, Mădălin Guţă, and Luis Artiles. "Minimax and Adaptive Estimation of the Wigner Function in Quantum Homodyne Tomography with Noisy Data". *The Annals of Statistics* 35.2 (2007), pp. 465–494 (cit. on p. 61).

[Cenni 2022] Marina F. B. Cenni et al. "Thermometry of Gaussian quantum systems using Gaussian measurements". *Quantum* 6 (June 2022), p. 743. DOI: 10.22331/q-2022-06-23-743 (cit. on p. 145).

[Chabaud 2021] Ulysse Chabaud et al. "Certification of Non-Gaussian States with Operational Measurements". *PRX Quantum* 2.2 (June 2021), p. 020333. DOI: 10.1103/PRXQuantum.2.020333 (cit. on pp. 28, 60).

[Chao 2016] Jerry Chao, E. Sally Ward, and Raimund J. Ober. "Fisher information theory for parameter estimation in single molecule microscopy: tutorial". *JOSA A* 33.7 (July 2016), B36–B57. DOI: 10.1364/JOSAA.33.000B36 (cit. on p. 94).

[Chapman 2022] Joseph C. Chapman et al. "Bayesian homodyne and heterodyne tomography". *Optics Express* 30.9 (Apr. 2022), pp. 15184–15200. DOI: 10.1364/OE.456597 (cit. on p. 63).

[Chiuri 2022] A. Chiuri et al. "Ghost imaging as loss estimation: Quantum versus classical schemes". *Physical Review A* 105.1 (Jan. 2022), p. 013506. DOI: 10.1103/PhysRevA.105.013506 (cit. on pp. 3, 47).

[Classen 2017] Anton Classen et al. "Superresolution via structured illumination quantum correlation microscopy". *Optica* 4.6 (June 2017), pp. 580–587. DOI: 10.1364/OPTICA.4.000580 (cit. on p. 3).

[Collaboration 2019] The Event Horizon Telescope Collaboration et al. "First M87 Event Horizon Telescope Results. IV. Imaging the Central Supermassive Black Hole". *The Astrophysical Journal Letters* 875.1 (Apr. 2019), p. L4. DOI: 10.3847/2041-8213/ab0e85 (cit. on p. 3).

[Combes 2004] Joshua Combes and H. M. Wiseman. "States for phase estimation in quantum interferometry". *Journal of Optics B: Quantum and Semiclassical Optics* 7.1 (Dec. 2004), p. 14. DOI: 10.1088/1464-4266/7/1/004 (cit. on p. 128).

[Costa Filho 2021] J. I. da Costa Filho. "Super-resolution microscopy with spatial modes of light". PhD thesis. University of Oxford, 2021 (cit. on pp. 80, 86, 149).

[Cramér 1946] H. Cramér. *Mathematical methods of statistics*. Mathematical methods of statistics. Princeton, NJ, US: Princeton University Press, 1946 (cit. on pp. 1, 35).

[DAngelo 2005a] M. D'Angelo and Y. H. Shih. "Quantum imaging". *Laser Physics Letters* 2.12 (Sept. 2005), p. 567. DOI: 10.1002/lapl.200510054 (cit. on p. 3).

[DAngelo 2005b] Milena D'Angelo et al. "Resolution of quantum and classical ghost imaging". *Physical Review A* 72.1 (July 2005), p. 013810. DOI: 10.1103/PhysRevA.72.013810 (cit. on p. 3).

[DAngelo 2016] Milena D'Angelo et al. "Correlation Plenoptic Imaging". *Physical Review Letters* 116.22 (June 2016), p. 223602. DOI: 10.1103/PhysRevLett.116.223602 (cit. on p. 4).

[DAriano 2003] G. Mauro D'Ariano, Matteo G. A. Paris, and Massimiliano F. Sacchi. *Quantum Tomography*. Feb. 2003 (cit. on pp. 60, 62).

[Datta 2020] Chandan Datta et al. "Sub-Rayleigh resolution of two incoherent sources by array homodyning". *Physical Review A* 102.6 (Dec. 2020), p. 063526. DOI: 10.1103/PhysRevA.102.063526 (cit. on pp. 80, 85).

[Datta 2021] Chandan Datta et al. "Sub-Rayleigh characterization of a binary source by spatially demultiplexed coherent detection". *Optics Express* 29.22 (Oct. 2021), pp. 35592–35601. DOI: 10.1364/OE.433990 (cit. on pp. 78, 80, 86).

[DAuria 2009] V. D'Auria et al. "Full Characterization of Gaussian Bipartite Entangled States by a Single Homodyne Detector". *Physical Review Letters* 102.2 (Jan. 2009), p. 020502. DOI: 10.1103/PhysRevLett.102.020502 (cit. on p. 130).

[De 2021] Syamsundar De et al. "Effects of coherence on temporal resolution". *Physical Review Research* 3.3 (July 2021), p. 033082. DOI: 10.1103/PhysRevResearch.3.033082 (cit. on p. 87).

[Defienne 2020] Hugo Defienne and Daniele Faccio. "Arbitrary spatial mode sorting in a multimode fiber". *Physical Review A* 101.6 (June 2020), p. 063830. DOI: 10.1103/PhysRevA.101.063830 (cit. on pp. 81, 82).

[Defienne 2022] Hugo Defienne et al. "Pixel super-resolution with spatially entangled photons". *Nature Communications* 13.1 (June 2022), p. 3566. DOI: 10.1038/s41467-022-31052-6 (cit. on p. 3).

[Delaubert 2006a] V. Delaubert et al. "${\mathrm{TEM}}_{10}$ homodyne detection as an optimal small-displacement and tilt-measurement scheme". *Physical Review A* 74.5 (Nov. 2006), p. 053823. DOI: 10.1103/PhysRevA.74.053823 (cit. on p. 4).

[Delaubert 2006b] Vincent Delaubert et al. "Quantum measurements of spatial conjugate variables: displacement and tilt of a Gaussian beam". *Optics Letters* 31.10 (May 2006), pp. 1537–1539. DOI: 10.1364/OL.31.001537 (cit. on p. 4).

[Demkowicz-Dobrzański 2012] Rafał Demkowicz-Dobrzański, Jan Kołodyński, and Mădălin Guţă. "The elusive Heisenberg limit in quantum-enhanced metrology". *Nature Communications* 3.1 (Sept. 2012), p. 1063. DOI: 10.1038/ncomms2067 (cit. on p. 1).

[Demkowicz-Dobrzański 2020] Rafał Demkowicz-Dobrzański, Wojciech Górecki, and Mădălin Guţă. "Multi-parameter estimation beyond quantum Fisher information". *Journal of Physics A: Mathematical and Theoretical* 53.36 (Aug. 2020), p. 363001. DOI: 10.1088/1751-8121/ab8ef3 (cit. on p. 59).

[Derkach 2020] Ivan Derkach, Vladyslav C. Usenko, and Radim Filip. "Squeezing-enhanced quantum key distribution over atmospheric channels". *New Journal of Physics* 22.5 (May 2020), p. 053006. DOI: 10.1088/1367-2630/ab7f8f (cit. on p. 128).

[Dertinger 2009] T. Dertinger et al. "Fast, background-free, 3D super-resolution optical fluctuation imaging (SOFI)". *Proceedings of the National Academy of Sciences* 106.52 (Dec. 2009), pp. 22287–22292. DOI: 10.1073/pnas.0907866106 (cit. on p. 3).

[Dertinger 2013] Thomas Dertinger et al. "Advances in superresolution optical fluctuation imaging (SOFI)". *Quarterly Reviews of Biophysics* 46.2 (May 2013), pp. 210–221. DOI: 10.1017/S0033583513000036 (cit. on p. 3).

[Deshler 2024] Nico Deshler, Sebastiaan Haffert, and Amit Ashok. *Achieving Quantum Limits of Exoplanet Detection and Localization*. Mar. 2024. DOI: 10.48550/arXiv.2403.17988 (cit. on p. 86).

[DeVoe 1996] R. G. DeVoe and R. G. Brewer. "Observation of Superradiant and Subradiant Spontaneous Emission of Two Trapped Ions". *Physical Review Letters* 76.12 (Mar. 1996), pp. 2049–2052. DOI: 10.1103/PhysRevLett.76.2049 (cit. on p. 119).

[Di Lena 2018] Francesco Di Lena et al. "Correlation Plenoptic Imaging: An Overview". *Applied Sciences* 8.10 (Oct. 2018), p. 1958. DOI: 10.3390/app8101958 (cit. on p. 4).

[Dodonov 2002] V. V. Dodonov. "'Nonclassical' states in quantum optics: a 'squeezed' review of the first 75 years". *Journal of Optics B: Quantum and Semiclassical Optics* 4.1 (Jan. 2002), R1. DOI: 10.1088/1464-4266/4/1/201 (cit. on p. 128).

[Donohue 2018] J. M. Donohue et al. "Quantum-Limited Time-Frequency Estimation through Mode-Selective Photon Measurement". *Physical Review Letters* 121.9 (Aug. 2018), p. 090501. DOI: 10.1103/PhysRevLett.121.090501 (cit. on p. 87).

[Eldar 2008] Yonina C. Eldar. "Rethinking Biased Estimation: Improving Maximum Likelihood and the Cramér–Rao Bound". *Foundations and Trends® in Signal Processing* 1.4 (July 2008), pp. 305–449. DOI: 10.1561/2000000008 (cit. on p. 34).

[Erkmen 2010] Baris I. Erkmen and Jeffrey H. Shapiro. "Ghost imaging: from quantum to classical to computational". *Advances in Optics and Photonics* 2.4 (Dec. 2010), pp. 405–450. DOI: 10.1364/AOP.2.000405 (cit. on p. 3).

[Esmaeil Zadeh 2021] Iman Esmaeil Zadeh et al. "Superconducting nanowire single-photon detectors: A perspective on evolution, state-of-the-art, future developments, and applications". *Applied Physics Letters* 118.19 (May 2021), p. 190502. DOI: 10.1063/5.0045990 (cit. on p. 20).

[Fabre 2000] C. Fabre, J. B. Fouet, and A. Maître. "Quantum limits in the measurement of very small displacements in optical images". *Optics Letters* 25.1 (Jan. 2000), pp. 76–78. DOI: 10.1364/OL.25.000076 (cit. on p. 4).

[Fabre 2020] C. Fabre and N. Treps. "Modes and states in quantum optics". *Reviews of Modern Physics* 92.3 (Sept. 2020), p. 035005. DOI: 10.1103/RevModPhys.92.035005 (cit. on pp. 7, 20).

[Fabre 2023] Claude Fabre. *Quantum Processes and Measurement: Theory and Experiment*. Cambridge University Press, July 2023 (cit. on p. 19).

[Fainsin 2023] David Fainsin. "Continuous Variable Multimode Quantum States at Telecommunication Wavelengths for Quantum Networks". PhD thesis. Sorbonne Université, Dec. 2023 (cit. on pp. 128, 134).

[Ferraro 2011] Pietro Ferraro, Adam Wax, and Zeev Zalevsky. *Coherent Light Microscopy: Imaging and Quantitative Phase Analysis*. Springer Science & Business Media, Feb. 2011 (cit. on p. 92).

[Fiderer 2021] Lukas J. Fiderer et al. "General Expressions for the Quantum Fisher Information Matrix with Applications to Discrete Quantum Imaging". *PRX Quantum* 2.2 (Apr. 2021), p. 020308. DOI: 10.1103/PRXQuantum.2.020308 (cit. on pp. 4, 86).

[Fischer 2010] Hans Fischer. *A History of the Central Limit Theorem: From Classical to Modern Probability Theory*. Springer Science & Business Media, Oct. 2010 (cit. on p. 38).

[Fisher 1925] R. A. Fisher. "Theory of Statistical Estimation". *Mathematical Proceedings of the Cambridge Philosophical Society* 22.5 (July 1925), pp. 700–725. DOI: 10.1017/S0305004100009580 (cit. on pp. 1, 34).

[Fontaine 2021] Nicolas K. Fontaine et al. "Hermite-Gaussian mode multiplexer supporting 1035 modes". *Optical Fiber Communication Conference (OFC) 2021 (2021), paper M3D.4*. Optica Publishing Group, June 2021, p. M3D.4. DOI: 10.1364/OFC.2021.M3D.4 (cit. on p. 80).

[Furusawa 1998] A. Furusawa et al. "Unconditional Quantum Teleportation". *Science* 282.5389 (Oct. 1998), pp. 706–709. DOI: 10.1126/science.282.5389.706 (cit. on p. 128).

[Gåsvik 2003] Kjell J. Gåsvik. *Optical Metrology*. John Wiley & Sons, Apr. 2003 (cit. on p. 1).

[Gatti 2004] A. Gatti et al. "Correlated imaging, quantum and classical". *Physical Review A* 70.1 (July 2004), p. 013802. DOI: 10.1103/PhysRevA.70.013802 (cit. on p. 3).

[Gatto Monticone 2014] D. Gatto Monticone et al. "Beating the Abbe Diffraction Limit in Confocal Microscopy via Nonclassical Photon Statistics". *Physical Review Letters* 113.14 (Sept. 2014), p. 143602. DOI: 10.1103/PhysRevLett.113.143602 (cit. on p. 3).

[Geissbuehler 2011] Stefan Geissbuehler, Claudio Dellagiacoma, and Theo Lasser. "Comparison between SOFI and STORM". *Biomedical Optics Express* 2.3 (Mar. 2011), pp. 408–420. DOI: 10.1364/BOE.2.000408 (cit. on p. 2).

[Gessner 2019] Manuel Gessner, Augusto Smerzi, and Luca Pezzè. "Metrological Nonlinear Squeezing Parameter". *Physical Review Letters* 122.9 (Mar. 2019), p. 090503. DOI: 10.1103/PhysRevLett.122.090503 (cit. on pp. 2, 32).

[Gessner 2020a] Manuel Gessner, Claude Fabre, and Nicolas Treps. "Superresolution Limits from Measurement Crosstalk". *Physical Review Letters* 125.10 (Aug. 2020), p. 100501. DOI: 10.1103/PhysRevLett.125.100501 (cit. on p. 85).

[Gessner 2020b] Manuel Gessner, Augusto Smerzi, and Luca Pezzè. "Multiparameter squeezing for optimal quantum enhancements in sensor networks". *Nature Communications* 11.1 (July 2020), p. 3817. DOI: 10.1038/s41467-020-17471-3 (cit. on pp. 2, 32, 44, 46).

[Gill 2013] Richard D. Gill and Mădălin I. Guţă. "On asymptotic quantum statistical inference". *From Probability to Statistics and Back: High-Dimensional Models and Processes – A Festschrift in Honor of Jon A. Wellner*. Vol. 9. Institute of Mathematical Statistics, Jan. 2013, pp. 105–128. DOI: 10.1214/12-IMSCOLL909 (cit. on p. 59).

[Giovannetti 2001] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. "Quantum-enhanced positioning and clock synchronization". *Nature* 412.6845 (July 2001), pp. 417–419. DOI: 10.1038/35086525 (cit. on p. 128).

[Giovannetti 2006] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. "Quantum Metrology". *Physical Review Letters* 96.1 (Jan. 2006), p. 010401. DOI: 10.1103/PhysRevLett.96.010401 (cit. on p. 1).

[Giovannetti 2011] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. "Advances in quantum metrology". *Nature Photonics* 5.4 (Apr. 2011), pp. 222–229. DOI: 10.1038/nphoton.2011.35 (cit. on p. 1).

[Glauber 1963] Roy J. Glauber. "The Quantum Theory of Optical Coherence". *Physical Review* 130.6 (June 1963), pp. 2529–2539. DOI: 10.1103/PhysRev.130.2529 (cit. on pp. 7, 20).

[Goodman 2005] Joseph W. Goodman. *Introduction to Fourier Optics*. Roberts and Company Publishers, 2005 (cit. on pp. 10, 66, 68).

[Gosalia 2023] Ronakraj K. Gosalia et al. *Quantum Super-Resolution with Balanced Homodyne Detection in Low-Earth-Orbit*. June 2023 (cit. on p. 86).

[Gottesman 2003] Daniel Gottesman and John Preskill. "Secure Quantum Key Distribution using Squeezed States". *Quantum Information with Continuous Variables*. Ed. by Samuel L. Braunstein and Arun K. Pati. Dordrecht: Springer Netherlands, 2003, pp. 317–356. DOI: 10.1007/978-94-015-1258-9_22 (cit. on p. 128).

[Grace 2020] Michael R. Grace et al. "Approaching quantum-limited imaging resolution without prior knowledge of the object location". *JOSA A* 37.8 (Aug. 2020), pp. 1288–1299. DOI: 10.1364/JOSAA.392116 (cit. on p. 86).

[Grace 2022] Michael R. Grace and Saikat Guha. "Identifying Objects at the Quantum Limit for Superresolution Imaging". *Physical Review Letters* 129.18 (Oct. 2022), p. 180502. DOI: 10.1103/PhysRevLett.129.180502 (cit. on p. 86).

[Greenwood 2023] Alexander B. Greenwood, Ruth Oulton, and Henkjan Gersen. "On the impact of realistic point sources in spatial mode demultiplexing super resolution imaging". *Quantum Science and Technology* 8.1 (Jan. 2023), p. 015024. DOI: 10.1088/2058-9565/aca0b7 (cit. on p. 87).

[Grote 2013] H. Grote et al. "First Long-Term Application of Squeezed States of Light in a Gravitational-Wave Observatory". *Physical Review Letters* 110.18 (May 2013), p. 181101. DOI: 10.1103/PhysRevLett.110.181101 (cit. on p. 128).

[Grynberg 2010] Gilbert Grynberg, Alain Aspect, and Claude Fabre. *Introduction to Quantum Optics: From the Semi-classical Approach to Quantized Light*. Cambridge University Press, Sept. 2010 (cit. on p. 12).

[Guerra 1990] John M. Guerra. "Photon tunneling microscopy". *Applied Optics* 29.26 (Sept. 1990), pp. 3741–3752. DOI: 10.1364/AO.29.003741 (cit. on p. 3).

[Guo 2020] Xueshi Guo et al. "Distributed quantum sensing in a continuous-variable entangled network". *Nature Physics* 16.3 (Mar. 2020), pp. 281–284. DOI: 10.1038/s41567-019-0743-x (cit. on p. 92).

[Gustafsson 2000] M. G. L. Gustafsson. "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy". *Journal of Microscopy* 198.2 (2000), pp. 82–87. DOI: 10.1046/j.1365-2818.2000.00710.x (cit. on p. 3).

[Gustafsson 2008] Mats G. L. Gustafsson et al. "Three-Dimensional Resolution Doubling in Wide-Field Fluorescence Microscopy by Structured Illumination". *Biophysical Journal* 94.12 (June 2008), pp. 4957–4970. DOI: 10.1529/biophysj.107.120345 (cit. on p. 3).

[Guţă 2020] M. Guţă et al. "Fast state tomography with optimal error bounds". *Journal of Physics A: Mathematical and Theoretical* 53.20 (Apr. 2020), p. 204001. DOI: 10.1088/1751-8121/ab8111 (cit. on p. 63).

[Haah 2016] Jeongwan Haah et al. "Sample-optimal tomography of quantum states". *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*. STOC '16. New York, NY, USA: Association for Computing Machinery, June 2016, pp. 913–925. DOI: 10.1145/2897518.2897585 (cit. on p. 63).

[Hand 2010] Eric Hand. "Telescope arrays give fine view of stars". *Nature* 464.7290 (Apr. 2010), pp. 820–821. DOI: 10.1038/464820a (cit. on p. 4).

[Hecht 2000] Bert Hecht et al. "Scanning near-field optical microscopy with aperture probes: Fundamentals and applications". *The Journal of Chemical Physics* 112.18 (May 2000), pp. 7761–7774. DOI: 10.1063/1.481382 (cit. on p. 3).

[Heide 2013] Felix Heide et al. "High-quality computational imaging through simple lenses". *ACM Transactions on Graphics* 32.5 (Oct. 2013), 149:1–149:14. DOI: 10.1145/2516971.2516974 (cit. on p. 3).

[Heintzmann 2017] Rainer Heintzmann and Thomas Huser. "Super-Resolution Structured Illumination Microscopy". *Chemical Reviews* 117.23 (Dec. 2017), pp. 13890–13908. DOI: 10.1021/acs.chemrev.7b00218 (cit. on p. 3).

[Hell 1994] Stefan W. Hell and Jan Wichmann. "Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy". *Optics Letters* 19.11 (June 1994), pp. 780–782. DOI: 10.1364/OL.19.000780 (cit. on p. 2).

[Hellinger 1909] E. Hellinger. "Neue Begründung der Theorie quadratischer Formen von unendlichvielen Veränderlichen." *Journal für die reine und angewandte Mathematik* 1909.136 (July 1909), pp. 210–271. DOI: 10.1515/crll.1909.136.210 (cit. on p. 35).

[Helstrom 1967] Carl W. Helstrom. "Image Restoration by the Method of Least Squares". *JOSA* 57.3 (Mar. 1967), pp. 297–303. DOI: 10.1364/JOSA.57.000297 (cit. on p. 4).

[Helstrom 1969] Carl W. Helstrom. "Quantum detection and estimation theory". *Journal of Statistical Physics* 1.2 (June 1969), pp. 231–252. DOI: 10.1007/BF01007479 (cit. on p. 54).

[Helstrom 1973] C. Helstrom. "Resolution of point sources of light as analyzed by quantum detection theory". *IEEE Transactions on Information Theory* 19.4 (July 1973), pp. 389–398. DOI: 10.1109/TIT.1973.1055052 (cit. on p. 86).

[Herman 2009] Gabor T. Herman. *Fundamentals of Computerized Tomography: Image Reconstruction from Projections*. Springer Science & Business Media, July 2009 (cit. on p. 61).

[Hess 2006] Samuel T. Hess, Thanu P. K. Girirajan, and Michael D. Mason. "Ultra-High Resolution Imaging by Fluorescence Photoactivation Localization Microscopy". *Biophysical Journal* 91.11 (Dec. 2006), pp. 4258–4272. DOI: 10.1529/biophysj.106.091116 (cit. on p. 2).

[Higham 2018] Catherine F. Higham et al. "Deep learning for real-time single-pixel video". *Scientific Reports* 8.1 (Feb. 2018), p. 2369. DOI: 10.1038/s41598-018-20521-y (cit. on p. 3).

[Hoffman 2021] David P. Hoffman, Isaac Slavitt, and Casey A. Fitzpatrick. "The promise and peril of deep learning in microscopy". *Nature Methods* 18.2 (Feb. 2021), pp. 131–132. DOI: 10.1038/s41592-020-01035-w (cit. on p. 3).

[Holevo 1973] A. S Holevo. "Statistical decision theory for quantum systems". *Journal of Multivariate Analysis* 3.4 (Dec. 1973), pp. 337–394. DOI: 10.1016/0047-259X(73)90028-6 (cit. on p. 59).

[Hradil 1997] Z. Hradil. "Quantum-state estimation". *Physical Review A* 55.3 (Mar. 1997), R1561–R1564. DOI: 10.1103/PhysRevA.55.R1561 (cit. on p. 62).

[Hradil 2019] Zdeněk Hradil et al. "Quantum Fisher information with coherence". *Optica* 6.11 (Nov. 2019), pp. 1437–1440. DOI: 10.1364/OPTICA.6.001437 (cit. on pp. 5, 87, 109).

[Hradil 2021] Zdeněk Hradil, Dominik Koutný, and Jaroslav Řeháček. "Exploring the ultimate limits: super-resolution enhanced by partial coherence". *Optics Letters* 46.7 (Apr. 2021), pp. 1728–1731. DOI: 10.1364/OL.417988 (cit. on pp. 5, 87, 109).

[Hsieh 2022a] Hsien-Yi Hsieh et al. "Direct Parameter Estimations from Machine Learning-Enhanced Quantum State Tomography". *Symmetry* 14.5 (May 2022), p. 874. DOI: 10.3390/sym14050874 (cit. on p. 141).

[Hsieh 2022b] Hsien-Yi Hsieh et al. "Extract the Degradation Information in Squeezed States with Machine Learning". *Physical Review Letters* 128.7 (Feb. 2022), p. 073604. DOI: 10.1103/PhysRevLett.128.073604 (cit. on pp. 63, 141).

[Huang 2008] Bo Huang et al. "Three-Dimensional Super-Resolution Imaging by Stochastic Optical Reconstruction Microscopy". *Science* 319.5864 (Feb. 2008), pp. 810–813. DOI: 10.1126/science.1153529 (cit. on p. 2).

[Huang 2009] Bo Huang, Mark Bates, and Xiaowei Zhuang. "Super-Resolution Fluorescence Microscopy". *Annual Review of Biochemistry* 78.Volume 78, 2009 (July 2009), pp. 993–1016. DOI: 10.1146/annurev.biochem.77.061906.092014 (cit. on p. 2).

[Huang 2021] Zixin Huang and Cosmo Lupo. "Quantum Hypothesis Testing for Exoplanet Detection". *Physical Review Letters* 127.13 (Sept. 2021), p. 130502. DOI: 10.1103/PhysRevLett.127.130502 (cit. on p. 86).

[Huang 2023] Zixin Huang, Christian Schwab, and Cosmo Lupo. "Ultimate limits of exoplanet spectroscopy: A quantum approach". *Physical Review A* 107.2 (Feb. 2023), p. 022409. DOI: 10.1103/PhysRevA.107.022409 (cit. on p. 86).

[Huang 2024] Jiahao Huang, Min Zhuang, and Chaohong Lee. *Entanglement-enhanced quantum metrology: from standard quantum limit to Heisenberg limit*. Feb. 2024 (cit. on p. 1).

[James 1993] Daniel F. V. James. "Frequency shifts in spontaneous emission from two interacting atoms". *Physical Review A* 47.2 (Feb. 1993), pp. 1336–1346. DOI: 10.1103/PhysRevA.47.1336 (cit. on p. 117).

[Jennison 1958] R. C. Jennison. "A Phase Sensitive Interferometer Technique for the Measurement of the Fourier Transforms of Spatial Brightness Distributions of Small Angular Extent". *Monthly Notices of the Royal Astronomical Society* 118.3 (June 1958), pp. 276–284. DOI: 10.1093/mnras/118.3.276 (cit. on p. 3).

[Kaiser 2016] F. Kaiser et al. "A fully guided-wave squeezing experiment for fiber quantum networks". *Optica* 3.4 (Apr. 2016), pp. 362–365. DOI: 10.1364/OPTICA.3.000362 (cit. on p. 134).

[Karuseichyk 2022] Ilya Karuseichyk et al. "Resolving mutually-coherent point sources of light with arbitrary statistics". *Physical Review Research* 4.4 (Oct. 2022), p. 043010. DOI: 10.1103/PhysRevResearch.4.043010 (cit. on pp. 5, 66, 77, 88).

[Karuseichyk 2024] Ilya Karuseichyk et al. "Exploiting separation-dependent coherence to boost optical resolution". *Physical Review A* 109.4 (Apr. 2024), p. 043524. DOI: 10.1103/PhysRevA.109.043524 (cit. on pp. 5, 66, 109).

[Kay 1998] Steven M. Kay. *Fundamentals of Statistical Signal Processing: Detection theory*. Prentice-Hall PTR, 1998 (cit. on pp. 1, 2, 32, 35–38, 45, 135, 138).

[Kellermann 2001] K. I. Kellermann and J. M. Moran. "The Development of High-Resolution Imaging in Radio Astronomy1". *Annual Review of Astronomy and Astrophysics* 39.Volume 39, 2001 (Sept. 2001), pp. 457–509. DOI: 10.1146/annurev.astro.39.1.457 (cit. on p. 3).

[Kelley 1964] P. L. Kelley and W. H. Kleiner. "Theory of Electromagnetic Field Measurement and Photoelectron Counting". *Physical Review* 136.2A (Oct. 1964), A316–A334. DOI: 10.1103/PhysRev.136.A316 (cit. on p. 20).

[Khetkeeree 2020] Suphongsa Khetkeeree. "Optimization of Lucy-Richardson Algorithm Using Modified Tikhonov Regularization for Image Deblurring". *Journal of Physics: Conference Series* 1438.1 (Jan. 2020), p. 012014. DOI: 10.1088/1742-6596/1438/1/012014 (cit. on p. 3).

[Kiss 1995] T. Kiss, U. Herzog, and U. Leonhardt. "Compensation of losses in photodetection and in quantum-state measurements". *Physical Review A* 52.3 (Sept. 1995), pp. 2433–2435. DOI: 10.1103/PhysRevA.52.2433 (cit. on p. 20).

[Klar 1999] Thomas A. Klar and Stefan W. Hell. "Subdiffraction resolution in far-field fluorescence microscopy". *Optics Letters* 24.14 (July 1999), pp. 954–956. DOI: 10.1364/OL.24.000954 (cit. on p. 2).

[Kolobov 1993] Mikhail I. Kolobov and Prem Kumar. "Sub-shot-noise microscopy: imaging of faint phase objects with squeezed light". *Optics Letters* 18.11 (June 1993), pp. 849–851. DOI: 10.1364/OL.18.000849 (cit. on p. 128).

[Kolobov 2000] Mikhail I. Kolobov and Claude Fabre. "Quantum Limits on Optical Resolution". *Physical Review Letters* 85.18 (Oct. 2000), pp. 3789–3792. DOI: 10.1103/PhysRevLett.85.3789 (cit. on p. 4).

[Köse 2022] Emre Köse, Gerardo Adesso, and Daniel Braun. "Quantum-enhanced passive remote sensing". *Physical Review A* 106.1 (July 2022), p. 012601. DOI: 10.1103/PhysRevA.106.012601 (cit. on p. 4).

[Köse 2023] Emre Köse and Daniel Braun. "Superresolution imaging with multiparameter quantum metrology in passive remote sensing". *Physical Review A* 107.3 (Mar. 2023), p. 032607. DOI: 10.1103/PhysRevA.107.032607 (cit. on p. 4).

[Kouadou 2023] Tiphaine Kouadou et al. "Spectrally shaped and pulse-by-pulse multiplexed multimode squeezed states of light". *APL Photonics* 8.8 (Aug. 2023), p. 086113. DOI: 10.1063/5.0156331 (cit. on p. 128).

[Koutný 2021] D. Koutný et al. "Axial superlocalization with vortex beams". *Quantum Science and Technology* 6.2 (Mar. 2021), p. 025021. DOI: 10.1088/2058-9565/abe8ca (cit. on p. 86).

[Koutný 2022] Dominik Koutný et al. "Neural-network quantum state tomography". *Physical Review A* 106.1 (July 2022), p. 012409. DOI: 10.1103/PhysRevA.106.012409 (cit. on p. 63).

[Kumar 2021] Santosh Kumar et al. "Time-Spatial Mode Selective Quantum Frequency Converter". *Physical Review A* 104.2 (Aug. 2021), p. 023506. DOI: 10.1103/PhysRevA.104.023506 (cit. on pp. 81, 87).

[Kumar 2022] Chandan Kumar and Arvind. "Estimation of the Wigner distribution of single-mode Gaussian states: A comparative study". *Physical Review A* 105.4 (Apr. 2022), p. 042419. DOI: 10.1103/PhysRevA.105.042419 (cit. on p. 145).

[Kundur 1996] D. Kundur and D. Hatzinakos. "Blind image deconvolution". *IEEE Signal Processing Magazine* 13.3 (May 1996), pp. 43–64. DOI: 10.1109/79.489268 (cit. on p. 3).

[Kupianskyi 2023] Hlib Kupianskyi, Simon A. R. Horsley, and David B. Phillips. "High-dimensional spatial mode sorting and optical circuit design using multi-plane light conversion". *APL Photonics* 8.2 (Feb. 2023), p. 026101. DOI: 10.1063/5.0128431 (cit. on pp. 80, 81).

[Kurdzialek 2022] Stanislaw Kurdzialek. "Back to sources – the role of losses and coherence in super-resolution imaging revisited". *Quantum* 6 (Apr. 2022), p. 697. DOI: 10.22331/q-2022-04-27-697 (cit. on pp. 5, 87, 92, 101, 109, 112).

[Kurdziałek 2023] Stanisław Kurdziałek and Rafał Demkowicz-Dobrzański. "Measurement Noise Susceptibility in Quantum Estimation". *Physical Review Letters* 130.16 (Apr. 2023), p. 160802. DOI: 10.1103/PhysRevLett.130.160802 (cit. on pp. 85, 90).

[Kyrillidis 2018] Anastasios Kyrillidis et al. "Provable compressed sensing quantum state tomography via non-convex methods". *npj Quantum Information* 4.1 (Aug. 2018), pp. 1–7. DOI: 10.1038/s41534-018-0080-4 (cit. on p. 60).

[Labroille 2014] Guillaume Labroille et al. "Efficient and mode selective spatial mode multiplexer based on multi-plane light conversion". *Optics Express* 22.13 (June 2014), pp. 15599–15607. DOI: 10.1364/OE.22.015599 (cit. on p. 80).

[Larson 2018] Walker Larson and Bahaa E. A. Saleh. "Resurgence of Rayleigh's curse in the presence of partial coherence". *Optica* 5.11 (Nov. 2018), pp. 1382–1389. DOI: 10.1364/OPTICA.5.001382 (cit. on pp. 5, 87, 109).

[Larson 2019] Walker Larson and Bahaa E. A. Saleh. "Resurgence of Rayleigh's curse in the presence of partial coherence: reply". *Optica* 6.4 (Apr. 2019), pp. 402–403. DOI: 10.1364/OPTICA.6.000402 (cit. on pp. 5, 87, 109).

[Laudenbach 2018] Fabian Laudenbach et al. "Continuous-Variable Quantum Key Distribution with Gaussian Modulation – The Theory of Practical Implementations". *Advanced Quantum Technologies* 1.1 (Aug. 2018), p. 1800011. DOI: 10.1002/qute.201800011 (cit. on p. 128).

[Laurenzis 2019] Martin Laurenzis and Frank Christnacher. "Active imaging". *Advanced Optical Technologies* 8.6 (Dec. 2019), pp. 423–424. DOI: 10.1515/aot-2019-0056 (cit. on p. 2).

[Lawrie 2019] B. J. Lawrie et al. "Quantum Sensing with Squeezed Light". *ACS Photonics* 6.6 (June 2019), pp. 1307–1318. DOI: 10.1021/acsphotonics.9b00250 (cit. on p. 128).

[Lee 2012] Sang-Hyuk Lee et al. "Counting single photoactivatable fluorescent molecules by photoactivated localization microscopy (PALM)". *Proceedings of the National Academy of Sciences* 109.43 (Oct. 2012), pp. 17436–17441. DOI: 10.1073/pnas.1215175109 (cit. on p. 2).

[Lee 2022] Kwan Kit Lee et al. *Quantum Multi-Parameter Adaptive Bayesian Estimation and Application to Super-Resolution Imaging*. June 2022. DOI: 10.48550/arXiv.2202.09980 (cit. on p. 87).

[Lehmberg 1970] R. H. Lehmberg. "Radiation from an N-Atom System. I. General Formalism". *Physical Review A* 2.3 (Sept. 1970), pp. 883–888. DOI: 10.1103/PhysRevA.2.883 (cit. on p. 116).

[Lemos 2014] Gabriela Barreto Lemos et al. "Quantum imaging with undetected photons". *Nature* 512.7515 (Aug. 2014), pp. 409–412. DOI: 10.1038/nature13586 (cit. on p. 3).

[Lemos 2022] Gabriela Barreto Lemos et al. "Quantum imaging and metrology with undetected photons: tutorial". *JOSA B* 39.8 (Aug. 2022), pp. 2200–2228. DOI: 10.1364/JOSAB.456778 (cit. on p. 3).

[Len 2020] Yink Loong Len et al. "Resolution limits of spatial mode demultiplexing with noisy detection". *International Journal of Quantum Information* 18.01 (Feb. 2020), p. 1941015. DOI: 10.1142/S0219749919410156 (cit. on pp. 78, 80, 85, 90).

[Leon-Saval 2013] Sergio G. Leon-Saval, Alexander Argyros, and Joss Bland-Hawthorn. "Photonic lanterns". *Nanophotonics* 2.5-6 (Dec. 2013), pp. 429–440. DOI: 10.1515/nanoph-2013-0035 (cit. on pp. 80, 81).

[Lesser 2015] Michael Lesser. "A Summary of Charge-Coupled Devices for Astronomy". *Publications of the Astronomical Society of the Pacific* 127.957 (Nov. 2015), p. 1097. DOI: 10.1086/684054 (cit. on p. 20).

[Leung 2011] Bonnie O. Leung and Keng C. Chou. "Review of Super-Resolution Fluorescence Microscopy for Biology". *Applied Spectroscopy* 65.9 (Sept. 2011), pp. 967–980. DOI: 10.1366/11-06398 (cit. on p. 2).

[Li 2024] Junyan Li and Shengshi Pang. *Quantum-limited superresolution of two incoherent point sources with unknown photon numbers*. Mar. 2024. DOI: 10.48550/arXiv.2403.13752 (cit. on p. 86).

[Liang 2021] Kevin Liang, S. A. Wadood, and A. N. Vamivakas. "Coherence effects on estimating two-point separation". *Optica* 8.2 (Feb. 2021), pp. 243–248. DOI: 10.1364/OPTICA.403497 (cit. on pp. 5, 75, 87, 109).

[Liang 2023a] Kevin Liang. "Off-axis aberrations improve the resolution limits of incoherent imaging". *Optics Express* 31.7 (Mar. 2023), pp. 11173–11184. DOI: 10.1364/OE.479799 (cit. on p. 69).

[Liang 2023b] Kevin Liang, S. A. Wadood, and A. N. Vamivakas. "Quantum Fisher information for estimating N partially coherent point sources". *Optics Express* 31.2 (Jan. 2023), pp. 2726–2743. DOI: 10.1364/OE.474036 (cit. on pp. 5, 87, 109).

[Liao 2022] Xiao-Jun Liao and Yong-Qi Fu. "Quantum metrology with multimode Gaussian states of multiple point sources". *Physical Review A* 106.2 (Aug. 2022), p. 022602. DOI: 10.1103/PhysRevA.106.022602 (cit. on p. 86).

[Lim 2024] Sumin Lim et al. *Multiphoton super-resolution imaging via virtual structured illumination*. Apr. 2024. DOI: 10.48550/arXiv.2404.11849 (cit. on p. 3).

[Linowski 2023] Tomasz Linowski et al. "Application range of crosstalk-affected spatial demultiplexing for resolving separations between unbalanced sources". *New Journal of Physics* 25.10 (Oct. 2023), p. 103050. DOI: 10.1088/1367-2630/ad0173 (cit. on pp. 85, 86).

[Liu 2019] Jing Liu et al. "Quantum Fisher information matrix and multiparameter estimation". *Journal of Physics A: Mathematical and Theoretical* 53.2 (Dec. 2019), p. 023001. DOI: 10.1088/1751-8121/ab5d4d (cit. on pp. 32, 54–58).

[Lohani 2020] Sanjaya Lohani et al. "Machine learning assisted quantum state estimation". *Machine Learning: Science and Technology* 1.3 (July 2020), p. 035007. DOI: 10.1088/2632-2153/ab9a21 (cit. on p. 63).

[Lopaeva 2013] E. D. Lopaeva et al. "Experimental Realization of Quantum Illumination". *Physical Review Letters* 110.15 (Apr. 2013), p. 153603. DOI: 10.1103/PhysRevLett.110.153603 (cit. on p. 3).

[Losero 2018] Elena Losero et al. "Unbiased estimation of an optical loss at the ultimate quantum limit with twin-beams". *Scientific Reports* 8.1 (May 2018), p. 7431. DOI: 10.1038/s41598-018-25501-w (cit. on p. 2).

[Lu 2018] Xiao-Ming Lu et al. "Quantum-optimal detection of one-versus-two incoherent optical sources with arbitrary separation". *npj Quantum Information* 4.1 (Dec. 2018), pp. 1–8. DOI: 10.1038/s41534-018-0114-y (cit. on pp. 74, 86).

[Lucy 1974] L. B. Lucy. "An iterative technique for the rectification of observed distributions". *The Astronomical Journal* 79 (June 1974), p. 745. DOI: 10.1086/111605 (cit. on p. 3).

[Lupo 2016] Cosmo Lupo and Stefano Pirandola. "Ultimate Precision Bound of Quantum and Subwavelength Imaging". *Physical Review Letters* 117.19 (Nov. 2016), p. 190802. DOI: 10.1103/PhysRevLett.117.190802 (cit. on pp. 47, 74–76, 85, 96, 101, 111, 112).

[Lupo 2020a] Cosmo Lupo. "Subwavelength quantum imaging with noisy detectors". *Physical Review A* 101.2 (Feb. 2020), p. 022323. DOI: 10.1103/PhysRevA.101.022323 (cit. on p. 85).

[Lupo 2020b] Cosmo Lupo, Zixin Huang, and Pieter Kok. "Quantum Limits to Incoherent Imaging are Achieved by Linear Interferometry". *Physical Review Letters* 124.8 (Feb. 2020), p. 080503. DOI: 10.1103/PhysRevLett.124.080503 (cit. on pp. 86, 100).

[Lvovsky 2002] A. I. Lvovsky and S. A. Babichev. "Synthesis and tomographic characterization of the displaced Fock state of light". *Physical Review A* 66.1 (July 2002), p. 011801. DOI: 10.1103/PhysRevA.66.011801 (cit. on p. 61).

[Lvovsky 2009] A. I. Lvovsky and M. G. Raymer. "Continuous-variable optical quantum-state tomography". *Reviews of Modern Physics* 81.1 (Mar. 2009), pp. 299–332. DOI: 10.1103/RevModPhys.81.299 (cit. on pp. 32, 59–63).

[Manzano 2020] Daniel Manzano. "A short introduction to the Lindblad Master Equation". *AIP Advances* 10.2 (Feb. 2020), p. 025106. DOI: 10.1063/1.5115323 (cit. on p. 19).

[Massaro 2023] Gianlorenzo Massaro et al. "Correlated-photon imaging at 10 volumetric images per second". *Scientific Reports* 13.1 (Aug. 2023), p. 12813. DOI: 10.1038/s41598-023-39416-8 (cit. on p. 4).

[Masters 2021] Barry R. Masters. *Superresolution Optical Microscopy: The Quest for Enhanced Resolution and Contrast*. Springer International Publishing, Aug. 2021 (cit. on p. 2).

[McCuller 2020] L. McCuller et al. "Frequency-Dependent Squeezing for Advanced LIGO". *Physical Review Letters* 124.17 (Apr. 2020), p. 171102. DOI: 10.1103/PhysRevLett.124.171102 (cit. on p. 128).

[Meda 2017] A. Meda et al. "Photon-number correlation for quantum enhanced imaging and sensing". *Journal of Optics* 19.9 (Aug. 2017), p. 094002. DOI: 10.1088/2040-8986/aa7b27 (cit. on p. 3).

[Mikhalychev 2018] A. Mikhalychev et al. "Synthesis of Quantum Antennas for Shaping Field Correlations". *Physical Review Applied* 9.2 (Feb. 2018), p. 024021. DOI: 10.1103/PhysRevApplied.9.024021 (cit. on p. 117).

[Mikhalychev 2019] A. B. Mikhalychev et al. "Efficiently reconstructing compound objects by quantum imaging with higher-order correlation functions". *Communications Physics* 2.1 (Oct. 2019), pp. 1–8. DOI: 10.1038/s42005-019-0234-5 (cit. on pp. 3, 4, 115).

[Mikhalychev 2021] A. B. Mikhalychev et al. "Lost photon enhances superresolution". *npj Quantum Information* 7.1 (Aug. 2021), pp. 1–10. DOI: 10.1038/s41534-021-00465-4 (cit. on p. 3).

[Monras 2007] Alex Monras and Matteo G. A. Paris. "Optimal Quantum Estimation of Loss in Bosonic Channels". *Physical Review Letters* 98.16 (Apr. 2007), p. 160401. DOI: 10.1103/PhysRevLett.98.160401 (cit. on p. 2).

[Morais 2022] Leonardo Assis Morais et al. *Precisely determining photon-number in real-time*. Apr. 2022. DOI: 10.48550/arXiv.2012.10158 (cit. on p. 20).

[Morizur 2010] Jean-François Morizur et al. "Programmable unitary spatial mode manipulation". *JOSA A* 27.11 (Nov. 2010), pp. 2524–2531. DOI: 10.1364/JOSAA.27.002524 (cit. on p. 80).

[Motka 2016] L. Motka et al. "Optical resolution from Fisher information". *The European Physical Journal Plus* 131.5 (May 2016), p. 130. DOI: 10.1140/epjp/i2016-16130-7 (cit. on p. 4).

[Nair 2016] Ranjith Nair and Mankei Tsang. "Far-Field Superresolution of Thermal Electromagnetic Sources at the Quantum Limit". *Physical Review Letters* 117.19 (Nov. 2016), p. 190801. DOI: 10.1103/PhysRevLett.117.190801 (cit. on p. 85).

[Nair 2018] Ranjith Nair. "Quantum-Limited Loss Sensing: Multiparameter Estimation and Bures Distance between Loss Channels". *Physical Review Letters* 121.23 (Dec. 2018), p. 230801. DOI: 10.1103/PhysRevLett.121.230801 (cit. on p. 2).

[Napoli 2019] Carmine Napoli et al. "Towards Superresolution Surface Metrology: Quantum Estimation of Angular and Axial Separations". *Physical Review Letters* 122.14 (Apr. 2019), p. 140505. DOI: 10.1103/PhysRevLett.122.140505 (cit. on p. 86).

[Oh 2021] Changhun Oh et al. "Quantum Limits of Superresolution in a Noisy Environment". *Physical Review Letters* 126.12 (Mar. 2021), p. 120502. DOI: 10.1103/PhysRevLett.126.120502 (cit. on p. 85).

[Ohtsu 1995] M. Ohtsu. "Progress of high-resolution photon scanning tunneling microscopy due to a nanometric fiber probe". *Journal of Lightwave Technology* 13.7 (July 1995), pp. 1200–1221. DOI: 10.1109/50.400696 (cit. on p. 3).

[Olivares 2019] Stefano Olivares et al. "Quantum tomography of light states by photon-number-resolving detectors". *New Journal of Physics* 21.10 (Oct. 2019), p. 103045. DOI: 10.1088/1367-2630/ab4afb (cit. on p. 60).

[Ortolano 2023] Giuseppe Ortolano et al. "Quantum-Enhanced Pattern Recognition". *Physical Review Applied* 20.2 (Aug. 2023), p. 024072. DOI: 10.1103/PhysRevApplied.20.024072 (cit. on p. 86).

[Paniate 2024] A. Paniate et al. "Light-field ghost imaging". *Physical Review Applied* 21.2 (Feb. 2024), p. 024032. DOI: 10.1103/PhysRevApplied.21.024032 (cit. on p. 4).

[Papoulis 2002] Athanasios Papoulis and S. Unnikrishna Pillai. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, 2002 (cit. on p. 52).

[Parniak 2018] Michał Parniak et al. "Beating the Rayleigh Limit Using Two-Photon Interference". *Physical Review Letters* 121.25 (Dec. 2018), p. 250503. DOI: 10.1103/PhysRevLett.121.250503 (cit. on pp. 86, 87).

[Patera 2019] Giuseppe Patera, Dmitri Horoshko, and Mikhail Kolobov. "Quantum temporal imaging with squeezed light". *EPJ Web of Conferences* 198 (2019), p. 00002. DOI: 10.1051/epjconf/201919800002 (cit. on p. 128).

[Paúr 2016] Martin Paúr et al. "Achieving the ultimate optical resolution". *Optica* 3.10 (Oct. 2016), pp. 1144–1147. DOI: 10.1364/OPTICA.3.001144 (cit. on p. 87).

[Paúr 2018] Martin Paúr et al. "Tempering Rayleigh's curse with PSF shaping". *Optica* 5.10 (Oct. 2018), pp. 1177–1180. DOI: 10.1364/OPTICA.5.001177 (cit. on p. 85).

[Paúr 2019] M. Paúr et al. "Reading out Fisher information from the zeros of the point spread function". *Optics Letters* 44.12 (June 2019), pp. 3114–3117. DOI: 10.1364/OL.44.003114 (cit. on p. 85).

[Pepe 2017] Francesco V. Pepe et al. "Diffraction-Limited Plenoptic Imaging with Correlated Light". *Physical Review Letters* 119.24 (Dec. 2017), p. 243602. DOI: 10.1103/PhysRevLett.119.243602 (cit. on p. 4).

[Peshko 2019] I. Peshko et al. "Quantum noise radar: superresolution with quantum antennas by accessing spatiotemporal correlations". *Optics Express* 27.20 (Sept. 2019), pp. 29217–29231. DOI: 10.1364/OE.27.029217 (cit. on pp. 3, 121).

[Pezze 2014] Luca Pezze' and Augusto Smerzi. *Quantum theory of phase estimation*. Nov. 2014. DOI: 10.48550/arXiv.1411.5164 (cit. on p. 128).

[Phillips 2019] D. S. Phillips et al. "Benchmarking of Gaussian boson sampling using two-point correlators". *Physical Review A* 99.2 (Feb. 2019), p. 023836. DOI: 10.1103/PhysRevA.99.023836 (cit. on p. 109).

[Pinel 2013] O. Pinel et al. "Quantum parameter estimation using general single-mode Gaussian states". *Physical Review A* 88.4 (Oct. 2013), p. 040102. DOI: 10.1103/PhysRevA.88.040102 (cit. on p. 56).

[Pirandola 2020] S. Pirandola et al. "Advances in quantum cryptography". *Advances in Optics and Photonics* 12.4 (Dec. 2020), pp. 1012–1236. DOI: 10.1364/AOP.361502 (cit. on p. 128).

[Prasad 2020] Sudhakar Prasad. "Quantum limited super-resolution of an unequal-brightness source pair in three dimensions". *Physica Scripta* 95.5 (Mar. 2020), p. 054004. DOI: 10.1088/1402-4896/ab573d (cit. on p. 86).

[Price 2011] Robert L. Price and W. Gray (Jay) Jerome. *Basic Confocal Microscopy*. Springer Science & Business Media, Apr. 2011 (cit. on p. 3).

[Pushkina 2021] A. A. Pushkina et al. "Superresolution Linear Optical Imaging in the Far Field". *Physical Review Letters* 127.25 (Dec. 2021), p. 253602. DOI: 10.1103/PhysRevLett.127.253602 (cit. on pp. 4, 80, 86, 87, 149).

[Ragy 2016] Sammy Ragy, Marcin Jarzyna, and Rafał Demkowicz-Dobrzański. "Compatibility in multiparameter quantum metrology". *Physical Review A* 94.5 (Nov. 2016), p. 052108. DOI: 10.1103/PhysRevA.94.052108 (cit. on p. 59).

[Rao 1994] Calyampudi Radhakrishna Rao and S. Das Gupta. *Selected Papers of C. R. Rao*. Wiley, 1994 (cit. on pp. 1, 35).

[Rayleigh 1879] Rayleigh. "XXXI. Investigations in optics, with special reference to the spectroscope". *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 8.49 (Oct. 1879), pp. 261–274. DOI: 10.1080/14786447908639684 (cit. on p. 70).

[Řeháček 2007] Jaroslav Řeháček et al. "Diluted maximum-likelihood algorithm for quantum tomography". *Physical Review A* 75.4 (Apr. 2007), p. 042108. DOI: 10.1103/PhysRevA.75.042108 (cit. on pp. 62, 63).

[Řeháček 2015] Jaroslav Řeháček et al. "Surmounting intrinsic quantum-measurement uncertainties in Gaussian-state tomography with quadrature squeezing". *Scientific Reports* 5.1 (July 2015), p. 12289. DOI: 10.1038/srep12289 (cit. on pp. 130, 145).

[Rehacek 2017] J. Rehacek et al. "Optimal measurements for resolution beyond the Rayleigh limit". *Optics Letters* 42.2 (Jan. 2017), pp. 231–234. DOI: 10.1364/OL.42.000231 (cit. on pp. 82, 83, 86).

[Řehaček 2017] J. Řehaček et al. "Multiparameter quantum metrology of incoherent point sources: Towards realistic superresolution". *Physical Review A* 96.6 (Dec. 2017), p. 062107. DOI: 10.1103/PhysRevA.96.062107 (cit. on p. 86).

[Řeháček 2018] J. Řeháček et al. "Optimal measurements for quantum spatial superresolution". *Physical Review A* 98.1 (July 2018), p. 012103. DOI: 10.1103/PhysRevA.98.012103 (cit. on p. 86).

[Rivenson 2017] Yair Rivenson et al. "Deep learning microscopy". *Optica* 4.11 (Nov. 2017), pp. 1437–1443. DOI: 10.1364/OPTICA.4.001437 (cit. on p. 3).

[Roggemann 1997] Michael C. Roggemann, Byron M. Welsh, and Robert Q. Fugate. "Improving the resolution of ground-based telescopes". *Reviews of Modern Physics* 69.2 (Apr. 1997), pp. 437–506. DOI: 10.1103/RevModPhys.69.437 (cit. on p. 3).

[Roman-Rodriguez 2024] Victor Roman-Rodriguez et al. *Multimode Squeezed State for Reconfigurable Quantum Networks at Telecommunication Wavelengths*. Feb. 2024. DOI: 10.48550/arXiv.2306.07267 (cit. on pp. 128, 134).

[Rosenberg 2022] Jason S. Rosenberg et al. "Observation of the Hanbury Brown–Twiss effect with ultracold molecules". *Nature Physics* 18.9 (Sept. 2022), pp. 1062–1066. DOI: 10.1038/s41567-022-01695-9 (cit. on p. 4).

[Rouvière 2024] Clémentine Rouvière et al. "Ultra-sensitive separation estimation of optical sources". *Optica* 11.2 (Feb. 2024), pp. 166–170. DOI: https://doi.org/10.1364/OPTICA.500039 (cit. on pp. 5, 81, 87).

[Rust 2006] Michael J. Rust, Mark Bates, and Xiaowei Zhuang. "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)". *Nature Methods* 3.10 (Oct. 2006), pp. 793–796. DOI: 10.1038/nmeth929 (cit. on p. 2).

[Sajjad 2024] Aqil Sajjad, Michael R. Grace, and Saikat Guha. "Quantum limits of parameter estimation in long-baseline imaging". *Physical Review Research* 6.1 (Feb. 2024), p. 013212. DOI: 10.1103/PhysRevResearch.6.013212 (cit. on p. 85).

[Samantaray 2017] Nigam Samantaray et al. "Realization of the first sub-shot-noise wide field microscope". *Light: Science & Applications* 6.7 (July 2017), e17005–e17005. DOI: 10.1038/lsa.2017.5 (cit. on p. 3).

[Sansavini 2023] Francesca Sansavini. "Time and Frequency Multiplexing for Continuous-Variable Quantum Networks". PhD thesis. Sorbonne Université ; Helsingin yliopisto (Finland), Dec. 2023 (cit. on p. 134).

[Santamaria 2023] Luigi Santamaria et al. "Spatial-mode demultiplexing for enhanced intensity and distance measurement". *Optics Express* 31.21 (Oct. 2023), pp. 33930–33944. DOI: 10.1364/OE.486617 (cit. on p. 87).

[Santamaria 2024] Luigi Santamaria, Fabrizio Sgobba, and Cosmo Lupo. "Single-photon sub-Rayleigh precision measurements of a pair of incoherent sources of unequal intensity". *Optica Quantum* 2.1 (Feb. 2024), pp. 46–56. DOI: 10.1364/OPTICAQ.505457 (cit. on pp. 81, 86, 87).

[Saxena 2015] Manish Saxena, Gangadhar Eluru, and Sai Siva Gorthi. "Structured illumination microscopy". *Advances in Optics and Photonics* 7.2 (June 2015), pp. 241–275. DOI: 10.1364/AOP.7.000241 (cit. on p. 3).

[Schlichtholz 2024] Konrad Schlichtholz et al. "Practical tests for sub-Rayleigh source discriminations with imperfect demultiplexers". *Optica Quantum* 2.1 (Feb. 2024), pp. 29–34. DOI: 10.1364/OPTICAQ.502459 (cit. on pp. 75, 86).

[Schnabel 2017] Roman Schnabel. "Squeezed states of light and their applications in laser interferometers". *Physics Reports*. Squeezed states of light and their applications in laser interferometers 684 (Apr. 2017), pp. 1–51. DOI: 10.1016/j.physrep.2017.04.001 (cit. on p. 128).

[Schuh 2012] H. Schuh and D. Behrend. "VLBI: A fascinating technique for geodesy and astrometry". *Journal of Geodynamics* 61 (Oct. 2012), pp. 68–80. DOI: 10.1016/j.jog.2012.07.007 (cit. on p. 4).

[Scully 1997] Marlan O. Scully and M. Suhail Zubairy. *Quantum Optics*. Cambridge University Press, Sept. 1997 (cit. on pp. 77, 89, 117).

[Scutaru 1998] H. Scutaru. "Fidelity for displaced squeezed thermal states and the oscillator semigroup". *Journal of Physics A: Mathematical and General* 31.15 (Apr. 1998), p. 3659. DOI: 10.1088/0305-4470/31/15/025 (cit. on p. 56).

[Sephton 2019] Bereneice Sephton et al. "Spatial mode detection by frequency upconversion". *Optics Letters* 44.3 (Feb. 2019), pp. 586–589. DOI: 10.1364/OL.44.000586 (cit. on p. 81).

[Shahram 2006] M. Shahram and P. Milanfar. "Statistical and Information-Theoretic Analysis of Resolution in Imaging". *IEEE Transactions on Information Theory* 52.8 (Aug. 2006), pp. 3411–3437. DOI: 10.1109/TIT.2006.878180 (cit. on p. 86).

[Shapiro 2008] Jeffrey H. Shapiro. "Computational ghost imaging". *Physical Review A* 78.6 (Dec. 2008), p. 061802. DOI: 10.1103/PhysRevA.78.061802 (cit. on p. 3).

[Shapiro 2012] Jeffrey H. Shapiro and Robert W. Boyd. "The physics of ghost imaging". *Quantum Information Processing* 11.4 (Aug. 2012), pp. 949–993. DOI: 10.1007/s11128-011-0356-5 (cit. on p. 3).

[Shapiro 2015] Jeffrey H. Shapiro, Dheera Venkatraman, and Franco N. C. Wong. "Classical Imaging with Undetected Photons". *Scientific Reports* 5.1 (May 2015), p. 10329. DOI: 10.1038/srep10329 (cit. on p. 3).

[Sherman 1950] Jack Sherman and Winifred J. Morrison. "Adjustment of an Inverse Matrix Corresponding to a Change in One Element of a Given Matrix". *The Annals of Mathematical Statistics* 21.1 (Mar. 1950), pp. 124–127. DOI: 10.1214/aoms/1177729893 (cit. on p. 90).

[Sibarita 2005] Jean-Baptiste Sibarita. "Deconvolution Microscopy". *Microscopy Techniques: -/-*. Ed. by Jens Rietdorf. Berlin, Heidelberg: Springer, 2005, pp. 201–243. DOI: 10.1007/b102215 (cit. on p. 3).

[Sidhu 2020] Jasminder S. Sidhu and Pieter Kok. "Geometric perspective on quantum parameter estimation". *AVS Quantum Science* 2.1 (Feb. 2020), p. 014701. DOI: 10.1116/1.5119961 (cit. on p. 55).

[Simon 1994] R. Simon, N. Mukunda, and Biswadeb Dutta. "Quantum-noise matrix for multimode systems: U(n) invariance, squeezing, and normal forms". *Physical Review A* 49.3 (Mar. 1994), pp. 1567–1583. DOI: 10.1103/PhysRevA.49.1567 (cit. on p. 26).

[Smith 2013] A. Smith et al. "Quantum state tomography by continuous measurement and compressed sensing". *Physical Review A* 87.3 (Mar. 2013), p. 030102. DOI: 10.1103/PhysRevA.87.030102 (cit. on p. 60).

[Soh 2023] Daniel Soh and Eric Chatterjee. "Label-free quantum super-resolution imaging using entangled multimode squeezed light". *New Journal of Physics* 25.9 (Sept. 2023), p. 093001. DOI: 10.1088/1367-2630/acf2ba (cit. on p. 128).

[Sorelli 2021a] Giacomo Sorelli et al. "Moment-based superresolution: Formalism and applications". *Physical Review A* 104.3 (Sept. 2021), p. 033515. DOI: 10.1103/PhysRevA.104.033515 (cit. on pp. 41, 75, 85).

[Sorelli 2021b] Giacomo Sorelli et al. "Optimal Observables and Estimators for Practical Superresolution Imaging". *Physical Review Letters* 127.12 (Sept. 2021), p. 123604. DOI: 10.1103/PhysRevLett.127.123604 (cit. on pp. 76, 78, 85, 90, 96).

[Sorelli 2022] Giacomo Sorelli et al. "Quantum limits for resolving Gaussian sources". *Physical Review Research* 4.3 (Aug. 2022), p. L032022. DOI: 10.1103/PhysRevResearch.4.L032022 (cit. on pp. 55, 76, 85, 87, 101, 102, 111).

[Sparrow 1916] C. M. Sparrow. "On Spectroscopic Resolving Power". *The Astrophysical Journal* 44 (Sept. 1916), p. 76. DOI: 10.1086/142271 (cit. on p. 70).

[Sroda 2020] Aleksandra Sroda et al. "SOFISM: Super-resolution optical fluctuation image scanning microscopy". *Optica* 7.10 (Oct. 2020), pp. 1308–1316. DOI: 10.1364/OPTICA.399600 (cit. on p. 3).

[Stein 2014] Manuel Stein, Amine Mezghani, and Josef A. Nossek. "A Lower Bound for the Fisher Information Measure". *IEEE Signal Processing Letters* 21.7 (July 2014), pp. 796–799. DOI: 10.1109/LSP.2014.2316008 (cit. on p. 46).

[Stoler 1970] David Stoler. "Equivalence Classes of Minimum Uncertainty Packets". *Physical Review D* 1.12 (June 1970), pp. 3217–3219. DOI: 10.1103/PhysRevD.1.3217 (cit. on p. 24).

[Ströhl 2016] Florian Ströhl and Clemens F. Kaminski. "Frontiers in structured illumination microscopy". *Optica* 3.6 (June 2016), pp. 667–677. DOI: 10.1364/OPTICA.3.000667 (cit. on p. 3).

[Suzuki 2020] Jun Suzuki, Yuxiang Yang, and Masahito Hayashi. "Quantum state estimation with nuisance parameters". *Journal of Physics A: Mathematical and Theoretical* 53.45 (Oct. 2020), p. 453001. DOI: 10.1088/1751-8121/ab8b78 (cit. on pp. 44, 104).

[Tan 2023a] Xiao-Jie Tan and Mankei Tsang. "Quantum limit to subdiffraction incoherent optical imaging. III. Numerical analysis". *Physical Review A* 108.5 (Nov. 2023), p. 052416. DOI: 10.1103/PhysRevA.108.052416 (cit. on p. 87).

[Tan 2023b] Xiao-Jie Tan et al. "Quantum-inspired superresolution for incoherent imaging". *Optica* 10.9 (Sept. 2023), pp. 1189–1194. DOI: 10.1364/OPTICA.493227 (cit. on p. 81).

[Tanaś 2004] R. Tanaś and Z. Ficek. "Entangling two atoms via spontaneous emission". *Journal of Optics B: Quantum and Semiclassical Optics* 6.3 (Mar. 2004), S90. DOI: 10.1088/1464-4266/6/3/015 (cit. on p. 121).

[Tang 2016] Zong Sheng Tang, Kadir Durak, and Alexander Ling. "Fault-tolerant and finite-error localization for point emitters within the diffraction limit". *Optics Express* 24.19 (Sept. 2016), pp. 22004–22012. DOI: 10.1364/OE.24.022004 (cit. on p. 87).

[Tao 2019] Xu Tao et al. "A high speed and high efficiency superconducting photon number resolving detector". *Superconductor Science and Technology* 32.6 (Apr. 2019), p. 064002. DOI: 10.1088/1361-6668/ab0799 (cit. on p. 20).

[Tata 1998] B. V. R. Tata and Baldev Raj. "Confocal laser scanning microscopy: Applications in material science and technology". *Bulletin of Materials Science* 21.4 (Aug. 1998), pp. 263–278. DOI: 10.1007/BF02744951 (cit. on p. 3).

[Teo 2017] Y. S. Teo et al. "Superiority of heterodyning over homodyning: An assessment with quadrature moments". *Physical Review A* 95.4 (Apr. 2017), p. 042322. DOI: 10.1103/PhysRevA.95.042322 (cit. on pp. 130, 145).

[Tham 2017] Weng-Kian Tham, Hugo Ferretti, and Aephraim M. Steinberg. "Beating Rayleigh's Curse by Imaging Using Phase Information". *Physical Review Letters* 118.7 (Feb. 2017), p. 070801. DOI: 10.1103/PhysRevLett.118.070801 (cit. on pp. 85, 87).

[Tiunov 2020] E. S. Tiunov et al. "Experimental quantum homodyne tomography via machine learning". *Optica* 7.5 (May 2020), pp. 448–454. DOI: 10.1364/OPTICA.389482 (cit. on p. 63).

[Trees 2001] Harry L. Van Trees. *Detection, Estimation, and Modulation Theory: Detection, estimation, and linear modulation theory*. Wiley, 2001 (cit. on pp. 1, 32, 37).

[Treps 2003] Nicolas Treps et al. "A Quantum Laser Pointer". *Science* 301.5635 (Aug. 2003), pp. 940–943. DOI: 10.1126/science.1086489 (cit. on pp. 4, 128).

[Treps 2004] N. Treps et al. "Nano-displacement measurements using spatially multimode squeezed light". *Journal of Optics B: Quantum and Semiclassical Optics* 6.8 (July 2004), S664. DOI: 10.1088/1464-4266/6/8/007 (cit. on pp. 4, 128).

[Treps 2005] N. Treps et al. "Quantum noise in multipixel image processing". *Physical Review A* 71.1 (Jan. 2005), p. 013820. DOI: 10.1103/PhysRevA.71.013820 (cit. on p. 4).

[Tsang 2015] Mankei Tsang. "Quantum limits to optical point-source localization". *Optica* 2.7 (July 2015), pp. 646–653. DOI: 10.1364/OPTICA.2.000646 (cit. on p. 86).

[Tsang 2016] Mankei Tsang, Ranjith Nair, and Xiao-Ming Lu. "Quantum Theory of Superresolution for Two Incoherent Optical Point Sources". *Physical Review X* 6.3 (Aug. 2016), p. 031033. DOI: 10.1103/PhysRevX.6.031033 (cit. on pp. 4, 21, 75, 80, 84, 85, 87, 94).

[Tsang 2018] Mankei Tsang. "Conservative classical and quantum resolution limits for incoherent imaging". *Journal of Modern Optics* 65.11 (June 2018), pp. 1385–1391. DOI: 10.1080/09500340.2017.1377306 (cit. on p. 87).

[Tsang 2019a] Mankei Tsang. "Quantum limit to subdiffraction incoherent optical imaging". *Physical Review A* 99.1 (Jan. 2019), p. 012305. DOI: 10.1103/PhysRevA.99.012305 (cit. on p. 86).

[Tsang 2019b] Mankei Tsang. "Resolving starlight: a quantum perspective". *Contemporary Physics* 60.4 (Oct. 2019), pp. 279–298. DOI: 10.1080/00107514.2020.1736375 (cit. on pp. 78, 84, 87).

[Tsang 2019c] Mankei Tsang. "Semiparametric estimation for incoherent optical imaging". *Physical Review Research* 1.3 (Oct. 2019), p. 033006. DOI: 10.1103/PhysRevResearch.1.033006 (cit. on pp. 4, 86).

[Tsang 2019d] Mankei Tsang and Ranjith Nair. "Resurgence of Rayleigh's curse in the presence of partial coherence: comment". *Optica* 6.4 (Apr. 2019), pp. 400–401. DOI: 10.1364/OPTICA.6.000400 (cit. on pp. 5, 87, 92, 101, 109, 112).

[Tsang 2020] Mankei Tsang, Francesco Albarelli, and Animesh Datta. "Quantum Semiparametric Estimation". *Physical Review X* 10.3 (July 2020), p. 031023. DOI: 10.1103/PhysRevX.10.031023 (cit. on p. 86).

[Tsang 2021] Mankei Tsang. "Quantum limit to subdiffraction incoherent optical imaging. II. A parametric-submodel approach". *Physical Review A* 104.5 (Nov. 2021), p. 052411. DOI: 10.1103/PhysRevA.104.052411 (cit. on p. 86).

[Tse 2019] M. Tse et al. "Quantum-Enhanced Advanced LIGO Detectors in the Era of Gravitational-Wave Astronomy". *Physical Review Letters* 123.23 (Dec. 2019), p. 231107. DOI: 10.1103/PhysRevLett.123.231107 (cit. on p. 128).

[Usenko 2018] Vladyslav C. Usenko. "Unidimensional continuous-variable quantum key distribution using squeezed states". *Physical Review A* 98.3 (Sept. 2018), p. 032321. DOI: 10.1103/PhysRevA.98.032321 (cit. on p. 128).

[Vahlbruch 2016] Henning Vahlbruch et al. "Detection of 15 dB Squeezed States of Light and their Application for the Absolute Calibration of Photoelectric Quantum Efficiency". *Physical Review Letters* 117.11 (Sept. 2016), p. 110801. DOI: 10.1103/PhysRevLett.117.110801 (cit. on p. 129).

[Valencia 2005] Alejandra Valencia et al. "Two-Photon Imaging with Thermal Light". *Physical Review Letters* 94.6 (Feb. 2005), p. 063601. DOI: 10.1103/PhysRevLett.94.063601 (cit. on p. 3).

[Villiers 2016] Geoffrey de Villiers and E. Roy Pike. *The Limits of Resolution*. CRC Press, Oct. 2016 (cit. on pp. 2, 70).

[Virgo Collaboration 2019] Virgo Collaboration et al. "Increasing the Astrophysical Reach of the Advanced Virgo Detector via the Application of Squeezed Vacuum States of Light". *Physical Review Letters* 123.23 (Dec. 2019), p. 231108. DOI: 10.1103/PhysRevLett.123.231108 (cit. on p. 128).

[Vlasenko 2020] S. Vlasenko et al. "Optimal correlation order in superresolution optical fluctuation microscopy". *Physical Review A* 102.6 (Dec. 2020), p. 063507. DOI: 10.1103/PhysRevA.102.063507 (cit. on p. 3).

[Vogel 2006] Werner Vogel and Dirk-Gunnar Welsch. *Quantum Optics*. John Wiley & Sons, Aug. 2006 (cit. on p. 20).

[Wadood 2024] S. A. Wadood et al. "Experimental demonstration of quantum-inspired optical symmetric hypothesis testing". *Optics Letters* 49.3 (Feb. 2024), pp. 750–753. DOI: 10.1364/OL.512320 (cit. on pp. 75, 86).

[Walschaers 2021] Mattia Walschaers. "Non-Gaussian Quantum States and Where to Find Them". *PRX Quantum* 2.3 (Sept. 2021), p. 030204. DOI: 10.1103/PRXQuantum.2.030204 (cit. on p. 26).

[Wang 2023] Ben Wang et al. "Quantum-limited resolution of partially coherent sources". *Chinese Optics Letters* 21.4 (Mar. 2023), p. 042601. DOI: 10.3788/COL202321.042601 (cit. on pp. 5, 87, 109, 115).

[Weedbrook 2012] Christian Weedbrook et al. "Gaussian quantum information". *Reviews of Modern Physics* 84.2 (May 2012), pp. 621–669. DOI: 10.1103/RevModPhys.84.621 (cit. on p. 128).

[Wetzstein 2011] Gordon Wetzstein et al. "Computational Plenoptic Imaging". *Computer Graphics Forum* 30.8 (2011), pp. 2397–2426. DOI: 10.1111/j.1467-8659.2011.02073.x (cit. on p. 4).

[Wick 1950] G. C. Wick. "The Evaluation of the Collision Matrix". *Physical Review* 80.2 (Oct. 1950), pp. 268–272. DOI: 10.1103/PhysRev.80.268 (cit. on p. 109).

[Wiener 1928] Norbert Wiener. "Coherency Matrices and Quantum". *Journal of Mathematics and Physics* 7.1-4 (1928), pp. 109–125. DOI: 10.1002/sapm192871109 (cit. on p. 11).

[Xu 2014] Li Xu et al. "Deep Convolutional Neural Network for Image Deconvolution". *Advances in Neural Information Processing Systems*. Vol. 27. Curran Associates, Inc., 2014 (cit. on p. 3).

[Yang 2017] Fan Yang et al. "Fisher information for far-field linear optical superresolution via homodyne or heterodyne detection in a higher-order local oscillator mode". *Physical Review A* 96.6 (Dec. 2017), p. 063829. DOI: 10.1103/PhysRevA.96.063829 (cit. on pp. 78, 85, 87).

[Yonezawa 2010] H. Yonezawa and A. Furusawa. "Continuous-variable quantum information processing with squeezed states of light". *Optics and Spectroscopy* 108.2 (Feb. 2010), pp. 288–296. DOI: 10.1134/S0030400X10020189 (cit. on p. 128).

[Yoshizawa 2017] Toru Yoshizawa. *Handbook of Optical Metrology: Principles and Applications, Second Edition*. CRC Press, July 2017 (cit. on p. 1).

[Yu 2018] Zhixian Yu and Sudhakar Prasad. "Quantum Limited Superresolution of an Incoherent Source Pair in Three Dimensions". *Physical Review Letters* 121.18 (Oct. 2018), p. 180504. DOI: 10.1103/PhysRevLett.121.180504 (cit. on p. 86).

[Yu 2020] Juan Yu et al. "Quantum Enhanced Optical Phase Estimation With a Squeezed Thermal State". *Physical Review Applied* 13.2 (Feb. 2020), p. 024037. DOI: 10.1103/PhysRevApplied.13.024037 (cit. on p. 128).

[Zanforlin 2022] Ugo Zanforlin et al. "Optical quantum super-resolution imaging and hypothesis testing". *Nature Communications* 13.1 (Sept. 2022), p. 5373. DOI: 10.1038/s41467-022-32977-8 (cit. on p. 87).

[Zhang 2007] Bo Zhang, Josiane Zerubia, and Jean-Christophe Olivo-Marin. "Gaussian approximations of fluorescence microscope point-spread function models". *Applied Optics* 46.10 (Apr. 2007), pp. 1819–1829. DOI: 10.1364/AO.46.001819 (cit. on p. 72).

[Zhang 2023a] Huan Zhang et al. "Quantum superresolution for imaging two pointlike entangled photon sources". *Physical Review A* 108.3 (Sept. 2023), p. 033713. DOI: 10.1103/PhysRevA.108.033713 (cit. on p. 93).

[Zhang 2023b] Yuanhang Zhang and Nicolas K. Fontaine. *Multi-Plane Light Conversion: A Practical Tutorial*. May 2023. DOI: 10.48550/arXiv.2304.11323 (cit. on p. 80).

[Zhou 2019a] Sisi Zhou and Liang Jiang. "Modern description of Rayleigh's criterion". *Physical Review A* 99.1 (Jan. 2019), p. 013808. DOI: 10.1103/PhysRevA.99.013808 (cit. on pp. 4, 86, 87).

[Zhou 2019b] Yiyu Zhou et al. "Quantum-limited estimation of the axial separation of two incoherent point sources". *Optica* 6.5 (May 2019), pp. 534–541. DOI: 10.1364/OPTICA.6.000534 (cit. on p. 86).