이학박사학위논문

Quantum metrology using Gaussian states and efficient Bayesian error certification

가우시안 상태를 이용한 양자 계측과 효율적인 베이지안 오류 검정

2020년 2월

서울대학교 대학원
물리천문학부
오창훈

Quantum metrology using Gaussian states and efficient Bayesian error certification

가우시안 상태를 이용한 양자 계측과 효율적인 베이지안 오류 검정

지도교수 정현석

이 논문을 이학박사 학위논문으로 제출함

2019년 12월

서울대학교 대학원

물리천문학부

오창훈

오창훈의 박사 학위논문을 인준함

2019년 12월

위 원 장	안 경 원	(인)
부위원장	정 현 석	(인)
위 원	신 용 일	(인)
위 원	김 도 헌	(인)
위 원	이 진 형	(인)

Quantum metrology using Gaussian states and efficient Bayesian error certification

Changhun Oh

Supervised by

Professor **Hyunseok Jeong**

A Dissertation

Submitted to the Faculty of

Seoul National University

in Partial Fulfillment of

the Requirements for the Degree of

Doctor of Philosophy

Dec. 2019

Department of Physics and Astronomy

The Graduate School

Seoul National University

Abstract

Quantum metrology using Gaussian states and efficient Bayesian error certification

Changhun Oh
Department of Physics and Astronomy
The Graduate School
Seoul National University

Precise measurement of physical quantities plays a crucial role in the development of science and technology. The main purpose of the dissertation is two-fold: to investigate the ultimate precision for estimation of physical quantities using Gaussian states and to propose an efficient method for certification of Bayesian error region in general quantum parameter estimation.

In the first part, we begin with analyzing sensitivity for estimating a phase difference in an optical interferometer. Optical interferometry is widely used in science and industry for measuring small displacements. Recently, a large-scale optical interferometer so-called the Laser interferometer Gravitational-Wave Observatory (LIGO) has succeeded in detecting a gravitational wave, the signal of which is extremely small. On the other hand, it has been shown that a non-classical feature of quantum states can improve the sensitivity of estimation, such as in optical interferometer, including the LIGO. From a practical point of view, we inspect the practically achievable precision using non-classical Gaussian states in Mach-Zehnder interferometer with feasible measurements and realistic photon loss. We then investigate the precision of single-mode phase estimation using Gaussian measurement, which can be realized by using homodyne detection, and

show that non-Gaussian measurement is necessary to utilize the power of Gaussian input probes optimally. Finally, we find the optimal measurement for general Gaussian quantum metrology and identify three distinct optimal measurements corresponding to different circumstances.

In the second part, we study the Bayesian error region, which is a crucial concept for a general estimation process. When estimating a physical quantity, one has to supply the error interval (single-parameter) or error region (multi-parameter) as well as the estimate. However, it has been shown that as the dimension of quantum systems of interest grows, it becomes intractable to calculate the size and credibility of Bayesian error regions. As an alternative, we derive an analytical expression for the properties, the size and credibility, of Bayesian error regions, in an asymptotic regime. We then propose an efficient numerical method to calculate them for high-dimensional quantum systems even in a non-asymptotic regime.

Keywords : Quantum Metrology, Gaussian quantum information, Bayesian credible region

Student Number : 2014-22371

Contents

List of Figures

Chapter 1

Introduction

Precise measurement of physical quantities is undoubtedly central for the development of science and technology. Especially, important recent discoveries in science, such as the observation of gravitational waves, have necessitated an extremely high precision physical device [1]. As much attention is paid on exploring a microscopic scale, the development of exceptionally accurate devices will be necessary for the advancement of science and industry in the future.

Ever since Carlton M. Caves, a pioneer of quantum metrology, proposed an enhanced optical interferometer [2], a remarkable breakthrough has started to emerge for developing high-resolution sensors through the power of quantum states [3, 4, 5]. As the power of quantum resources is established, such as coherence [6] and entanglement [7], and the experimental manipulation of quantum objects is accessible, the so-called second quantum revolution is in progress. The revolution is upgrading existing information processing into quantum information processing and classical computers into quantum computers. Particularly, quantum sensors are expected to be the first realizable quantum devices that will replace classical counterparts.

One critical enhancement that the revolution will bring is the precision of sensors. One of the most interesting improvement is optical interferometers [8], which are extensively used in science and industry for the measurement of tiny displacements. As the development of a large-scale interferometer enables detecting gravitational waves, quantum-enhanced gravitational wave detectors are being further established for detecting smaller signals which cannot be detected by the existing classical interferometers [2, 9]. Besides

optical interferometers, advanced atomic interferometers have also been proposed to detect a gravitational constant more precisely than conventional methods [10, 11]. Another implication of quantum-enhanced sensors is a magnetometer that is used for measuring nuclear magnetic resonance (NMR) [12]. NMR spectroscopy is evidently a core technique for medical diagnosis and biochemical studies.

In this dissertation, we investigate practically attainable precision in quantum metrology, such as optical interferometers using Gaussian states. Quantum metrology is the study of the improvement of sensors by using quantum resources. In the first chapter, we study quantum estimation using Gaussian states. Gaussian states have been paid attention due to their simple mathematical structure and relatively easy experimental generation and manipulation. In addition, they have been used for testing various fundamental quantum features such as Bell violation [13] and entanglement [14] and enable diverse quantum information processing such as quantum computing [15]. We investigate a practically achievable sensitivity using Gaussian states and find the optimal setup to exploit the ultimate power of Gaussian states in general Gaussian quantum metrology.

In the second chapter, we study Bayesian error certification [16, 17, 18] in quantum estimation theory. Providing an error interval or error region together with the estimate is necessary to show the reliability of the estimate. However, when inspecting many parameters in high-dimensional quantum systems, calculating an error region and its credibility is demanding because typically the error region is extremely small compared to the entire space. Furthermore, due to the positivity constraint in the case of quantum state tomography, precise calculation of size and credibility of Bayesian is in general NP-hard [19]. In order to resolve the difficulty, we derive an approximated analytical expression of the size and credibility of the credible region in an asymptotic regime. We then propose a novel numerical method to compute them efficiently.

This dissertation is based on the following publications:

- C. Oh, S.-Y. Lee, H. Nha, and H. Jeong, "Practical resources and measurements for lossy optical quantum metrology," Phys. Rev. A **96,** 062304 (2017).

- Y. S. Teo, C. Oh, and H. Jeong, "Bayesian error regions in quantum estimation I: analytical reasonings," New J. Phys. **20,** 093009 (2018).

- C. Oh, Y. S. Teo, and H. Jeong, "Bayesian error regions in quantum estimation II: region accuracy and adaptive methods," New J. Phys. **20,** 093010 (2018).

- C. Oh, C. Lee, C. Rockstuhl, H. Jeong, J. Kim, H. Nha, and S.-Y. Lee, "Optimal Gaussian measurements for phase estimation in single-mode Gaussian metrology," npj Quantum Information **5,** 10 (2019).

- C. Oh, C. Lee, L. Banchi, S.-Y. Lee, C. Rockstuhl, and H. Jeong, "Optimal measurements for quantum fidelity between Gaussian states and its relevance to quantum metrology," Phys. Rev. A **100,** 012323 (2019).

- C. Oh, Y. S. Teo, and H. Jeong, "Probing Bayesian credible regions intrinsically: a feasible error certification for physical systems," Phys. Rev. Lett. **123,** 040602 (2019).

- C. Oh, Y. S. Teo, and H. Jeong, "Efficient Bayesian credible-region certification for quantum-state tomography," Phys. Rev. A **100,** 012345 (2019).

Chapter 2

Preliminary

2.1 Continuous variable system

The n-mode bosonic continuous variable system is described by the Hamiltonian of n harmonic oscillators,

$$\hat{H} = \sum_{i=1}^{n} \hat{H}_i = \sum_{i=1}^{n} \hbar\omega_i \left(\hat{a}_i^\dagger \hat{a}_i + \frac{1}{2} \right), \tag{2.1}$$

where $\hbar$ is the reduced Planck constant, ω_i is the frequency of the mode, and $\hat{a}_i(\hat{a}_i^\dagger)$ is called a annihilation (creation) operator of ith mode. The annihilation and creation operators satisfy the following commutation relation,

$$[\hat{a}_i, \hat{a}_j^\dagger] = \delta_{ij}. \tag{2.2}$$

For simplicity, we set $\hbar = \omega = 1$ throughout the dissertation. The Hamiltonian of the each mode defines the Fock basis $\{|n_i\rangle\}$ such that

$$\hat{H}_i |n_i\rangle = \left(\hat{a}_i^\dagger \hat{a}_i + \frac{1}{2} \right) |n_i\rangle = \left(n_i + \frac{1}{2} \right) |n_i\rangle. \tag{2.3}$$

The annihilation and creation operators transform the Fock basis as

$$\hat{a}_i |n_i\rangle = \sqrt{n_i - 1}\,|n_i - 1\rangle, \quad \hat{a}_i^\dagger |n_i\rangle = \sqrt{n_i + 1}\,|n_i + 1\rangle. \tag{2.4}$$

The annihilation and creation operators define the quadrature operators,

$$\hat{x}_i = \frac{\hat{a}_i + \hat{a}_i^\dagger}{\sqrt{2}}, \quad \hat{p}_i = \frac{\hat{a}_i - \hat{a}_i^\dagger}{\sqrt{2}i}, \tag{2.5}$$

which obey the canonical commutation relation (CCR),

$$[\hat{x}_i, \hat{p}_j] = i\delta_{ij}. \tag{2.6}$$

If we define the n-mode quadrature operator vector,

$$\hat{Q} = (\hat{x}_1, \hat{p}_1, ..., \hat{x}_n, \hat{p}_n)^{\mathrm{T}}, \tag{2.7}$$

the CCR is simply written as

$$[\hat{Q}_j, \hat{Q}_k] = i\Omega_{jk}, \quad \Omega = \mathbb{1}_n \otimes \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{2.8}$$

where $\mathbb{1}_n$ is the $n \times n$ identity matrix. Transformations of the quadrature operators that preserve the CCR are called symplectic transformations represented by symplectic matrices S,

$$\hat{Q} \to \hat{U}^\dagger \hat{Q} \hat{U} = S\hat{Q}, \quad S^{\mathrm{T}} \Omega S = \Omega. \tag{2.9}$$

Another transformation preserving the CCR is the so-called displacement operation or Weyl transformation, which displaces the quadrature operators such as

$$\hat{Q} \to \hat{V}^\dagger(\vec{\xi}) \hat{Q} \hat{V}(\vec{\xi}) = \hat{Q} + (\Omega\vec{\xi})^{\mathrm{T}}, \quad \hat{V}(\vec{\xi}) = \exp(i\vec{\xi}^{\mathrm{T}}\hat{Q}), \tag{2.10}$$

where $\vec{\xi}$ is a $2n$-dimensional real vector. Equivalently, we can define displacement

operators of n-dimensional complex vector arguments $\vec{\alpha} = (\alpha_1, \alpha_2, ..., \alpha_n)^{\mathrm{T}}$,

$$\hat{D}(\vec{\alpha}) = \exp\left[\sum_{i=1}^{n}\left(\alpha_i \hat{a}_i^{\dagger} - \alpha_i^{*}\hat{a}_i\right)\right], \tag{2.11}$$

which displaces the annihilation operator as

$$\hat{a}_i \to \hat{D}^{\dagger}(\vec{\alpha})\hat{a}_i\hat{D}(\vec{\alpha}) = \hat{a}_i + \alpha_i. \tag{2.12}$$

Throughout the dissertation, when a $2n$-dimensional real-vector argument $\vec{d}$ instead of a n-dimensional complex-vector $\vec{\alpha}$ is applied to the displacement operator $\hat{D}(\cdot)$, it is understood as displacing the quadrature vector as

$$\hat{Q} \to \hat{Q} + \vec{d}, \tag{2.13}$$

which is equivalent to the Weyl transformation $\hat{V}(-\Omega\vec{d})$. If a displacement operator is applied to a vacuum state, then we obtain a so-called coherent state [20],

$$\hat{D}(\vec{\alpha})|0\rangle^{\otimes n} = |\alpha_1\rangle|\alpha_2\rangle...|\alpha_n\rangle = |\vec{\alpha}\rangle, \tag{2.14}$$

where each coherent state is an eigenstate of the corresponding annihilation operator, $\hat{a}_i|\alpha_i\rangle = \alpha_i|\alpha_i\rangle$. The displacement operator has the following properties,

$$\hat{D}^{\dagger}(\vec{\alpha}) = \hat{D}(-\vec{\alpha}), \tag{2.15}$$

$$\mathrm{Tr}[\hat{D}(\vec{\alpha})] = \pi^n \delta^{(2n)}(\vec{\alpha}), \tag{2.16}$$

$$\hat{D}(\vec{\alpha}_1)\hat{D}(\vec{\alpha}_2) = D(\vec{\alpha}_1 + \vec{\alpha}_2)\exp\left[-\frac{1}{2}(\vec{\alpha}_1^{\dagger}\vec{\alpha}_2 - \vec{\alpha}_2^{\dagger}\vec{\alpha}_1)\right]. \tag{2.17}$$

The last property can be obtained by using the Baker-Campbell-Hausdorff formula.

The following formula connects displacement operator with functions of the number operator,

$$\nu^{\hat{a}^\dagger \hat{a}} = \int_{\mathbb{C}} \frac{d^2 z}{\pi(1-\nu)} \exp\left[-\frac{1}{2}\frac{1+\nu}{1-\nu}|z|^2 \right] \hat{D}(z), \tag{2.18}$$

with special cases

$$|0\rangle\langle 0| = \int_{\mathbb{C}} \frac{d^2 z}{\pi} \exp\left[-\frac{1}{2}|z|^2 \right] \hat{D}(z), \quad \hat{\Pi} \equiv (-1)^{\hat{a}^\dagger \hat{a}} = \int_{\mathbb{C}} \frac{d^2 z}{2\pi} \hat{D}(z). \tag{2.19}$$

One useful property of the parity operator $\hat{\Pi}$ is $\hat{\Pi}\hat{D}(\alpha)\hat{\Pi} = \hat{D}(-\alpha)$.

The set of Weyl operators $\hat{V}(\vec{\xi})$ is complete, in a sense that any operator $\hat{O}$ can be written as [21]

$$\hat{O} = \int_{\mathbb{R}^{2n}} \frac{d^{2n}\vec{\xi}}{(2\pi)^n} \mathrm{Tr}[\hat{O}\hat{V}(\vec{\xi})]\hat{V}^\dagger(\vec{\xi}), \tag{2.20}$$

where we define $\chi[\hat{O}](\vec{\xi}) \equiv \mathrm{Tr}[\hat{O}\hat{V}(\vec{\xi})]$ as the characteristic function of the operator $\hat{O}$. Equivalently, the set of displacement operators is complete as

$$\hat{O} = \int_{\mathbb{C}^n} \frac{d^{2n}\vec{\alpha}}{\pi^n} \mathrm{Tr}[\hat{O}\hat{D}(\vec{\alpha})]\hat{D}^\dagger(\vec{\alpha}). \tag{2.21}$$

Here the characteristic function of an operator $\hat{O}$ is $\chi[\hat{O}](\vec{\alpha}) = \mathrm{Tr}[\hat{O}\hat{D}(\vec{\alpha})]$. The characteristic function generates moments of symmetrized quadratures for a density operator $\hat{\rho}$ of a single-mode quantum state,

$$\frac{\partial^{p+q}}{\partial\xi_1^p \partial\xi_2^q} \frac{1}{i^{p+q}} \chi[\hat{\rho}](\vec{\xi})|_{\vec{\xi}=0} = \mathrm{Tr}\left[\hat{\rho}[\hat{x}^p \hat{p}^q]_{\mathrm{S}}\right], \tag{2.22}$$

which can be shown by using the definition of the Weyl operator as

$$\begin{aligned}\hat{V}(\xi) &= \sum_{k=0}^{\infty} \frac{1}{k!}(i\xi_1\hat{x} + i\xi_2\hat{p})^k = \sum_{k=0}^{\infty} \frac{i^k}{k!} \sum_{l=0}^{k} \binom{k}{l} \xi_1^l \xi_2^{k-l} [\hat{x}^l \hat{p}^{k-l}]_S \\ &= \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{\xi_1^k \xi_2^l}{k!l!} i^{k+l} [\hat{x}^k \hat{p}^l]_S. \end{aligned} \tag{2.23}$$

The generalization to an arbitrary number of modes is straightforward. Similarly, for a single-mode complex-displacement operator $\hat{D}(\alpha)$, moments of symmetric operators of $\hat{a}$ and $\hat{a}^\dagger$ can be generated by the characteristic function as

$$\mathrm{Tr}[\hat{\rho}[(\hat{a}^\dagger)^p(\hat{a})^q]_S] = (-1)^q \frac{\partial^{p+q}}{\partial\alpha_k^p \partial\alpha_l^{*q}} \chi(\alpha)|_{\alpha=0} \tag{2.24}$$

which can be shown as

$$\begin{aligned}\hat{D}(\alpha) &= \sum_{k=0}^{\infty} \frac{1}{k!}(\alpha\hat{a}^\dagger - \alpha^*\hat{a})^k = \sum_{k=0}^{\infty} \frac{(-1)^l}{k!} \sum_{l=0}^{k} \binom{k}{l} \alpha^{k-l}\alpha^{*l} [(\hat{a}^\dagger)^{k-l}\hat{a}^l]_S \\ &= \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} (-1)^l \frac{\alpha^k \alpha^{*l}}{k!l!} [(\hat{a}^\dagger)^{k-l}\hat{a}^l]_S. \end{aligned} \tag{2.25}$$

Finally, we introduce the Wigner function of an operator $\hat{O}$ that is defined as the Fourier transformation of the characteristic function,

$$W[\hat{O}](\vec{\beta}) = \int_{\mathbb{C}^n} \frac{d^{2n}\vec{\alpha}}{\pi^{2n}} \exp(\vec{\alpha}^\dagger\vec{\beta} - \vec{\beta}^\dagger\vec{\alpha})\chi[\hat{O}](\vec{\alpha}). \tag{2.26}$$

Similar to characteristic function, any operator can be written by Wigner function as

$$\hat{O} = 2^n \int_{\mathbb{C}} d^{2n}\vec{\alpha} W(\vec{\alpha}) \hat{D}(\vec{\alpha}) \hat{\Pi} \hat{D}^\dagger(\vec{\alpha}) \tag{2.27}$$

In addition, Wigner function of any operator can be written as

$$W[\hat{O}](\vec{\beta}) = \left(\frac{2}{\pi}\right)^n \mathrm{Tr}[\hat{O}\hat{D}(\vec{\beta})\hat{\Pi}\hat{D}^\dagger(\vec{\beta})]. \tag{2.28}$$

Note that all the properties presented for complex-displacement operators can be equivalently written for Weyl operators and the Wigner function in terms of real numbers is defined as

$$W[\hat{O}](\vec{x}) = \int_{\mathbb{R}^{2n}} \frac{d^{2n}\xi}{(2\pi)^{2n}} \exp(-i\vec{x}\cdot\vec{\xi})\chi[\hat{O}](\xi). \tag{2.29}$$

2.2 Gaussian states

Gaussian states are an important class of quantum states in the continuous variable system not only because their mathematical structure is simple to analyze their properties but also because they are relatively easy to generate in practice [22, 23, 24, 25, 26].

They are defined as states whose Wigner function follows Gaussian distribution,

$$W[\hat{\rho}](\vec{x}) = \frac{1}{(2\pi)^n\sqrt{|\Gamma|}} \exp\left[-\frac{1}{2}\left(\vec{x}-\vec{d}\right)^{\mathrm{T}}\Gamma^{-1}\left(\vec{x}-\vec{d}\right)\right], \tag{2.30}$$

or, equivalently, states that have Gaussian characteristic function,

$$\chi[\hat{\rho}](\vec{\xi}) = \exp\left[-\frac{1}{2}\vec{\xi}^{\mathrm{T}}\Gamma\vec{\xi} + i\vec{d}^{\mathrm{T}}\xi\right]. \tag{2.31}$$

Thus, Gaussian states are fully characterized by their covariance matrix Γ and the first moment vector $\vec{d}$, which are defined as $\Gamma_{ij} = \mathrm{Tr}[\hat{\rho}\{\hat{Q}_i - d_i, \hat{Q}_j - d_j\}]/2$ and $d_i = \mathrm{Tr}[\hat{\rho}\hat{Q}_i]$. Here, $\{\hat{A},\hat{B}\} \equiv \hat{A}\hat{B} + \hat{B}\hat{A}$. The covariance matrix represents a legitimate

quantum state ($\hat{\rho} \geq 0$ and $\mathrm{Tr}[\hat{\rho}] = 1$) if and only if [27, 28]

$$\Gamma + \frac{i}{2}\Omega \geq 0. \tag{2.32}$$

Since the covariance matrix Γ is real symmetric, it can be diagonalized by a symplectic matrix S according to Williamson decomposition such that [27, 28]

$$\Gamma = SDS^{\mathrm{T}}, \tag{2.33}$$

where the diagonal matrix D has a form of $(d_1, d_1, d_2, d_2, ..., d_n, d_n)$, which is the so-called Williamson decomposition. Noting that a single-mode Gaussian state of the covariance matrix of $\mathrm{diag}(d, d)$ represents a thermal state of a mean photon number $\bar{n} = (2d-1)/2$,

$$\hat{\rho}_{\mathrm{T}} = \sum_{k=0}^{\infty} \frac{1}{\bar{n}+1} \left(\frac{\bar{n}}{\bar{n}+1} \right)^k |k\rangle\langle k|, \tag{2.34}$$

and thus D represents a product of thermal states, arbitrary multimode Gaussian states can be written as

$$\hat{\rho}_{\mathrm{G}} = \hat{U}_{\mathrm{S}} \left(\otimes_{i=1}^{n} \hat{\rho}_{\mathrm{T}i} \right) \hat{U}_{\mathrm{S}}^{\dagger}, \tag{2.35}$$

where $\hat{U}_{\mathrm{S}}$ is a unitary operation corresponding to the symplectic matrix S by the relation Eq. (2.9). Furthermore, by Bloch-Messiah decomposition

$$S = O_2 \left[\oplus_{i=1}^{n} \mathrm{diag}(s_i, s_i^{-1}) \right] O_1 \tag{2.36}$$

with orthogonal matrices O_1 and O_2 and $s_i > 0$ for all i, arbitrary multimode Gaus-

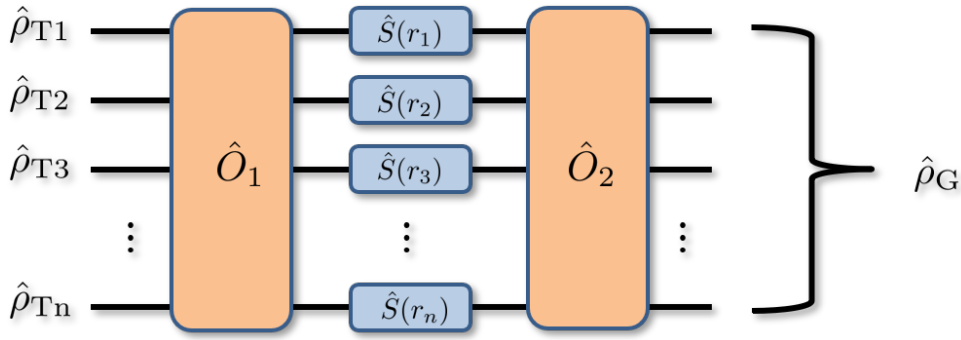


Figure 1: Generation of arbitrary Gaussian states.

sian states can be written as

$$\hat{\rho}_{\mathrm{G}} = \hat{O}_2 \left[\otimes_{i=1}^{n} \hat{S}(r_i)\right] \hat{O}_1 \left(\otimes_{i=1}^{n} \hat{\rho}_{\mathrm{T}i}\right) \hat{O}_1^{\dagger} \left[\otimes_{i=1}^{n} \hat{S}^{\dagger}(r_i)\right] \hat{O}_2^{\dagger}, \tag{2.37}$$

where $\hat{O}_1$ and $\hat{O}_2$ are a combination of beam splitters and phase shifters, corresponding to the orthogonal matrices O_1 and O_2, and $\hat{S}(r_i)$ is a squeezing operator corresponding to a diagonal symplectic matrix $\mathrm{diag}(s_i, 1/s_i)$, which are depicted in Fig. 1. Since the implementation of beam splitters and phase shifters in practice is simple, the only obstacle is to execute squeezing operation of high squeezing parameters. By the virtue of the decomposition, one can easily see that any single-mode Gaussian states can be written as

$$\hat{\rho} = \hat{D}(\alpha)\hat{S}(\xi)\hat{\rho}_{\mathrm{T}}\hat{S}^{\dagger}(\xi)\hat{D}^{\dagger}(\alpha), \tag{2.38}$$

where $\alpha = |\alpha|e^{i\theta_{\mathrm{c}}}, \xi = re^{i\theta_{\mathrm{s}}} \in \mathbb{C}$ are displacement and squeezing parameters.

2.3 Quantum estimation theory

Precise measurement of physical quantities is essential in development of science and technology. In this section, we exhibit a gist of quantum estimation theory.

Let us begin with an arbitrary situation in which one wants to estimate a physical quantity of interest. A general estimation procedure can be described as following. There is a physical quantity of interest which changes our probe so that we obtain the information of the quantity by analyzing the transformed probe. If the transformation or the result is probabilistic, we may describe the probability of obtaining the outcome x when the true value θ_{True} of the quantity of interest is a certain value θ by a conditional probability $p(x|\theta_{\text{True}} = \theta)$. After collecting outcomes $\mathbb{D} = \{x_1, x_2, ..., x_N\}$, we process the data to estimate the true value by constructing an estimator, which is a function of data,

$$\hat{\theta} = \hat{\theta}(\mathbb{D}). \tag{2.39}$$

If we use an unbiased estimator, i.e., $\langle \hat{\theta}(\mathbb{D}) \rangle_{\mathbb{D}} = \theta_{\text{True}}$ for all θ_{True}, the mean squared error of the estimator is bounded below by the inverse of the so-called (classical) Fisher information [29],

$$\Delta^2 \hat{\theta} \equiv \langle (\hat{\theta} - \theta)^2 \rangle \geq \frac{1}{NF(\theta)}, \tag{2.40}$$

which is called the Cramér-Rao inequality and N is the number of copies. The overall estimation procedure and the Cramér-Rao inequality are illustrated in Fig. 2. The inequality can be asymptotically saturated by using the maximum likelihood estimator [30].

If we employ quantum resources such a way that a quantum state $\hat{\rho}$ is used as an input probe and evolves by a quantum operation $\mathcal{E}_\theta$ encoding a quantity of interest θ and is measured by POVM $\{\hat{E}_x\}$, the conditional probability, which is essential

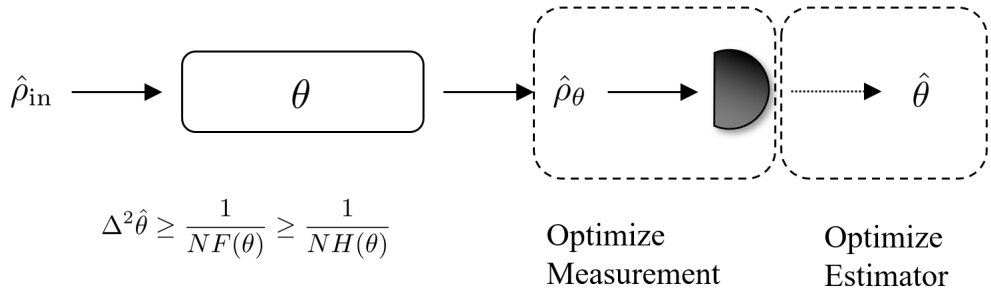


Figure 2: General quantum estimation procedure and Quantum Cramér-Rao inequality. An input probe $\hat{\rho}_{\rm in}$ evolves by a quantum operation encoding the information of an unknown paramter θ and is measured. Finally, the measurement outcome is post-processed to obtain an estimate. We can optimize the estimator to reduce the estimation error with fixing a measurement, which gives classical Cramér-Rao inequality. If we further optimize the measurement, we finally obtain the quantum Cramér-Rao inequality. See the main text for the details.

for the estimation, can be described as

$$p(x|\theta) = \mathrm{Tr}\left[\hat{\rho}_\theta \hat{E}_x\right]. \tag{2.41}$$

where we defined $\hat{\rho}_\theta = \mathcal{E}_\theta(\hat{\rho})$. In the quantum case, we can choose our input probe and the measurement to change $p(x|\theta)$ which is related to the Fisher information. For a fixed input probe $\hat{\rho}$, the Fisher information can be maximized over all possible measurement $\{\hat{E}_x\}$,

$$H(\theta) = \max_{\{\hat{E}_x\}} F(\theta), \tag{2.42}$$

where $H(\theta)$ is called the quantum Fisher information [31]. Mathematically, it is simplified as

$$H(\theta) = \mathrm{Tr}\left[\hat{\rho}_\theta \hat{L}_\theta^2\right], \tag{2.43}$$

where $\hat{L}_\theta$ is the so-called symmetric logarithmic derivative (SLD) operator satisfying

$$\frac{\partial\hat{\rho}_\theta}{\partial\theta} = \frac{\hat{\rho}_\theta\hat{L}_\theta + \hat{L}_\theta\hat{\rho}_\theta}{2}. \tag{2.44}$$

If we write the state in a diagonal form $\hat{\rho}_\theta = \sum_i p_i|\psi_i\rangle\langle\psi_i|$, the SLD operator can be written as

$$\hat{L}_\theta = 2\sum_{i,j}\frac{\langle\psi_i|\partial_\theta\hat{\rho}_\theta|\psi_j\rangle}{p_i+p_j}|\psi_i\rangle\langle\psi_j|. \tag{2.45}$$

The quantum Fisher information is thus given by

$$H(\theta) = 2\sum_{i,j}\frac{|\langle\psi_i|\partial_\theta\hat{\rho}_\theta|\psi_j\rangle|^2}{p_i+p_j}. \tag{2.46}$$

When the input state is pure, $\hat{\rho} = |\psi\rangle\langle\psi|$, and the quantum operation encoding the parameter θ is unitary, $\mathcal{E}_\theta(\cdot) = \hat{U}_\theta\cdot\hat{U}_\theta^\dagger = e^{-i\hat{H}\theta}\cdot e^{i\hat{H}\theta}$, the quantum Fisher information is further simplified as [32]

$$H(\theta) = 4\langle\Delta^2\hat{H}\rangle_{|\psi\rangle}. \tag{2.47}$$

The optimal measurement achieving Eq. (2.42) has to satisfy [33]

$$\hat{E}_x^{1/2}(\hat{\rho}_\theta^{1/2} - \lambda_x\hat{L}_\theta\hat{\rho}_\theta^{1/2}) = 0, \tag{2.48}$$

$$\mathrm{Tr}[\hat{E}_x\hat{\rho}_\theta\hat{L}_\theta] \in \mathbb{R}, \tag{2.49}$$

If one prepares measurement consisting of projection operators onto the eigenbasis of the SLD operator, it achieves the maximum sensitivity that quantum Fisher

information gives.

Chapter 3

Quantum Metrology using Gaussian states

The contents of the present chapter are largely based on the papers [C. Oh, S.-Y. Lee, H. Nha, and H. Jeong, "Practical resources and measurements for lossy optical quantum metrology," Phys. Rev. A **96,** 062304 (2017)] Ref. [34], [C. Oh, C. Lee, C. Rockstuhl, H. Jeong, J. Kim, H. Nha, and S.-Y. Lee, "Optimal Gaussian measurements for phase estimation in single-mode Gaussian metrology," npj Quantum Information **5,** 10 (2019)] Ref. [35], and [C. Oh, C. Lee, L. Banchi, S.-Y. Lee, C. Rockstuhl, and H. Jeong, "Optimal measurements for quantum fidelity between Gaussian states and its relevance to quantum metrology," Phys. Rev. A **100,** 012323 (2019)] Ref. [36].

3.1 Introduction

Gaussian states have been considered as important resources for precision measurement since Carlton M. Caves, a pioneer of quantum metrology, proposed to use a squeezed vacuum state in a gravitational wave detector for improving the sensitivity of the Michelson interferometer [2]. A conventional Michelson interferometer operates by injecting a laser on only one input arm and letting the other arm empty. Having realized that a vacuum fluctuation on the empty arm causes a severe uncertainty on the phase sensitivity, he proposed to inject a squeezed vacuum state, which has a property that the uncertainty on an axis is reduced while that on the other axis is inevitably amplified due to Heisenberg's uncertainty principle [37]. By

the virtue of the reduced uncertainty on an axis, consequently, the error caused by the photon-number uncertainty is significantly reduced whereas that induced by a radiation pressure only slightly increases. This proposal has enlightened physicists to begin developing many sensing devices using quantum resources to improve sensitivities.

Due to his proposal, a squeezed state, which is one of the most representative and important nonclassical states, begins to be considered as an essential resource in quantum metrology. Besides a coherent & squeezed vacuum (CSV) state that Caves proposed, many advantageous quantum states have been proposed for advanced optical interferometer such as NOON state [12, 38, 39], twin Fock state [40], two-mode squeezed vacuum (TMSV) state [41], and entangled coherent state [42, 43]. Although there are a plenty of proposals and experiments for generation of the useful nonclassical states [44, 45, 46], only a CSV state and a TMSV state, which are Gaussian states, are practically accessible for a sufficiently large number of photons.

First of all, in this chapter, we analyze and compare the phase sensitivity attained by a CSV state and a TMSV state in the Mach-Zehnder interferometer (MZI). Notice that the mathematical description of MZI is equivalent to that of the Michelson interferometer [47, 48]. We first derive the quantum Fisher information of the states for estimating a phase difference. We then compare classical Fisher information based on the parity measurement and homodyne detection. Homodyne detection is currently available in laboratory, and parity detection using photon-number resolving detection (PNRD) is expected to be accessible for low photon-number regime [49, 50]. Based on the analysis, we conclude that homodyne detection, which is a subset of Gaussian measurement, allows one to achieve high sensitivity and is robust to photon loss.

The conclusion motivates us to proceed to inspect the capability of general Gaussian measurement for single-mode phase estimation. Gaussian states are often cooperated with Gaussian measurements, and it has been shown that Gaussian measurements enable the full characterization of all Gaussian states [51]. Furthermore, they can be used for testing the inseparability of Gaussian states with a necessary and sufficient condition [52], and Gaussian measurements are sufficient to constitute the optimal set of POVMs for a minimization involved in the computation of quantum discord for Gaussian states [53, 54]. The main question that we tackle is whether or not general Gaussian measurement is optimal for any Gaussian input states for phase estimation. Our answer to the question is that Gaussian measurement is not optimal in general except for displaced thermal states and squeezed vacuum states. We also find the general optimal measurement for Gaussian single-mode phase estimation.

In the final section of chapter, we identify the optimal measurement for general Gaussian quantum metrology by finding the optimal measurement for quantum fidelity between general Gaussian states. Focusing on the single-mode case, we found three distinct optimal measurements depending on the circumstance. Finally, our method can be straightforwardly generalized to multimode cases for establishing the general optimal measurement beyond Gaussian measurement for Gaussian metrology.

3.2 Advanced Mach-Zehnder Interferometer

In this section, we begin with briefly explaining the structure of the Mach-Zehnder inteferometer (MZI), which is depicted in Fig. 3. The MZI consists of two 50:50 beam splitters, the first of which makes two paths to encode a phase difference

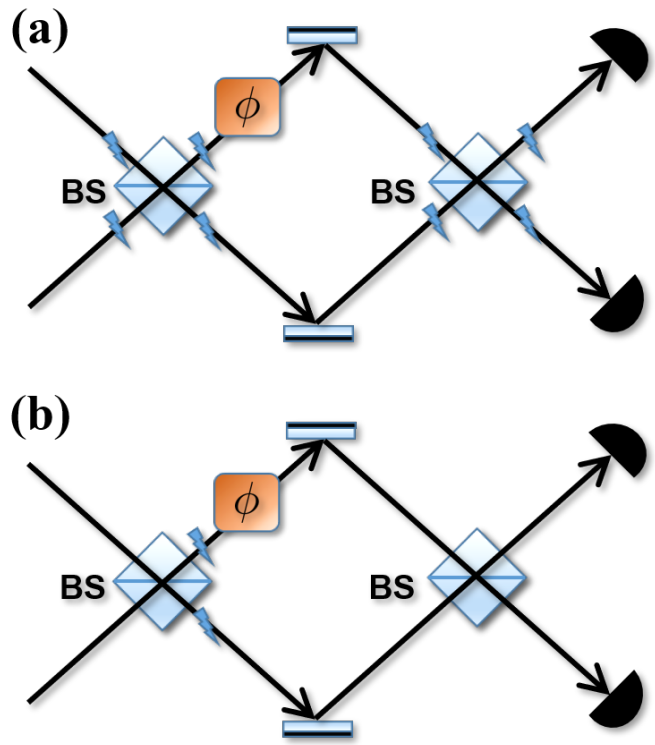


Figure 3: Lossy MZI considering all possible photon losses. When the photon loss rates at each arm are equal, we can simplify (a) to (b).

that we are interested in. A conventional MZI uses a laser, which is described as a coherent state in quantum optics, as an input state on one of two input ports and injects nothing on the other port. The coherent state is split to two coherent states with the equal amplitude by the first 50:50 beam splitter. Due to the difference of a path length of interest, two coherent states obtain different phases before they mix again by the second beam splitter. Finally, we measure the intensity difference on the two output ports by which we estimate the phase difference or the path length difference.

In the language of quantum optics, the MZI can be illustrated as follows. We have two input modes described by the annihilation operators $\hat{a}$ and $\hat{b}$ for each mode. The

input modes mix by the first beam splitter, which is described by a unitary operator

$$\hat{B}_1 = \exp\left[\frac{\pi}{4}(\hat{a}\hat{b}^\dagger - \hat{a}^\dagger\hat{b})\right]. \tag{3.1}$$

In the Heisenberg picture, the beam splitter transforms the mode operators as, by using Baker-Campbell-Hausdorff formula,

$$\hat{a} \to \hat{B}_1^\dagger \hat{a}\hat{B}_1 = \frac{\hat{a}+\hat{b}}{\sqrt{2}}, \quad \hat{b} \to \hat{B}_1^\dagger \hat{b}\hat{B}_1 = \frac{-\hat{a}+\hat{b}}{\sqrt{2}}. \tag{3.2}$$

The quantum state of light then evolves by the phase shifter of a relative phase ϕ between two modes,

$$\hat{P}(\phi) = \exp\left(i\phi\hat{a}^\dagger\hat{a}\right), \tag{3.3}$$

which transforms the mode operator as

$$\hat{a} \to \hat{P}^\dagger(\phi)\hat{a}\hat{P}(\phi) = \hat{a}e^{i\phi}, \quad \hat{b} \to \hat{P}^\dagger(\phi)\hat{b}\hat{P}(\phi) = \hat{b}. \tag{3.4}$$

Finally, after the second beam splitter $\hat{B}_2 = \hat{B}_1^\dagger$, we measure the average of an observable $\hat{O} = \hat{a}^\dagger\hat{a} - \hat{b}^\dagger\hat{b}$, which represents the intensity difference. Since the input state is $|\psi_{\text{in}}\rangle = |\alpha\rangle \otimes |0\rangle$, where $\alpha = |\alpha|e^{i\theta_c}$, one may derive the uncertainty by using the error propagation relation,

$$\Delta^2\phi = \frac{1}{N}\frac{\langle\hat{O}^2\rangle - \langle\hat{O}\rangle^2}{|\partial\langle\hat{O}\rangle/\partial\phi|^2} = \frac{1}{N|\alpha|^2\sin^2\phi}, \tag{3.5}$$

where N is the number of repetition. Thus, if the true value of the relative phase is around $\pi/2$, the uncertainty becomes $\Delta^2\phi = \left(N|\alpha|^2\right)^{-1} = (N\bar{n})^{-1}$, which is the so-called shot-noise limit. As mentioned before, the sensitivity can be improved by

using quantum resources such as squeezed states. If we replace the vacuum state by a squeezed state $|\xi\rangle$ with a squeezing parameter $\xi = re^{i\theta_s}$, we may obtain the uncertainty as [55],

$$\Delta^2\phi = \frac{1}{N}\left(\frac{|\alpha|^2 e^{-2r} + \sinh^2 r}{(|\alpha|^2 - \sinh^2 r)^2} + \frac{|\alpha|^2 + 2\sinh^2 r \cosh^2 r}{(|\alpha|^2 - \sinh^2 r)^2 \tanh^2 \phi}\right), \tag{3.6}$$

where we have chosen the optimal angles of the coherent state and the squeezed state. If we use a high intensity coherent state, $|\alpha|^2 \gg \sinh^2 r$, the uncertainty becomes

$$\Delta^2\phi = \frac{1}{N}\frac{e^{-2r}}{\bar{n}}, \tag{3.7}$$

where $\bar{n} = |\alpha|^2 + \sinh^2 r \simeq |\alpha|^2$. Moreover, if we employ a photon-number resolving detection instead of measuring the intensity difference of the output modes, the sensitivity improves further as [55],

$$\Delta^2\phi = \frac{1}{N}\frac{1}{|\alpha|^2 e^{2r} + \sinh^2 r}, \tag{3.8}$$

which reached the so-called Heisenberg scaling,

$$\Delta^2\phi = \frac{1}{N\bar{n}^2}, \tag{3.9}$$

when $|\alpha|^2 \simeq \sinh^2 r \simeq \bar{n}/2$ and $\bar{n}, N \gg 1$. Hence, the squeezed state significantly reduces the uncertainty of the estimation, and consequently classical resources cannot achieve the same uncertainty using the same energy. Thus, quantum resources such as squeezed states enable exceptionally precise estimation of phase. It is worth noting that as presented before, there have been many proposals for advantageous

quantum states such as NOON state [12, 38, 39], twin Fock state [40], TMSV state [41].

From a practical point of view, on the other hand, consideration of photon loss and the analysis of the robustness of the system is important since it is practically impossible to perfectly isolate the system from the environment [56, 57]. A description of photon-loss has been extensively studied, and the most frequently used model is the following master equation [56, 58, 59],

$$\frac{d\hat{\rho}(t)}{dt} = \frac{\gamma}{2}\left(\bar{n}_e L[\hat{a}^\dagger] + (\bar{n}_e + 1)L[\hat{a}^\dagger]\right)\hat{\rho}(t), \tag{3.10}$$

where $L[\hat{o}]\hat{\rho} = (2\hat{o}\hat{\rho}\hat{o}^\dagger - \hat{o}^\dagger\hat{o}\hat{\rho} - \hat{\rho}\hat{o}^\dagger\hat{o})$ with a damping rate of $\gamma \geq 0$, and $n_e \geq 0$ represents the average number of thermal photons of the environment. Note that Eq. (3.10) is written in the interaction picture. When we describe a pure-loss model, i.e. $\bar{n}_e = 0$, the photon loss can be equivalently described by a beam splitter model,

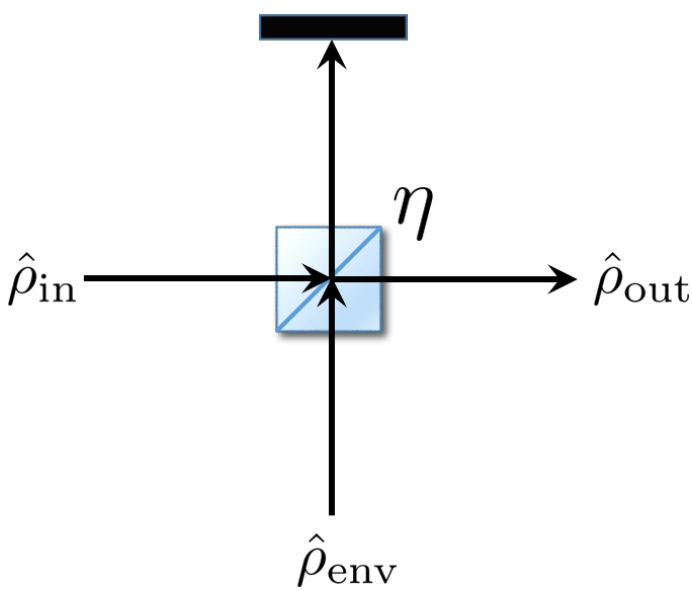


Figure 4: Beam splitter model for the description of photon loss. η is the transmissivity of the beam splitter.

which is shown in Fig. 4, where the state for the environment is usually described by the vacuum and one of two output modes is inaccessible. Here, the transmissivity η of the beam splitter is characterized by $\eta = e^{-\gamma t}$.

3.2.1 Comparison between Coherent & Squeezed vacuum state and two-mode squeezed vacuum state

In this section, we analyze the sensitivity of a CSV state and a TMSV state based on quantum Fisher information and classical Fisher information for parity measurement and homodyne measurement. In order to do that, we first derive the covariance matrix and the first-moment vector of the phase-encoded states. First of all, when the input state is $\hat{\rho}_{\text{in}}$, the phase-encoded state, which is the state before the second beam splitter, can be written as

$$\hat{\rho}(\phi) = \hat{P}(\phi)\hat{B}_1(\pi/4)\hat{\rho}_{\text{in}}\hat{B}_1^{\dagger}(\pi/4)\hat{P}^{\dagger}(\phi). \tag{3.11}$$

A CSV state and a TMSV state are respectively given by

$$|\psi_{\text{CSV}}\rangle = \hat{S}_1(r)\hat{D}_2(\alpha)|0\rangle_{12}, \quad |\psi_{\text{TMSV}}\rangle = \hat{S}_{12}(s)|0\rangle_{12}, \tag{3.12}$$

where $\hat{S}_1(r) = \exp[r(\hat{a}^{\dagger 2} - \hat{a}^2)/2]$ is a single-mode squeezing operator of a real squeezing parameter $r \in \mathbb{R}$ on the first mode, $\hat{D}_2(\alpha) = \exp(\alpha\hat{a}^{\dagger} - \alpha^*\hat{a})$ is a displacement operator of a displacement parameter $\alpha \in \mathbb{C}$, and $\hat{S}_{12}(s) = \exp[s(\hat{a}^{\dagger}\hat{b}^{\dagger} - \hat{a}\hat{b})]$ is a two-mode squeezing operator of a squeezing parameter $s \in \mathbb{R}$. The total average photon numbers of each state are given by $\bar{n}_{\text{CSV}} = \langle\psi_{\text{CSV}}|\hat{a}^{\dagger}\hat{a} + \hat{b}^{\dagger}\hat{b}|\psi_{\text{CSV}}\rangle = \alpha^2 + \sinh^2 r$ and $\bar{n}_{\text{TMSV}} = \langle\psi_{\text{TMSV}}|\hat{a}^{\dagger}\hat{a} + \hat{b}^{\dagger}\hat{b}|\psi_{\text{TMSV}}\rangle = 2\sinh^2 s$. Applying a unitary operation $\hat{U}$ on the density operator of a Gaussian state is equivalently described by

applying the corresponding symplectic transformation $S_{\hat{U}}$ to the covariance matrix, which are obtained by the relation $\hat{U}^{\dagger}\hat{Q}\hat{U} = S_{\hat{U}}\hat{Q}$; consequently, the covariance matrix and the first-moment vector transform to

$$\Gamma \to S_{\hat{U}}\Gamma S_{\hat{U}}^{\mathrm{T}}, \quad \vec{d} \to S_{\hat{U}}\vec{d}. \tag{3.13}$$

The transformation rule enables one to obtain the covariance matrices of a CSV state and a TMSV state after phase-encoding, which are written as

$$\Gamma_{\mathrm{CSV}} = \frac{1}{2}\begin{pmatrix} \cosh r^2 + \frac{1}{2}\cos 2\phi \sinh 2r & \sin 2\phi \sinh 2r & -e^{r}\cos\phi\sinh r & -e^{-r}\sin\phi\sinh r \\ \sin 2\phi \sinh 2r & \cosh r^2 - \frac{1}{2}\cos 2\phi \sinh 2r & -e^{r}\sin\phi\sinh r & e^{-r}\cos\phi\sinh r \\ -e^{r}\cos\phi\sinh r & -e^{r}\sin\phi\sinh r & e^{r}\cosh r & 0 \\ -e^{-r}\sin\phi\sinh r & e^{-r}\cos\phi\sinh r & 0 & e^{-r}\cosh r \end{pmatrix}, \tag{3.14}$$

$$\Gamma_{\mathrm{TMSV}} = \frac{1}{2}\begin{pmatrix} \cosh 2r + \cos 2\phi \sinh 2r & \sin 2\phi \sinh 2r & 0 & 0 \\ \sin 2\phi \sinh 2r & \cosh 2r - \cos 2\phi \sinh 2r & 0 & 0 \\ 0 & 0 & e^{-2r} & 0 \\ 0 & 0 & 0 & e^{2r} \end{pmatrix}. \tag{3.15}$$

The first moment vector of a TMSV state is zero, and that of a CSV state is

$$\vec{d}_{\mathrm{CSV}} = (\alpha\cos\phi, \alpha\sin\phi, -\alpha, 0)^{\mathrm{T}}, \tag{3.16}$$

where α is assumed to be real.

3.2.1.1 Quantum Fisher information

Once we have the covariance matrix and the first moment of the phase-encoded state, the quantum Fisher information of the state can be calculated by using a closed-form of quantum fidelity between two-Gaussian states [60]. More specifi-

cally, the quantum fidelity between two Gaussian states can be written as

$$F(\hat{\rho}_0,\hat{\rho}_1) \equiv \left(\mathrm{Tr}\sqrt{\hat{\rho}_1^{1/2}\hat{\rho}_0\hat{\rho}_1^{1/2}}\right)^2 = F_0(\Gamma_0,\Gamma_1)\exp\left[-\frac{1}{2}\vec{\delta}_u^{\mathrm{T}}(\Gamma_0+\Gamma_1)^{-1}\vec{\delta}_u\right], \tag{3.17}$$

where $F_0(\Gamma_0,\Gamma_1)$ has a closed analytical form in terms of Γ_0 and Γ_1, and $\vec{\delta}_u = \vec{d}_1 - \vec{d}_0$. Particularly for two-mode Gaussian states, we have

$$F_0 = \frac{1}{\sqrt{\Gamma}+\sqrt{\Lambda}-\sqrt{(\sqrt{\Gamma}+\sqrt{\Lambda})^2-\Delta}}, \tag{3.18}$$

where $\Delta \equiv |\Gamma_0+\Gamma_1|, \Gamma \equiv 2^4|\Omega\Gamma_0\Omega\Gamma_1 - \mathbb{1}/4|$, and $\Lambda \equiv 2^4|\Gamma_0 + i\Omega/2||\Gamma_1 + i\Omega/2|$. Using the quantum fidelity formula, one can calculate the quantum Fisher information by using the relation,

$$H(\phi) = \frac{4[1-F(\hat{\rho}(\phi),\hat{\rho}(\phi+d\phi))]}{d\phi^2}. \tag{3.19}$$

Let us first compare the quantum Fisher information of CSV states and TMSV states. If there is no photon-loss, we obtain

$$H^{\mathrm{CSV}} = \alpha^2 e^{2r} + \sinh^2 r + \alpha^2 + \frac{\sinh^2 2r}{2} \tag{3.20}$$

$$\xrightarrow{\alpha=0} \bar{n}_{\mathrm{CSV}}(2\bar{n}_{\mathrm{CSV}}+3), \tag{3.21}$$

$$H^{\mathrm{TMSV}} = \cosh 4s - 1 = 2\bar{n}_{\mathrm{TMSV}}(\bar{n}_{\mathrm{TMSV}}+2). \tag{3.22}$$

If we consider the photon-loss of transmissivity η on each mode, the quantum

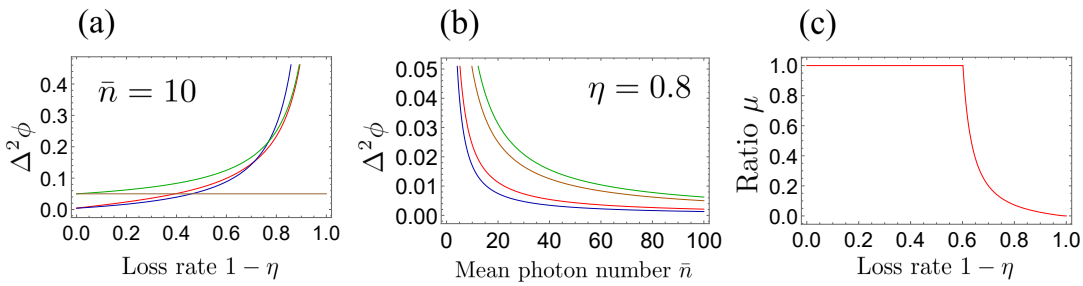


Figure 5: Quantum Fisher information of CSV state and TMSV state. (a) The red curve represents the inverse of quantum Fisher information for a CSV state, the blue curve for a TMSV state, and the green curve for a coherent state, which is a benchmark for a classical limit. The brown line represents the inverse of quantum Fisher information for coherent state without considering photon loss. the Both states show a quantum enhancement even when photon loss exists. (b) The inverse of quantum Fisher information for different mean photon number $\bar{n}$. (c) Optimized ratio of squeezed state in CSV states. When loss rate is small, injecting a squeezed state without a coherent state is optimal while the ratio decreases as loss rate increases.

Fisher information is degraded as

$$H_\eta^{\mathrm{CSV}} = \frac{\eta \sinh^2 r[1+\eta^2+(2-\eta)\eta\cosh 2r]}{1+2\eta(1-\eta)\sinh^2 r} + \frac{2\alpha^2\eta(e^r-\eta\sinh r)}{e^r-2\eta\sinh r}, \tag{3.23}$$

$$H_\eta^{\mathrm{TMSV}} = \frac{2\eta^2\sinh^2 2s}{1+2\eta(1-\eta)\sinh^2 s}. \tag{3.24}$$

The comparison of quantum Fisher information of CSV state and TMSV state is shown in Fig. 5 (a), (b). When it comes to quantum Fisher information, both CSV state and TMSV state maintain quantum enhancement even under a moderate loss rate $1-\eta < 0.4$ for a given $\bar{n} = 10$. Thus, if we can construct the optimal measurement for each state, quantum enhancement can be attained.

On the other hand, non-classical states generally tend to be more fragile to photon-loss than classical states. Thus, it is important to investigate the optimal ratio of the mean photon number of the squeezed state and coherent state. We have also analyzed the optimal ratio of the mean photon number of the squeezed state to the

total mean photon number,

$$\mu = \frac{\sinh^2 r}{|\alpha|^2 + \sinh^2 r}. \tag{3.25}$$

When loss rate is small, using only squeezed states is more beneficial than using coherent states. As expected, however, as loss rate increases we need more coherent states that are more robust to loss for high sensitivity, which is shown in Fig. 5 (c).

3.2.1.2 Parity measurement

The parity measurement is a measurement that distinguishes between even and odd photon numbers [61]. It is described by a parity operator $\hat{\Pi} \equiv (-1)^{\hat{a}^\dagger \hat{a}}$; it yields measurement outcome 0 for even photon numbers and measurement outcome 1 for odd photon numbers. Once the Wigner function of the state $W(x,p)$ is known, the expectation value of the parity operator can be easily calculated by $\langle \hat{\Pi} \rangle = \pi W(0,0) = \exp(-\vec{d}^{\mathrm{T}} \Gamma^{-1} \vec{d})/2\sqrt{|\Gamma|}$. Parity measurement has been demonstrated to be optimal for CSV state and TMSV state [41, 55, 62, 63]. The linearized error based on the parity operator, which is actually the same as the inverse of the Fisher information of the parity operator, is given by

$$\Delta^2 \phi_{\hat{\Pi}} = \frac{\langle \hat{\Pi}^2 \rangle - \langle \hat{\Pi} \rangle^2}{|\partial \langle \hat{\Pi} \rangle / \partial \phi|^2} = \frac{1 - \langle \hat{\Pi} \rangle^2}{|\partial \langle \hat{\Pi} \rangle / \partial \phi|^2}. \tag{3.26}$$

Let us consider a sensitivity based on the parity measurement performed on the first mode $\hat{\Pi}_1$ after the second beam splitter. When there is no photon-loss, the

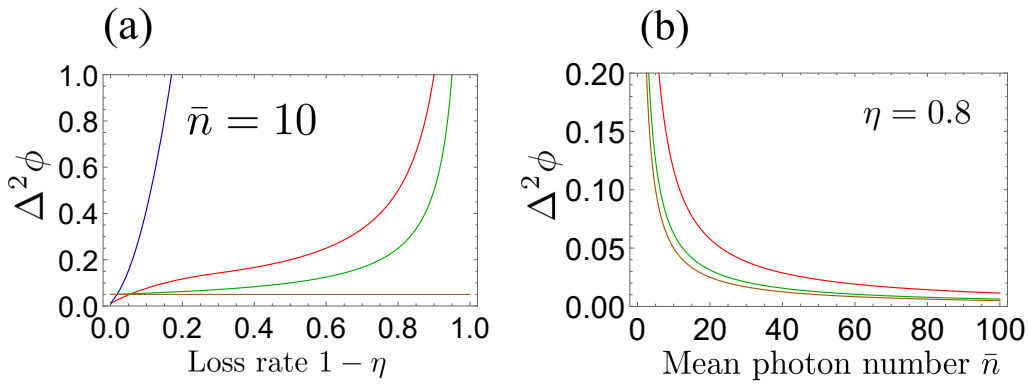


Figure 6: Classical Fisher information based on parity for CSV state and TMSV state. Each colored curve represents the quantity of the same state as Fig. 5, replacing quantum Fisher information by classical Fisher information. Parity measurement is so fragile to photon loss that both CSV state and TMSV state do not show quantum enhancement when $1-\eta > 0.05$.

sensitivities are given by

$$\Delta^2\phi^{\mathrm{CSV}}_{\hat{\Pi}_1} = \frac{1}{\alpha^2 e^{2r} + \sinh^2 r} \sim \frac{1}{\bar{n}^2_{\mathrm{CSV}}}, \tag{3.27}$$

$$\Delta^2\phi^{\mathrm{TMSV}}_{\hat{\Pi}_1} = \frac{1}{\bar{n}_{\mathrm{TMSV}}(\bar{n}_{\mathrm{TMSV}}+2)}, \tag{3.28}$$

which clearly achieve the Heisenberg scaling, although they do not attain the ultimate sensitivity of quantum Fisher information. Fig. 6 displays the sensitivity of CSV and TMSV states by parity measurement, which shows that the sensitivity via parity measurement is extremely fragile to photon-loss. Consequently, parity measurement does not enable one to see quantum enhancement in a realistic circumstance.

3.2.1.3 Homodyne measurement

Finally, we consider homodyne measurement [64]. The homodyne measurement measures the quadrature of the state $\hat{x}_\theta = (\hat{a}e^{-i\theta} + \hat{a}^\dagger e^{i\theta})/\sqrt{2}$. We calculate the clas-

sical Fisher information of homodyne detection on two modes. When an input state is a single-mode Gaussian state characterized by $(\Gamma, \vec{d})$, the outcome probability distribution of homodyne detection on the input state can be described by its covariance matrix, which is the 2×2 matrix composed of the first and third columns and rows of $S_{\hat{P}(\theta_1,\theta_2)} \Gamma S^{\mathrm{T}}_{\hat{P}(\theta_1,\theta_2)}$, and first moment vector, $([S_{\hat{P}(\theta_1,\theta_2)} \vec{d}]_1, [S_{\hat{P}(\theta_1,\theta_2)} \vec{d}]_3)^{\mathrm{T}}$, where $\hat{P}(\theta_1, \theta_2)$ is a phase shifter on the first and second mode for homodyne detection angles θ_1 and θ_2, respectively.

In the absence of photon-loss, the classical Fisher information for two-mode homodyne detection can be derived with appropriate homodyne angles, which are written as

$$F_{\mathrm{HD}}^{\mathrm{CSV}} \overset{\alpha=0}{=} \bar{n}_{\mathrm{CSV}}(2\bar{n}_{\mathrm{CSV}} + 3), \qquad F_{\mathrm{HD}}^{\mathrm{TMSV}} = 2\bar{n}_{\mathrm{TMSV}}(\bar{n}_{\mathrm{TMSV}} + 2). \tag{3.29}$$

We have numerically optimized the Fisher information of homodyne detection over

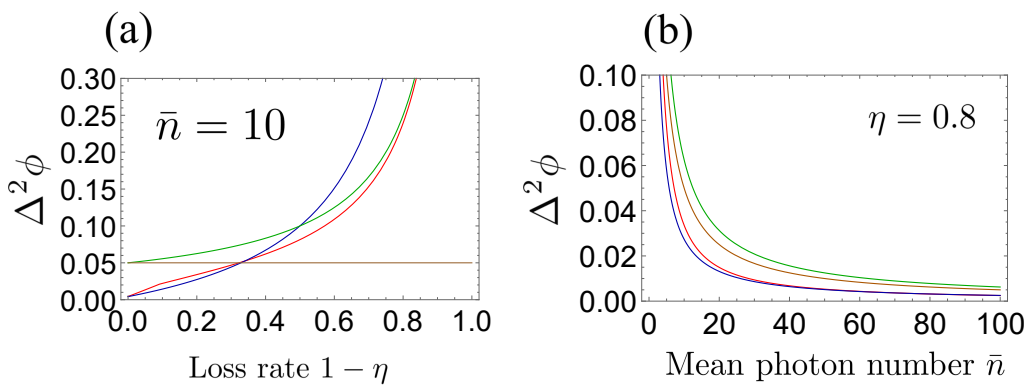


Figure 7: Classical Fisher information based on homodyne detection for CSV state and TMSV state. Each colored curve represents the quantity of the same state as Fig. 5, replacing quantum Fisher information by classical Fisher information. (a) For a fixed $\bar{n} = 10$, estimation errors of CSV state and TMSV state, implied by the inverse of classical Fisher information based on homodyne detection, is smaller than classical limit and that of coherent state under photon loss. (b) For a given fixed loss $1 - \eta = 0.2$, it shows that quantum enhancement is obtained for different $0 < \bar{n} \leq 100$.

homodyne angle, which is shown in Fig. 7. It shows that quantum enhancement is maintained even when loss rate is not very small, $1-\eta < 0.32$, which is in the contrast to parity measurement. Also, for a given loss rate $1-\eta = 0.2$, quantum enhancement is considerable for different mean photon number $0 < \bar{n} \leq 100$. Although we have focused on two-mode homodyne detection, single-mode homodyne detection also achieves a robust sensitivity under photon-loss [34].

3.2.1.4 Comparison among different measurement schemes

We have analyzed the sensitivity of each measurement scheme for a CSV state and a TMSV state, which is shown in Fig. 8. In this section, we compare the sensitivities to see which measurement is most robust. One interesting result is that for CSV state, homodyne detection is nearly optimal in a sense that the gap between the estimation errors based on homodyne detection and quantum Fisher information is small. Also, homodyne detection is robust for both states against to photon loss, so that it is expected to be employed in practice for quantum enhanced phase estimation. Finally, as we have already seen, parity measurement is so fragile that the quantum enhancement is lost even under a small photon-loss.

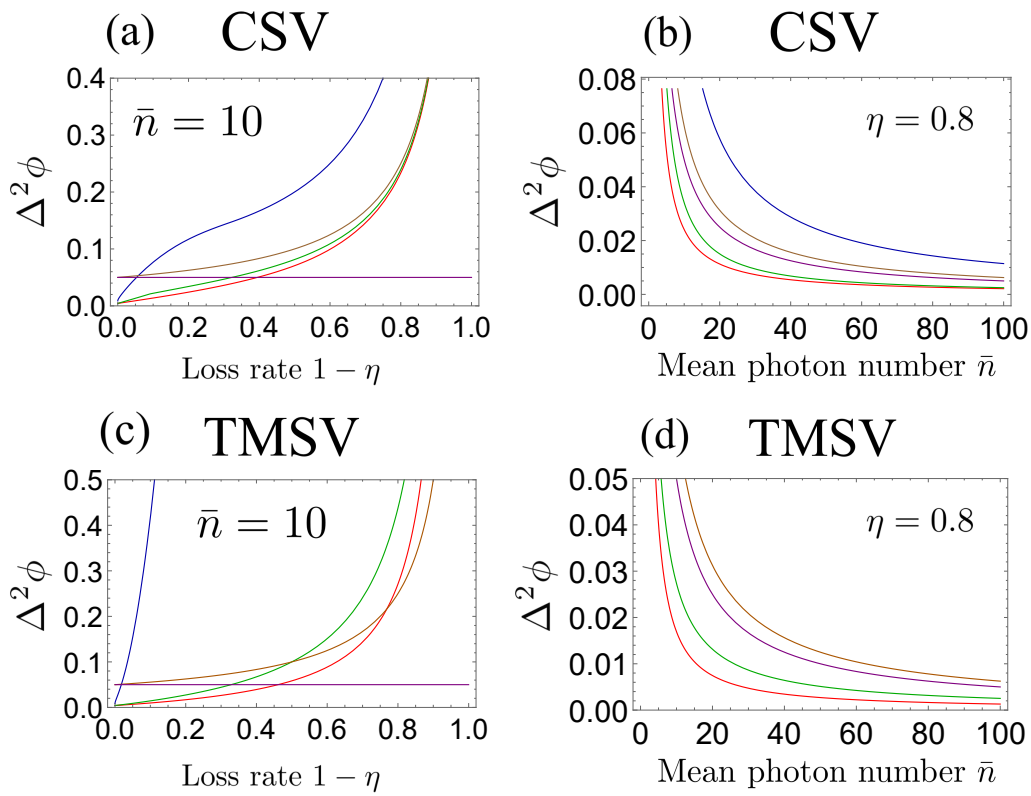


Figure 8: Comparison of sensitivities based on quantum Fisher information, parity, and homodyne measurement. Red curve represents the inverse of quantum Fisher information, blue curve the inverse of classical Fisher information based on parity, green curve the homodyne, brown the inverse of quantum Fisher information for coherent state, and purple the standard quantum limit.

3.2.2 Remarks

We have investigated the sensitivity for phase estimation in MZI using a CSV state and a TMSV state. First of all, we conclude that under a photon-loss, parity detection for a CSV state and a TMSV state is extremely fragile, while homodyne detection is much more robust against the loss. More importantly, even under a photon-loss, homodyne detection renders a nearly optimal sensitivity. In the following section, we investigate Gaussian measurement, which is a generalized version of homodyne detection, and inspect if it is optimal for phase estimation.

In this section, we have fixed the total mean photon number as $\bar{n} = 10$. Under the current technology, it is possible to generate a two-mode squeezed vacuum state with $\bar{n} = 10$. In experiment, the generation of $15dB$ single-mode squeezed vacuum states was reported [65], which corresponds to $\bar{n} \approx 7$. Injecting each single-mode squeezed vacuum state with $\bar{n} = 7$ to a 50:50 beam splitter, we can obtain the two-mode squeezed vacuum state with $\bar{n} = 14$. Although the coherent and squeezed vacuum state may not approach the range of $\bar{n} = 10$, we obtain similar phenomena to the results shown in this dissertation for the case of $\bar{n} = 7$ [34]. It is worthwhile to mention that although one may have thought of using different states other than a squeezed state when one employs a coherent state on one of two modes, the optimal choice of the state on the other port is proved to be a squeezed vacuum state [66].

Finally, it is worth emphasizing that the phase shift operator of Eq. (3.3) assumes a strong reference beam outside of the interferometer which can be used for homodyne detection or other types of coherent measurements. If one does not want to assume the reference beam, an appropriate phase randomization or multiparameter estimation has to be properly considered. More detailed discussion is provided in Ref. [67]

3.3 Gaussian measurements for single-mode phase estimation with Gaussian states

In this section, we verify that Gaussian measurements are not sufficient to achieve the ultimate sensitivity for a single-mode phase estimation using Gaussian states. We focus on a single-mode phase estimation with Gaussian states and calculate quantum Fisher information and compare it with classical Fisher information of general Gaussian measurements.

Recall that any single-mode Gaussian states can be written as

$$\hat{\rho} = \hat{D}(\alpha)\hat{S}(\xi)\hat{\rho}_{\mathrm{T}}\hat{S}^{\dagger}(\xi)\hat{D}^{\dagger}(\alpha), \tag{3.30}$$

where $\alpha = |\alpha|e^{i\theta_{\mathrm{c}}}, \xi = re^{i\theta_{\mathrm{s}}} \in \mathbb{C}$ are displacement and squeezing parameters, and $\hat{\rho}_{\mathrm{T}}$ is a thermal state. We consider a single-mode phase estimation, which is depicted in Fig. 9. The input state $\hat{\rho}_{\mathrm{in}}$ evolves to $\hat{\rho}_{\phi} = \hat{P}(\phi)\hat{\rho}_{\mathrm{in}}\hat{P}^{\dagger}(\phi)$ through a phase shifter $\hat{P}(\phi) = e^{-i\hat{a}^{\dagger}\hat{a}\phi}$. The goal is to estimate ϕ as precise as possible. We shall analyze the optimal sensitivity and then the sensitivity attainable by using general Gaussian measurement. Finally, we find the optimal measurement to achieve the optimal precision.

3.3.1 Optimal Sensitivity

We first calculate the quantum Fisher information for the unknown parameter ϕ. When the parameter of interest ϕ is generated by a Hamiltonian $\hat{H}$, the corresponding quantum Fisher information can be written in terms of the Hamiltonian and the input state, diagonalized to $\hat{\rho} = \sum_n p_n|\psi_n\rangle\langle\psi_n|$, as,

$$H(\phi) = 2\sum_{n,m}\frac{(p_n - p_m)^2}{p_n + p_m}|\langle\psi_n|\hat{H}|\psi_m\rangle|^2. \tag{3.31}$$

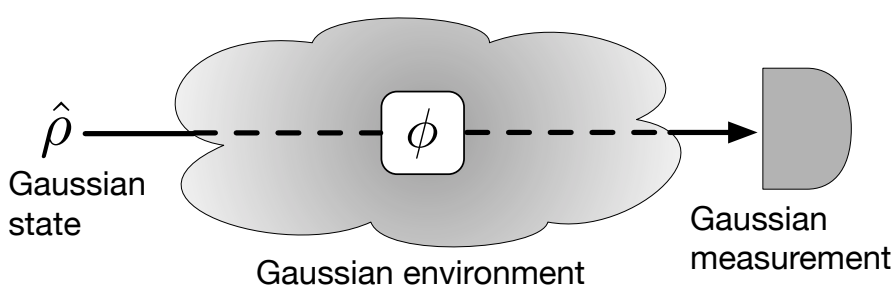


Figure 9: Single-mode phase estimation

For phase estimation, illustrated in Fig. 9, the Hamiltonian is written as $\hat{H} = \hat{a}^\dagger \hat{a}$ and ϕ represents the phase of interest. Also, from Eq. (3.30), we have a diagonalized form of Gaussian states as

$$\hat{\rho} = \sum_{n=0}^{\infty} p_n \hat{D}(\alpha)\hat{S}(\xi)|n\rangle\langle n|\hat{S}^\dagger(\xi)\hat{D}^\dagger(\alpha), \quad p_n = \frac{1}{\bar{n}+1}\left(\frac{\bar{n}}{\bar{n}+1}\right)^n, \tag{3.32}$$

where $\bar{n}$ is the average photon number of the thermal state $\hat{\rho}_{\mathrm{T}}$, and thus, the eigenstates of the density matrix are $|\psi_n\rangle = \hat{D}(\alpha)\hat{S}(\xi)|n\rangle$. After a straightforward algebra, we obtain

$$H(\phi) = \frac{2(2\bar{n}+1)^2 \sinh^2 2r}{2\bar{n}^2+2\bar{n}+1} + \frac{4|\alpha|^2}{2\bar{n}+1}|\cosh r - e^{i(\theta_s - 2\theta_c)}\sinh r|^2. \tag{3.33}$$

Before we proceed, let us consider the Gaussian environment that introduces noise on the state with preserving the Gaussian property. Let us recall that the dynamics of the state under a typical phase-insensitive Gaussian channel in thermal equilibrium can be generally described by the quantum master equation, written in the interaction picture as

$$\frac{d\hat{\rho}(t)}{dt} = \frac{\gamma}{2}\left(\bar{n}_e L[\hat{a}^\dagger] + (\bar{n}_e+1)L[\hat{a}^\dagger]\right)\hat{\rho}(t), \tag{3.34}$$

where $L[\hat{o}]\hat{\rho} = (2\hat{o}\hat{\rho}\hat{o}^\dagger - \hat{o}^\dagger\hat{o}\hat{\rho} - \hat{\rho}\hat{o}^\dagger\hat{o})$ with a damping rate of $\gamma \geq 0$, and $n_e \geq 0$ represents the average number of thermal photons of the environment. For Gaussian states, the quantum master equation can be easily solved and the state after time t can be characterized by

$$\Gamma(t) = (1-\eta)\Gamma_\infty + \eta\Gamma_{\mathrm{in}}, \quad \vec{d} = \sqrt{\eta}\vec{d}_{\mathrm{in}}, \tag{3.35}$$

where $\Gamma_\infty = (\bar{n}_e + 1/2)\mathbb{1}_2$ and $\eta = e^{-\gamma t}$. If we prepared an input Gaussian state of a squeezing parameter r_{in} and a displacement α_{in} with a thermal photon number $\bar{n}_{\text{th,in}}$, the state after time t has the parameters as

$$\alpha = \sqrt{\eta}\alpha_{\text{in}}, \tag{3.36}$$

$$r = \frac{1}{2}\ln\left[\frac{(1-\eta)(2\bar{n}_e+1)+\eta(2\bar{n}_{\text{th,in}}+1)e^{2r_{\text{in}}}}{\sqrt{A+B}}\right], \tag{3.37}$$

$$\bar{n} = \frac{1}{2}\sqrt{A+B} - \frac{1}{2}. \tag{3.38}$$

where $A = \eta(2\bar{n}_{\text{th,in}}+1) + (1-\eta)(2\bar{n}_e+1)]^2, B = 4\eta(1-\eta)(2\bar{n}_{\text{th,in}}+1)(2\bar{n}_e + 1)\sinh^2 r_{\text{in}}$. An important remark is that the loss channel maps a Gaussian state into another Gaussian state, the form of which is again written as Eq. (3.30)

3.3.2 Optimal Gaussian measurement

Gaussian measurement is defined, throughout the present dissertation, as a measurement that gives Gaussian distribution if it is performed on Gaussian states [24]. From a practical perspective, any Gaussian measurement is implementable only by adding Gaussian ancilla and using homodyne detection with Gaussian operations, as depicted in Fig. 10. More rigorously, the elements of general single-mode Gaussian measurement can always be written as

$$\hat{E}_y = \frac{1}{\pi}\hat{D}(y)\hat{\Pi}^0\hat{D}^\dagger(y), \tag{3.39}$$

where $\hat{\Pi}^0$ is a density matrix of a general Gaussian state characterizing the measurement, and $y \in \mathbb{C}$ is the measurement outcome. Thus, the resulting distribution of the measurement on a quantum state $\hat{\rho}$ is $P(y) = \text{Tr}[\hat{\rho}\hat{E}_y]$. Since the probability distribution of the measurement outcome for $\hat{\Pi}^0$ being a squeezed thermal state

can be decomposed into a mixture of those for $\hat{\Pi}^0$ being squeezed vacuum states, without loss of generality, we can assume $\hat{\Pi}^0$ to be a squeezed vacuum state of a squeezing parameter $se^{i\psi}$ by the virtue of data processing inequality [68, 69]. The implementation of general Gaussian measurements is described in Fig. 11. Particularly important examples of Gaussian measurement are homodyne detection and heterodyne detection. Homodyne detection corresponds to the case where $\hat{\Pi}^0$ is an infinitely squeezed state, and heterodyne detection corresponds to the case where $\hat{\Pi}^0$ is a vacuum state. The classical Fisher information of Gaussian measurement of characterized by γ, which is the covariance matrix of $\hat{\Pi}^0$, for a Gaussian state characterized by $\{\Gamma, \vec{d}\}$ is given by

$$F(\phi) = 4\frac{\partial \vec{d}^{\mathrm{T}}}{\partial \phi}(\Gamma+\gamma)^{-1}\frac{\partial \vec{d}}{\partial \phi} + \frac{1}{2}\mathrm{Tr}\left[(\Gamma+\gamma)^{-1}\frac{\partial \Gamma}{\partial \phi}(\Gamma+\gamma)^{-1}\frac{\partial \Gamma}{\partial \phi}\right]. \tag{3.40}$$

The derivation is supplied in Appendix.

The squeezing parameter of $se^{i\psi}$ with $s \geq 0$, characterizing $\hat{\Pi}_0$, can be controlled in the general-dyne measurement setup by adjusting a transmittance τ of the beam

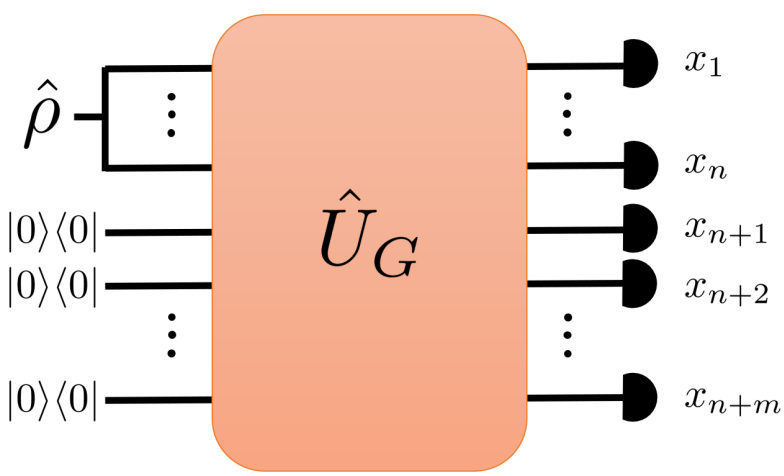


Figure 10: Gaussian measurement.

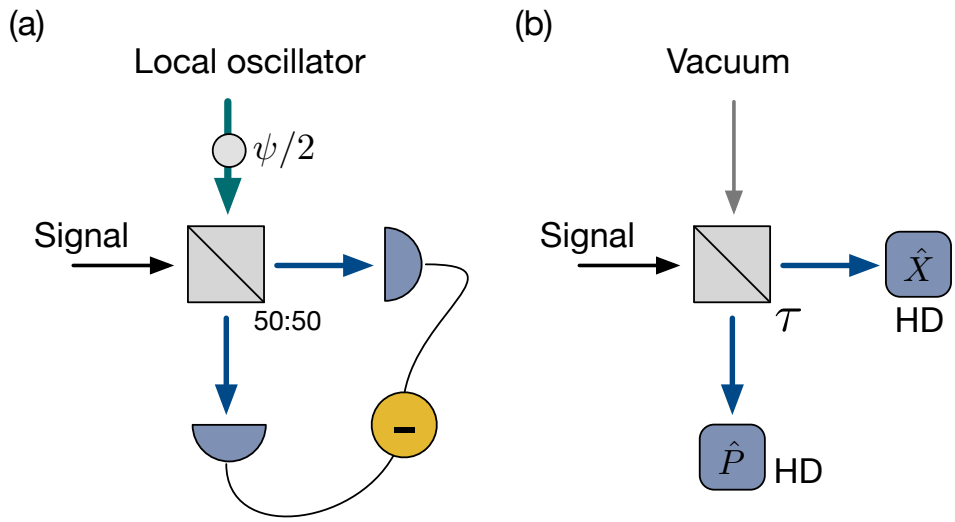


Figure 11: Implementation of (a) homodyne detection and (b) Gaussian measurement. The homodyne detection can be performed by injecting the signal and a strong laser as a local oscillator into a 50:50 beam splitter and measuring the intensity difference of the output modes. The general Gaussian measurement can be done by injecting the signal and a vacuum into a beam splitter with a transmissivity τ and performing double homodyne detection with an appropriate angle.

splitter shown in Fig. 11(b), i.e., $s = \ln\sqrt{\tau/(1-\tau)}$ with $\tau \geq 1/2$, and the phase ψ can be tuned by varying phases of the local oscillator modes in two-homodyne detections. Consequently, the outcome y is then obtained as [70],

$$y = \frac{1}{\sqrt{2\tau}} X_{\psi/2} e^{i\psi/2} + \frac{i}{\sqrt{2(1-\tau)}} P_{\psi/2} e^{i\psi/2}, \tag{3.41}$$

where $X_{\psi/2}$ and $P_{\psi/2}$ are the rotated quadrature variables, being measured in the respective output ports of the beam splitter.

Let us first consider a displaced thermal state and a squeezed thermal state. Displaced thermal states (DTS) are written as

$$\hat{\rho}_{\mathrm{DTS}} = \hat{D}(\alpha)\hat{\rho}_{\mathrm{T}}\hat{D}^{\dagger}(\alpha), \tag{3.42}$$

and the quantum Fisher information is obtained by Eq. (3.33),

$$H_{\mathrm{DTS}} = \frac{4|\alpha|^2}{2\bar{n}+1}. \tag{3.43}$$

The Fisher information for a general Gaussian measurement is obtained by using Eq. (3.40) as

$$F_{\mathrm{DTS}} = \frac{2|\alpha|^2(1+2\bar{n}+\cosh 2s_{\mathrm{DTS}} - \cos\chi_{\mathrm{DTS}}\sinh 2s_{\mathrm{DTS}})}{1+2\bar{n}(\bar{n}+1)+(2\bar{n}+1)\cosh 2s_{\mathrm{DTS}}}, \tag{3.44}$$

where $\chi_{\mathrm{DTS}} = 2(\theta_c - \phi) - \psi$. We find that when $s_{\mathrm{DTS}} \to \infty$ and $\chi_{\mathrm{DTS}} = \pi$, F_{DTS} is equal to H_{DTS}. It means that homodyne detection is optimal when we use a displaced thermal state.

Let us consider squeezed thermal states, which are written as

$$\hat{\rho}_{\mathrm{STS}} = \hat{S}(\xi)\hat{\rho}_{\mathrm{T}}\hat{S}^{\dagger}(\xi). \tag{3.45}$$

The quantum Fisher information of STS is obtained by Eq. (3.33),

$$H_{\mathrm{STS}} = C_{\mathrm{H}}\sinh^2 2r, \tag{3.46}$$

where $C_{\mathrm{H}} = 2(2\bar{n}+1)^2/(2\bar{n}^2+2\bar{n}+1)$. When we use a squeezed vacuum state, it is known that a homodyne detection ($s \to \infty$) with a homodyne angle $\chi = \tanh 2r$ is optimal. However, when the input state is a squeezed thermal state, the parameters of general Gaussian measurement have to be optimized and we get two different

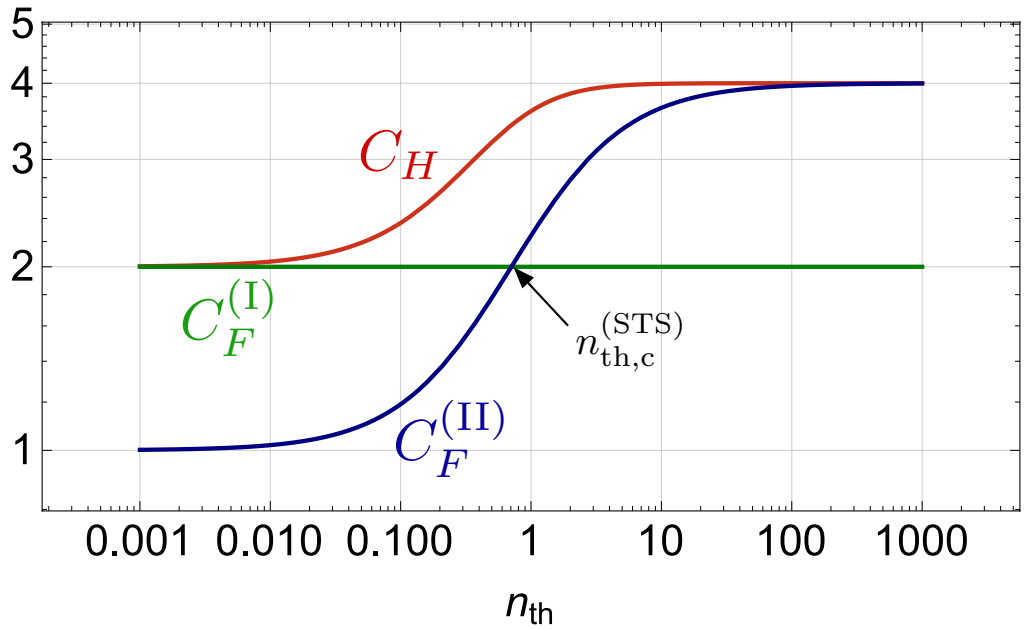


Figure 12: Quantum Fisher information and classical Fisher information with general Gaussian measurement of squeezed thermal states. Here, $n_{\mathrm{th,c}}^{(\mathrm{STS})} = 1/\sqrt{2}$.

optimal measurement for different parameters of the input state,

$$F_{\mathrm{STS}}^{(I)} = C_F^{(I)} \sinh^2 2r, \tag{3.47}$$

$$F_{\mathrm{STS}}^{(II)} = C_F^{(II)} \sinh^2 2r, \tag{3.48}$$

where $C_F^{(I)} = 2$, and $C_F^{(II)} = [(2\bar{n}+1)/(\bar{n}+1)]^2$. $F_{\mathrm{STS}}^{(I)}$ is obtained by a homodyne detection ($s \to \infty$) with a homodyne angle $\cos\chi_{\mathrm{STS}}^{(I)} = \tanh 2r$ and $F_{\mathrm{STS}}^{(II)}$ is obtained by a general Gaussian measurement with $s = r$ and $\cos\chi_{\mathrm{STS}}^{(II)} = 1$. Thus, when $\bar{n} \leq 1/\sqrt{2}$, the homodyne detection is the optimal Gaussian measurement. Otherwise, the latter is the optimal Gaussian measurement. More importantly, neither of the Gaussian measurements achieve the optimal sensitivity implied by quantum Fisher information of Eq. (3.46), as shown in Fig. 12.

3.3.3 Optimal measurement

We have shown that Gaussian measurement is not optimal for phase estimation with Gaussian states in general. In this section, we find the optimal measurement for phase estimation with general Gaussian states. The optimal measurement can be found by using symmetric logarithmic derivative (SLD) operator $\hat{L}_\phi$ satisfying the equation $\partial\hat{\rho}_\phi/\partial\phi = (\hat{\rho}_\phi\hat{L}_\phi + \hat{L}_\phi\hat{\rho}_\phi)/2$. It is well-known that a measurement composed of the projections onto the eigenbasis of the SLD operator is the optimal measurement, which is unique if the encoded state is full-rank [33, 71]. Since any non-pure Gaussian states are full-rank, the measurement derived by the SLD operator is always the unique optimal measurement. The SLD operator for a parameter-encoded state $\hat{\rho}_\phi = \sum_n p_n|\psi_n\rangle\langle\psi_n|$ with $\langle\psi_n|\psi_m\rangle = \delta_{m,n}$ is written as

$$\hat{L}_\phi = 2\sum_{n,m}\frac{\langle\psi_m|\partial_\phi\hat{\rho}_\phi|\psi_n\rangle}{p_n+p_m}|\psi_m\rangle\langle\psi_n|, \tag{3.49}$$

where the summation is taken over n,m for which $p_n + p_m \neq 0$. Using the spectral decomposition of a single-mode Gaussian state of Eq. (3.30), when $r \neq 0$, one can simplify the SLD operator,

$$\begin{aligned}\hat{L}_\phi = c\mathbb{1}_2 - &\frac{(2\bar{n}+1)\sinh 2r}{2\bar{n}^2+2\bar{n}+1}\times \\ &\hat{D}(\alpha_\phi)\hat{D}(\omega)\hat{P}(\phi-\theta_s/2)(\hat{x}\hat{p}+\hat{p}\hat{x})\hat{P}^\dagger(\phi-\theta_s/2)\hat{D}^\dagger(\omega)\hat{D}^\dagger(\alpha_\phi),\end{aligned} \tag{3.50}$$

where

$$\mathrm{Re}\omega = \frac{|\alpha|}{2\bar{n}+1}[\cos(\phi+\theta_c-\theta_s)\cosh 2r + \cos(\phi-\theta_c)\sinh 2r], \tag{3.51}$$

$$\mathrm{Im}\omega = -\frac{|\alpha|}{2\bar{n}+1}[\sin(\phi+\theta_c-\theta_s)\cosh 2r + \sin(\phi-\theta_c)\sinh 2r], \tag{3.52}$$

and the constant c can be determined from $\mathrm{Tr}[\hat{\rho}_\phi \hat{L}_\phi] = 0$. It implies that the optimal measurement is the projections onto the eigenbasis of $\hat{x}\hat{p} + \hat{p}\hat{x}$ followed by Gaussian unitary operations, including displacement operator and phase shifter.

For STS, i.e. $\alpha = 0$, the SLD operator becomes

$$\hat{L}_\phi = -\frac{(2\bar{n}+1)\sinh r}{2\bar{n}^2 + 2\bar{n} + 1} \sinh 2r (\hat{x}_{\theta_s/2-\phi}\hat{p}_{\theta_s/2-\phi} + \hat{p}_{\theta_s/2-\phi}\hat{x}_{\theta_s/2-\phi}), \tag{3.53}$$

which implies that if we use a squeezed state as an input probe, which is the optimal state in a lossless case, and photon-loss occurs, we need to construct the measurement setting for $\hat{x}\hat{p} + \hat{p}\hat{x}$ to achieve the optimal sensitivity.

On the other hand, when $r = 0$, i.e., for DTS, the SLD operator becomes

$$\hat{L}_\phi = -\frac{2\sqrt{2}|\alpha|}{2\bar{n}+1} \hat{x}_{\theta_c - \phi - \pi/2}. \tag{3.54}$$

Hence, the optimal measurement is projections onto quadrature operator, which can be implemented by homodyne detection [64].

3.3.4 Remarks

In this section, we have investigated a single-mode phase estimation with Gaussian states. One of the most intriguing results is that Gaussian measurement is not optimal for phase estimation with Gaussian states except for displaced thermal states and squeezed vacuum states. It suggests that certain non-Gaussian measurements have to be prepared to achieve the optimal sensitivity, which turns out to be the projections onto the eigenbasis of $\hat{x}\hat{p} + \hat{p}\hat{x}$ followed by Gaussian unitary operations, including displacement operator and phase shifter. Since the experimental implementation of the optimal measurement has not been found, an important future

work is to find an appropriate measurement setup for it.

In general SLD operators of single-mode Gaussian states can be always written as [26, 72, 73]

$$\hat{L}_\phi = L_\phi^{(0)} + \vec{L}_\phi^{(1)\mathrm{T}} \hat{Q} + \hat{Q}^\mathrm{T} L_\phi^{(2)} \hat{Q}, \tag{3.55}$$

where $\hat{Q} = (\hat{x}, \hat{p})^\mathrm{T}$ is the quadrature vector, $L_\phi^{(0)}$ is a real constant, $\vec{L}_\phi^{(1)}$ is a real 2-dimensional vector, and $L_\phi^{(2)}$ is a 2×2 real symmetric matrix. $L_\phi^{(2)}$ of displaced thermal states for phase estimation is a zero matrix because the SLD operator consists of linear terms of $\hat{a}$ and $\hat{a}^\dagger$ as shown in Eq. (3.54). Thus the SLD operator is proportional to a rotated quadrature operator, so that the eigenbasis of SLD operator is the quadrature eigenstate. In this case, it is evident that a homodyne detection along a rotated quadrature operator is the optimal measurement. On the other hand, the SLD operator of displaced squeezed thermal states has non-zero $L_\phi^{(2)}$. Eq. (3.50) implies that the SLD operator can be written as

$$\hat{L}_\phi = L_\phi^{(0)} + \mathcal{A} \hat{Q}'^\mathrm{T} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \hat{Q}', \tag{3.56}$$

where $\hat{Q}'$ is a transformed quadrature operator vector by symplectic matrices. The fact that Gaussian measurement is not optimal for displaced squeezed thermal states suggests that non-Gaussian measurement is required to implement optimal measurement for the SLD operator Eq. (3.56). For pure states, the eigenbasis of Eq. (3.55) is not enough to provide full information for optimal measurements. Note that the SLD operator is not unique for pure states. Thus, in that case, whether Gaussian measurement is optimal or not cannot be determined solely by Eq. (3.55).

It is worth comparing the SLDs of loss parameter in Gaussian metrology [74]. For

the case of loss parameter estimation in Gaussian metrology, the SLD operator is written as

$$\hat{L}_\phi = L_\phi^{(0)} + \mathcal{K}\hat{Q}'^{\mathrm{T}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \hat{Q}' = L_\phi^{(0)} + \mathcal{K}(\hat{x}'^2 + \hat{p}'^2) = L_\phi^{(0)} + \mathcal{K}\hat{a}^{\dagger\prime}\hat{a}' \qquad (3.57)$$

where $\mathcal{K}$ is some real constant and $\hat{O}'$ is symplectic-transformed operators. In this case, photon number counting after Gaussian operations is the optimal measurement. Inspired by the above observation, in the following section, we find optimal measurements for general Gaussian quantum metrology.

3.4 Optimal measurements for Quantum fidelity and Quantum Fisher information of Gaussian states

In this final section of the chapter, we identify the optimal measurement for general Gaussian metrology. So far, we have concluded that general Gaussian measurement is not enough for Gaussian metrology, so that certain non-Gaussian measurement is necessary to attain the optimal sensitivity that the input state allows. In this section, we find the optimal measurement for general Gaussian metrology and classify three different types of optimal measurements for single-mode Gaussian metrology depending on the circumstances: an excitation-number-resolving detection, a projection onto the eigenbasis of operator $\hat{x}\hat{p} + \hat{p}\hat{x}$, and a quadrature variable detection. In order to derive the optimal measurement, we first study quantum fidelity and the optimal measurement for it.

3.4.1 Optimal measurement for Gaussian quantum fidelity

Before we start to study quantum fidelity, let us consider two probability distributions $p_0(x)$ and $p_1(x)$ over the same domain $\mathcal{X}$. For these distributions, we may define a closeness between the distribution that measures how similar they are. One particularly interesting and useful measure is the so-called Bhattacharyya coefficient [75],

$$\mathrm{BC}(p_0,p_1) = \int_{x\in\mathcal{X}} dx\sqrt{p_0(x)p_1(x)} \quad \text{if continuous variable} \tag{3.58}$$

$$= \sum_{x\in\mathcal{X}} \sqrt{p_0(x)p_1(x)} \qquad \text{if discrete variable,} \tag{3.59}$$

which is a kind of an overlap between the distributions. It enjoys a property that $0 \leq \mathrm{BC}(p_0,p_1) \leq 1$ where $\mathrm{BC}(p_0,p_1) = 0$ if and only if the supports of the distributions are disjoint, i.e., $p_0(x)p_1(x) = 0$ for all $x \in \mathcal{X}$, and $\mathrm{BC}(p_0,p_1) = 1$ if and only if $p_0(x) = p_1(x)$ for all $x \in \mathcal{X}$. Thus, it properly measures the similarity of given two distributions. Note that the Bhattacharyya coefficient is not a distance in a strict sense because it does not obey the triangle inequality which a distance has to satisfy.

We may extend the concept of the Bhattacharyya coefficient for quantum states $\hat{\rho}_0$ and $\hat{\rho}_1$ by introducing a measurement or POVM (positive-operator valued measure) $\{\hat{E}_x\}_{x\in\mathcal{X}}$ satisfying $\hat{E}_x \geq 0$ and $\sum_{x\in\mathcal{X}} \hat{E}_x = \mathbb{1}$. Thus, one may define a measurement dependent measure as

$$\mathrm{BC_Q}(\hat{\rho}_0,\hat{\rho}_1;\{\hat{E}_x\}) = \mathrm{BC}\left(\mathrm{Tr}[\hat{\rho}_0\hat{E}_x],\mathrm{Tr}[\hat{\rho}_1\hat{E}_x]\right), \tag{3.60}$$

which represents the similarity between the probability distribution obtained by a POVM. If we optimize this quantity over all possible POVMs, we finally obtain

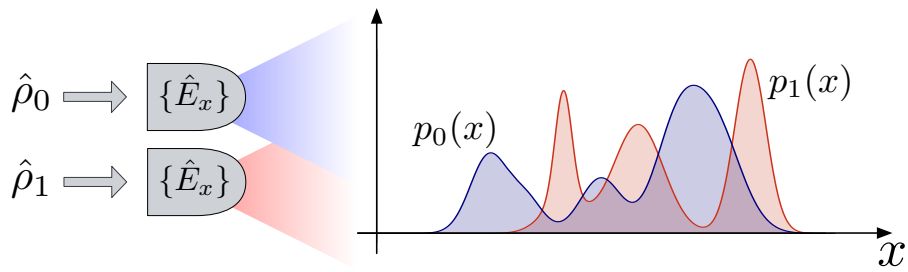


Figure 13: Quantum Fidelity

quantum fidelity [76, 77]

$$F(\hat{\rho}_0,\hat{\rho}_1)=\min_{\{\hat{E}_x\}}\mathrm{BC_Q}(\hat{\rho}_0,\hat{\rho}_1;\{\hat{E}_x\}). \tag{3.61}$$

Quantum fidelity also enjoys a similar property that $0\leq F(\hat{\rho}_0,\hat{\rho}_1)\leq 1$ where $F(\hat{\rho}_0,\hat{\rho}_1)=0$ if and only if the supports of the quantum states are disjoint, and $F(\hat{\rho}_0,\hat{\rho}_1)=1$ if and only if $\hat{\rho}_0=\hat{\rho}_1$. Thus, it also properly measures the similarity of given two quantum states. The quantum fidelity can be further simplified as [76]

$$F(\hat{\rho}_0,\hat{\rho}_1)=\left(\mathrm{Tr}\sqrt{\hat{\rho}_1^{1/2}\hat{\rho}_0\hat{\rho}_1^{1/2}}\right)^2. \tag{3.62}$$

If we focus on the operational meaning of quantum fidelity, which is the maximal distinguishability of the resultant distribution over all POVMs, it is crucial to find and prepare an optimal POVM to discriminate the quantum states the most. It has been found that the optimal POVMs satisfy [77]

$$\hat{E}_x^{1/2}(\hat{\rho}_1^{1/2}-\mu_x\hat{\rho}_0^{1/2}\hat{W}^\dagger)=0, \tag{3.63}$$

$$\mathrm{Tr}(\hat{W}\hat{\rho}_0^{1/2}\hat{E}_x\hat{\rho}_1^{1/2})\in\mathbb{R}, \tag{3.64}$$

where $\hat{W}$ is a unitary operator satisfying $\hat{W}\hat{\rho}_0^{1/2}\hat{\rho}_1^{1/2}=\sqrt{\hat{\rho}_1^{1/2}\hat{\rho}_0\hat{\rho}_1^{1/2}}$ and μ_x is a

constant. If the quantum states of interest are full-rank, the unique optimal measurement consists of projections onto the eigenbasis of a Hermitian operator

$$\hat{M}(\hat{\rho}_0,\hat{\rho}_1)=\hat{\rho}_1^{-1/2}\sqrt{\hat{\rho}_1^{1/2}\hat{\rho}_0\hat{\rho}_1^{1/2}}\hat{\rho}_1^{-1/2}. \tag{3.65}$$

Notice that for any unitary operator $\hat{U}$, the quantum fidelity and the operator $\hat{M}$ satisfy

$$F(\hat{U}\hat{\rho}_0\hat{U}^\dagger,\hat{U}\hat{\rho}_1\hat{U}^\dagger)=F(\hat{\rho}_0,\hat{\rho}_1), \tag{3.66}$$

$$\hat{M}(\hat{U}\hat{\rho}_0\hat{U}^\dagger,\hat{U}\hat{\rho}_1\hat{U}^\dagger)=\hat{U}\hat{M}(\hat{\rho}_0,\hat{\rho}_1)\hat{U}^\dagger, \tag{3.67}$$

which are straightforward but highly useful in this section.

Let us now consider quantum fidelity between arbitrary Gaussian states. Due to its importance, there have been numerous studies to find a closed form of quantum fidelity between any Gaussian states [78, 79, 80, 81, 82, 83, 84, 85, 86, 60]. Only recently, an analytical expression of quantum fidelity between arbitrary multimode Gaussian states has been found [60]. In order to achieve it, we need to find the optimal measurement satisfying Eqs. (3.63) and (3.64) or to simplify the expression of Eq. (3.65). To do those, it is crucial to introduce and employ the Gibbs exponential form of Gaussian states [60, 87],

$$\hat{\rho}_{\mathrm{G}}[G,\vec{d}]\equiv\frac{1}{Z_\Gamma}\exp\left[-\frac{1}{2}(\hat{Q}-\vec{d})^{\mathrm{T}}G(\hat{Q}-\vec{d})\right], \tag{3.68}$$

where $\vec{d}$ is the first moment vector, G is the Gibbs matrix defined as $G=2i\Omega\coth^{-1}(2\Gamma i\Omega)$ with the covariance matrix Γ, and $Z_\Gamma=\det(\Gamma+i\Omega/2)^{1/2}$ is a normalization factor, which we hereinafter omit for simplicity.

Let us consider two arbitrary Gaussian states $\hat{\rho}_0$ and $\hat{\rho}_1$, characterized by $(G_0,\vec{d}_0)$

and $(G_1, \vec{d}_1)$, respectively. Plugging the Gibbs representation of the states into Eq. (3.65) and simplifying it, we obtain

$$\hat{M} \propto \hat{D}(\vec{d}_1) \exp\left[-\frac{1}{2}\hat{Q}^{\mathrm{T}} G_{\mathrm{M}} \hat{Q} - \vec{v}_{\mathrm{M}}^{\mathrm{T}} \hat{Q}\right] \hat{D}^{\dagger}(\vec{d}_1), \tag{3.69}$$

where the matrix G_{M} is the solution of the equation

$$e^{i\Omega G_{\mathrm{M}}} e^{i\Omega G_1} e^{i\Omega G_{\mathrm{M}}} = e^{i\Omega G_0}, \tag{3.70}$$

and $\vec{v}_M$ is a real vector, which can be explicitly expressed for particular cases as below. First of all, when the Gibbs matrices are the same, $G_0 = G_1 = (S^{-1})^{\mathrm{T}}(\oplus_{j=1}^{n} g_j \mathbb{1}_2) S^{-1}$ with g_j being the symplectic spectrum, Eq. (3.70) has a trivial solution $G_{\mathrm{M}} = 0$, and Eq. (3.69) reduces to a simple form of $\hat{M} \propto e^{\vec{v}_{\mathrm{M}}^{\mathrm{T}}(\hat{Q} - \vec{d}_1)}$ where

$$\vec{v}_{\mathrm{M}} = (S^{-1})^{\mathrm{T}}(\oplus_{j=1}^{n} \tanh(g_j/2)\mathbb{1}_2)(\vec{d}_0 - \vec{d}_1). \tag{3.71}$$

The eigenbasis is the same as that of $\vec{v}_{\mathrm{M}}^{\mathrm{T}} \hat{Q}$, and therefore the homodyne detection is optimal. On the other hand, if $G_0 \neq G_1$, the operator $\hat{M}$ reduces to

$$\hat{M} \propto \hat{D}(\vec{d}_1) \hat{D}(\vec{d}_{\mathrm{M}}) \hat{\rho}_{\mathrm{G}}[G_{\mathrm{M}}, 0] \hat{D}^{\dagger}(\vec{d}_{\mathrm{M}}) \hat{D}^{\dagger}(\vec{d}_1), \tag{3.72}$$

where $\vec{v}_{\mathrm{M}} = G_{\mathrm{M}} \vec{d}_{\mathrm{M}}$ and the expression of $\vec{d}_{\mathrm{M}}$ is provided in Appendix. Note that when the first moment vectors of the states are the same, i.e., $\vec{d}_0 = \vec{d}_1$, one gets $\vec{v}_{\mathrm{M}} = 0$.

3.4.2 Optimal measurements for single-mode Gaussian states

In this section, let us focus on the single-mode case and find the optimal measurements. Any single-mode Gaussian states can be written as

$$\hat{\rho} = \hat{D}(\vec{d})\hat{S}(\xi)\hat{\rho}_{\mathrm{T}}\hat{S}^{\dagger}(\xi)\hat{D}^{\dagger}(\vec{d}), \tag{3.73}$$

where $\hat{\rho}_{\mathrm{T}} = \sum_{n=0}^{\infty} \bar{n}^n/(\bar{n}+1)^{n+1}|n\rangle\langle n|$ is a thermal state, and $\hat{S}(\xi)$ represents a single-mode squeezing operator. When $\xi = r \in \mathbb{R}$, the Gibbs matrix of the state is given by

$$G = 2\coth^{-1}(2\bar{n}+1)\begin{pmatrix} e^{2r} & 0 \\ 0 & e^{-2r} \end{pmatrix}. \tag{3.74}$$

Due to the property of $\hat{M}$ of Eq. (3.67), we can assume one of the two states, $\hat{\rho}_1$, to be a thermal state and the other, $\hat{\rho}_0$, to be a general Gaussian state without loss of generality up to Gaussian unitary operations. Thus, the Gibbs matrix of $\hat{\rho}_1$ is given by $g_1\mathbb{1}_2$. Since we already concluded that homodyne detection is optimal for $G_0 = G_1$, we focus only on the case $G_0 \neq G_1$ where $\hat{M}$ has a form of $\hat{M} \propto \hat{\rho}_{\mathrm{G}}[G_{\mathrm{M}}, 0]$ up to Gaussian unitary operations including displacement operations in Eq. (3.72). Thus, the crucial factor determining the optimal measurement is the matrix G_{M}. Since G_{M} is a symmetric matrix, we can classify it by the signs of its eigenvalues μ_1 and μ_2. The identified types are listed below.

(i) If the signs of the eigenvalues are the same, i.e., G_{M} is positive or negative definite, then the eigenbasis of $\hat{M}$ is that of the number operator $\hat{n} = (\hat{x}^2 + \hat{p}^2 - 1)/2$ followed by Gaussian unitary operations including a squeezing operation that makes the magnitude of the eigenvalues the same. Thus, the number resolving detection

along appropriate Gaussian unitary operations is optimal.

(ii) If the signs of the eigenvalues are different, then the eigenbasis of $\hat{M}$ is that of $\hat{x}\hat{p}+\hat{p}\hat{x}$ followed by similar Gaussian unitary operations to type (i). Thus, a measurement performing projection onto the eigenbasis of $\hat{x}\hat{p}+\hat{p}\hat{x}$ is optimal.

(iii) If one of the eigenvalues is zero, then the eigenbasis of $\hat{M}$ is that of a quadrature operator. Thus, homodyne detection is optimal.

The above classification is described in Fig. 14. It is worth emphasizing that type (ii) measurement is non-Gaussian, which has been shown in Sec. 3.3. Thus, except for type (iii), which is homodyne detection, non-Gaussian measurement is necessary for optimality.

Let us consider special cases that correspond to each type. The first case is that $\hat{\rho}_0$

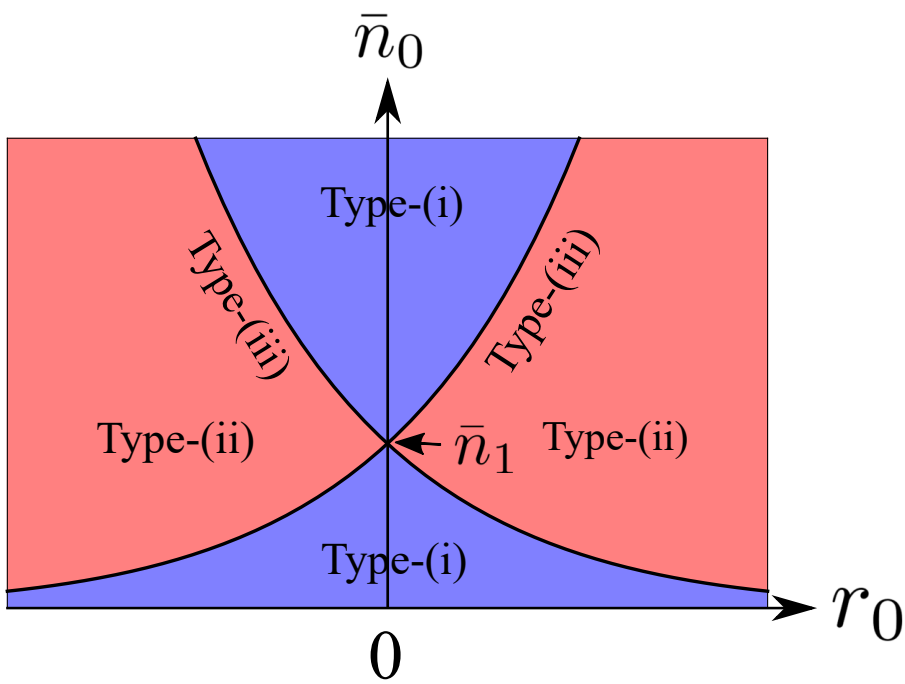


Figure 14: Classification of optimal measurements. $\bar{n}_1$ represents the average photon number of a state and $\bar{n}_0$ and r_0 represent the average photon number and squeezing parameter of the other state.

is a displaced thermal state. Thus, its Gibbs matrix is $g_0\mathbb{1}_2$. If $g_0 = g_1$, we have $G_0 = G_1$; thus type (iii) is optimal. If $g_0 \neq g_1$, we have $G_{\rm M} = (g_0 - g_1)\mathbb{1}/2$; thus the optimal measurement is type (i). Finally, when the average thermal photon numbers of the states are different each other while $G_0 \neq G_1$, one can find the signs of the eigenvalues of $G_{\rm M}$ are alway opposite, which corresponds to type (ii).

3.4.3 Optimal measurement for Gaussian quantum Fisher information

Quantum Fisher information is closely related to quantum fidelity by the following relation [33],

$$H(\theta) = \frac{4[1 - F(\hat{\rho}_\theta, \hat{\rho}_{\theta+d\theta})]}{d\theta^2}. \tag{3.75}$$

The relation implies that quantum parameter estimation is a task of distinguishing two infinitesimally close states $\hat{\rho}_\theta$ and $\hat{\rho}_{\theta+d\theta}$. Let us recall that the optimal POVM for quantum Fisher information has to satisfy [33]

$$\hat{E}_x^{1/2}(\hat{\rho}_\theta^{1/2} - \lambda_x \hat{L}_\theta \hat{\rho}_\theta^{1/2}) = 0, \tag{3.76}$$

$$\mathrm{Tr}[\hat{E}_x \hat{\rho}_\theta \hat{L}_\theta] \in \mathbb{R}, \tag{3.77}$$

which looks similar to Eqs. (3.63) and (3.64). Indeed, we prove that the above conditions of Eqs. (3.76) and (3.77) are equivalent to those of Eqs. (3.63) and (3.64) when $d\theta$ is infinitesimal, which results in the relation between $\hat{M}$ and the SLD operator $\hat{L}_\theta$ such that,

$$\hat{M}(\hat{\rho}_\theta, \hat{\rho}_{\theta+d\theta}) \simeq 1 + \hat{L}_\theta d\theta/2. \tag{3.78}$$

The proof is supplied in Appendix. Using the form of the operator $\hat{M}$ for arbitrary Gaussian states, one can find the expression for two infinitesimally close Gaussian states $\hat{\rho}_\theta$ and $\hat{\rho}_{\theta+d\theta}$,

$$\hat{M}(\hat{\rho}_\theta, \hat{\rho}_{\theta+d\theta}) \simeq 1 - \hat{D}(\vec{d}_\theta)(\hat{Q}^{\mathrm{T}} G_{\mathrm{M}} \hat{Q}/2 - \vec{v}_{\mathrm{M}}^{\mathrm{T}} \hat{Q})\hat{D}^\dagger(\vec{d}_\theta), \tag{3.79}$$

where $\vec{d}_\theta$ is the first-moment vector of $\hat{\rho}_\theta$, and G_{M} and $\vec{v}_{\mathrm{M}}$ are associated of the states as in Eq.(3.69). Consequently, one can derive the expression for the SLD operator,

$$\hat{L}_\theta d\theta = -\hat{D}(\vec{d}_\theta)(\hat{Q}^{\mathrm{T}} G_{\mathrm{M}} \hat{Q} - 2\vec{v}_{\mathrm{M}}^{\mathrm{T}} \hat{Q})\hat{D}^\dagger(\vec{d}_\theta) + \nu, \tag{3.80}$$

where $\nu = \mathrm{Tr}[\hat{D}^\dagger(\vec{d}_\theta)\hat{\rho}_\theta \hat{D}(\vec{d}_\theta)\hat{Q}^{\mathrm{T}} G_{\mathrm{M}} \hat{Q}]$ can be determined from $\mathrm{Tr}[\hat{\rho}_\theta \hat{L}_\theta] = 0$. In the infinitesimal limit of $d\theta$, one can show that Eq. (3.70) for G_{M} is simplified as

$$4\Gamma_\theta G_{\mathrm{M}} \Gamma_\theta + \Omega G_{\mathrm{M}} \Omega + 2d\theta \frac{\partial \Gamma_\theta}{\partial \theta} = 0, \tag{3.81}$$

and the solution of which can be formally written as

$$G_{\mathrm{M}} = i\Omega \sum_{m=0}^{\infty} W_\theta^{-m-1} \frac{\partial W_\theta}{\partial \theta} W_\theta^{-m-1} d\theta, \tag{3.82}$$

and $\vec{v}_{\mathrm{M}} = \Gamma_\theta^{-1}(\partial \vec{d}_\theta/\partial\theta)d\theta/2$, where Γ_θ is the covariance matrix of the state $\hat{\rho}_\theta$. Especially for isothermal Gaussian states of the thermal photon number $\bar{n}$, the G_{M} is simplified to

$$G_{\mathrm{M}} = -\frac{1}{2\bar{n}^2 + 2\bar{n} + 1} \Omega \frac{\partial \Gamma_\theta}{\partial \theta} \Omega. \tag{3.83}$$

Using the expression of the SLD operator for Gaussian states, we can now derive the quantum Fisher information for general Gaussian states,

$$H(\theta) = \mathrm{Tr}[\hat{\rho}_\theta \hat{L}_\theta^2] = -\mathrm{Tr}\left[\frac{\partial \Gamma_\theta}{\partial \theta} G_{\mathrm{M}}\right] + \frac{\partial \vec{d}_\theta^{\mathrm{T}}}{\partial \theta} \Gamma_\theta^{-1} \frac{\partial \vec{d}_\theta}{\partial \theta}. \tag{3.84}$$

Furthermore, a general expression of quantum Fisher information matrix of Gaussian states for a multiparameter case can be derived,

$$H_{ij}(\vec{\theta}) = \mathrm{Tr}[\hat{\rho}_{\vec{\theta}} \{\hat{L}_{\theta_i}, \hat{L}_{\theta_j}\}/2] = -\mathrm{Tr}\left[\frac{\partial \Gamma_{\vec{\theta}}}{\partial \theta_i} G_{\mathrm{M}}^j\right] + \frac{\partial \vec{d}_{\vec{\theta}}^{\mathrm{T}}}{\partial \theta_i} \Gamma_{\vec{\theta}}^{-1} \frac{\partial \vec{d}_{\vec{\theta}}}{\partial \theta_j}. \tag{3.85}$$

Let us consider displacement, phase, squeezing and loss estimation and find the optimal measurement. For a single-mode Gaussian probe state $\hat{\rho}_{\mathrm{in}}$ of $(\Gamma_{\mathrm{in}}, \vec{d}_{\mathrm{in}})$, the displacement operation $\hat{D}(\vec{\zeta})$ changes the first-moment vector only.

$$\Gamma_{\mathrm{in}} \to \Gamma_{\mathrm{in}}, \quad \vec{d}_{\mathrm{in}} \to \vec{d}_{\mathrm{in}} + \vec{\zeta}, \tag{3.86}$$

where $\vec{\zeta} \in \mathbb{R}^2$. Therefore, the first-moment vectors and covariance matrices of $\hat{\rho}_{\vec{\zeta}}$ and $\hat{\rho}_{\vec{\zeta}+d\vec{\zeta}}$ for $d\vec{\zeta}^{\mathrm{T}} = (d\zeta_1, 0)$ are related as

$$\Gamma_{\vec{\zeta}+d\vec{\zeta}} = \Gamma_{\vec{\zeta}}, \quad \vec{d}_{\vec{\zeta}+d\vec{\zeta}} = \vec{d}_{\zeta} + (d\zeta_1, 0)^{\mathrm{T}}, \tag{3.87}$$

respectively. Since the displacement transforms only the first-moment vector, we can immediately find that the optimal measurement for quantum fidelity between $\hat{\rho}_{\vec{\zeta}}$ and $\hat{\rho}_{\vec{\zeta}+d\vec{\zeta}}$ is type (iii), namely, homodyne detection. Using the expression of $\vec{v}_{\mathrm{M}}$, one can easily find the SLD operator and quantum Fisher information for displace-

ment parameter $\vec{\zeta} = (\zeta_1, 0)$,

$$\begin{aligned}\hat{L}_{\zeta_1} &= \hat{D}(\vec{\zeta})([\Gamma_{\zeta_1}^{-1}]_{11}\hat{x} + [\Gamma_{\zeta_1}^{-1}]_{12}\hat{p})\hat{D}^\dagger(\vec{\zeta}) \\ &= [\Gamma_{\zeta_1}^{-1}]_{11}(\hat{x} - \zeta_1) + [\Gamma_{\zeta_1}^{-1}]_{12}\hat{p}, \qquad (3.88)\\ H(\zeta_1) &= [\Gamma_{\zeta_1}^{-1}]_{11}. \qquad (3.89)\end{aligned}$$

Let us consider a phase estimation of a phase shifter $\hat{P}(\theta) = e^{-i\theta\hat{Q}^{\mathrm{T}}\hat{Q}}$ with a single-mode Gaussian input probe $\hat{\rho}_{\mathrm{in}}$. Since the first-moment vector of the input probe does not change the type of optimal measurement, we will focus on a squeezed thermal state input,

$$\hat{\rho}_{\mathrm{in}} = \hat{S}(\xi)\hat{\rho}_{\mathrm{T}}\hat{S}^\dagger(\xi) \to \hat{\rho}_\theta = \hat{P}(\theta)\hat{S}(\xi)\hat{\rho}_{\mathrm{T}}\hat{S}^\dagger(\xi)\hat{P}^\dagger(\theta). \qquad (3.90)$$

We assume $\xi = r \in \mathbb{R}$ hereinafter. The relevant states under investigation are $\hat{\rho}_\theta$ and $\hat{\rho}_{\theta+d\theta}$, but since the full expression for arbitrary θ is complicated and the value of θ does not change the optimal type of measurement, we set $\theta = 0$ for simplicity; thus, the relevant states are $\hat{\rho}_0$ and $\hat{\rho}_{d\theta}$.

Let us find the optimal measurement for $\hat{\rho}_0$ and $\hat{\rho}_\theta$. The covariance matrix of $\hat{\rho}_\theta$ is given by

$$\Gamma_\theta = \frac{2\bar{n}+1}{2}\begin{pmatrix} \cosh 2r + \cos 2\theta \sinh 2r & \sinh 2r \sin 2\theta \\ \sinh 2r \sin 2\theta & \cosh 2r - \cos 2\theta \sinh 2r \end{pmatrix}, \qquad (3.91)$$

and that of $\hat{\rho}_0$ is obtained by substituting $\theta = 0$. Since the average thermal photon numbers of the state are equal but the covariance matrices are not as long as $\theta \neq 0$ and $r \neq 0$, we can conclude that the optimal measurement is type (ii). One can

actually find the expression for G_{M} which is given by

$$G_{\mathrm{M}} = A \begin{pmatrix} -\sin\theta & \cos\theta \\ \cos\theta & \sin\theta \end{pmatrix}, \tag{3.92}$$

where a constant A is given by $\cos A = (4\bar{n}^2+4\bar{n}+2)/[(4\bar{n}^2+2\bar{n}+1)(4\bar{n}^2+6\bar{n}+3)+(2\bar{n}+1)^2\cos 2\theta+2(2\bar{n}+1)^2\cosh 4r\sin^2\theta]^{1/2}$. Indeed, the signs of the eigenvalues of G_{M} are different; thus, it confirms that type-(ii) is optimal. Explicitly, after taking a limit $\theta \to d\theta$, G_{M} becomes

$$G_{\mathrm{M}} = \frac{(2\bar{n}+1)\sinh 2r}{2\bar{n}^2+2\bar{n}+1} d\theta \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{3.93}$$

Therefore,

$$\hat{M} = 1 - \frac{(2\bar{n}+1)\sinh 2r}{2(2\bar{n}^2+2\bar{n}+1)} d\theta(\hat{x}\hat{p}+\hat{p}\hat{x}) = 1+\hat{L}_\theta d\theta/2, \tag{3.94}$$

where $\hat{L}_\theta$ is the SLD operator that is the same as one that we have derived in Sec. 3.3. Even though we have derived the matrix G_{M} directly, one may derive the same result by using Eq. (3.82).

In squeezing parameter estimation, an input Gaussian state transforms as

$$\hat{\rho}_{\mathrm{in}} = \hat{D}(\vec{d})\hat{S}(\xi)\hat{\rho}_{\mathrm{T}}\hat{S}^\dagger(\xi)\hat{D}^\dagger(\vec{d}) \tag{3.95}$$

$$\to \hat{\rho}_s = \hat{S}(s)\hat{D}(\vec{d})\hat{S}(\xi)\hat{\rho}_{\mathrm{T}}\hat{S}^\dagger(\xi)\hat{D}^\dagger(\vec{d})\hat{S}^\dagger(s), \tag{3.96}$$

where we assume $\xi = s$ is a real number. It means that we estimate the squeezing magnitude while the squeezing angle is already known. Since $\hat{\rho}_s$ and $\hat{\rho}_{s+ds}$ have a different squeezing magnitude and the equal average thermal photon number, one

can immediately find that the optimal measurement is type-(ii). Explicitly, we can find the expression of the SLD operator and the quantum Fisher information by using Eq. (3.82), which are written as

$$\begin{aligned}\hat{L}_s = &\frac{2\bar{n}+1}{2\bar{n}^2+2\bar{n}+1}\hat{D}(\vec{d})\hat{Q}^{\mathrm{T}}\mathrm{diag}[e^{-2s}(\cosh 2r-\cos\theta_s\sinh 2r)\\ &-e^{2s}(\cosh 2r+\cos\theta_s\sinh 2r)]\hat{Q}\hat{D}^{\dagger}(\vec{d})+\nu, \qquad (3.97)\\ H(s) = &\frac{2(2\bar{n}+1)^2}{2\bar{n}^2+2\bar{n}+1}(\cosh^2 2r-\cos^2\theta_s\sinh^2 2r)\\ &+\frac{4|\alpha|^2}{2\bar{n}+1}[\cosh 2r-\sinh 2r\cos(2\theta_c+\theta_s)], \qquad (3.98)\end{aligned}$$

where we have defined $\vec{d}=\sqrt{2}|\alpha|(\cos\theta_c,\sin\theta_c)^{\mathrm{T}}$. The SLD operator shows that the optimal measurement is indeed type-(ii).

Finally, let us consider estimation of a loss parameter in a phase-insensitive loss channel, the dynamics of which is described by the quantum master equations as

$$\frac{d\hat{\rho}}{dt}=\frac{\gamma}{2}(2\hat{a}\hat{\rho}\hat{a}^{\dagger}-\hat{a}^{\dagger}\hat{a}\hat{\rho}-\hat{\rho}\hat{a}^{\dagger}\hat{a}), \qquad (3.99)$$

where $\hat{a}=(\hat{x}+i\hat{p})/\sqrt{2}$ is the annihilation operator and γ is the unknown loss parameter of interest. The solution of the master equation can be given in terms of the first-moment vector and the covariance matrix as [22]

$$\vec{d}_{t=0}\to\vec{d}_t=e^{-\gamma t/2}\vec{d}_0, \qquad (3.100)$$

$$\Gamma_{t=0}\to\Gamma_t=e^{-\gamma t}\Gamma_0+(1-e^{-\gamma t})\mathbb{1}_2/2. \qquad (3.101)$$

After some algebra, one can easily obtain the expression for G_{M} and the quantum

Fisher information,

$$G_{\mathrm{M}} = A \times \mathrm{diag}(\sin^4\phi - e^{2r}\cos^4\phi, \sin^4\phi - e^{2r}\cos^4\phi) t d\gamma, \tag{3.102}$$

$$H(\gamma) = \frac{\cos^2\phi(1-2\sin^2\phi\cos^2\phi)\sinh^2 r}{\sin^2\phi(1+2\sin^2\phi\cos^2\phi\sinh^2 r)} t^2, \tag{3.103}$$

where $A \equiv 4/(-2\sinh^2 r\cos 4\phi + \cosh 2r + 7)\sin^2\phi$ and $\cos^2\phi \equiv e^{-\gamma t}$, and we have assumed a zero-displacement of the input state for simplicity. The signs of the eigenvalues of G_{M} is always negative regardless of the parameters, which suggests that the optimal measurement is type-(ii). The same result has already been derived in Ref. [74, 88]. It is worth emphasizing that the optimality of type-(ii) holds for any loss parameters of phase-insensitive channels.

3.4.4 Remarks

In this final section of the present chapter, we have identified the optimal measurements for quantum fidelity between two Gaussian states. Especially for a single-mode case, we found that there are three different types of optimal measurement depending on a type of a parameter of interest : (i) Number resolving detection, (ii) $\hat{x}\hat{p} + \hat{p}\hat{x}$ type measurement, (iii) homodyne detection. One may generalize the result of a single-mode case by investigating the normal form of symplectic matrices. Our finding is applied for parameter estimation of a Gaussian operation preserving Gaussian properties of states. The same classification holds due to the relation between quantum Fisher information and quantum fidelity.

3.5 Conclusion

In this chapter, we have studied about Gaussian quantum metrology, beginning with analysis of advanced Mach-Zehnder interferometer with coherent & squeezed state and two-mode squeezed state. Concluding that homodyne detection is robust and promising measurement for quantum enhanced phase estimation, we studied about Gaussian measurement, which is a generalization of homodyne detection, for single-mode phase estimation and found that Gaussian measurement is not sufficient to achieve the optimal sensitivity of phase estimation. Finally, we have identified three different types of optimal measurement in single-mode Gaussian metrology and provided a general method to find the optimal measurement in general Gaussian metrology.

A remaining interesting question is to find an experimental scheme to realize $\hat{x}\hat{p} + \hat{p}\hat{x}$ type of measurement which is the optimal measurement for phase estimation when squeezed thermal state is employed. Since there are only three different types of optimal measurement, which are $\hat{x}\hat{p} + \hat{p}\hat{x}$ type of measurement, homodyne measurement, and number resolving detection, and homodyne detection and number resolving detection are accessible in a certain regime, study about the realization of $\hat{x}\hat{p} + \hat{p}\hat{x}$ type of measurement will complete the optimal set of measurement for single-mode Gaussian metrology.

Gaussian states are important resources that are accessible in experiments. Our theoretical analysis is expected to be demonstrated in experiment in the near future and to play a useful role in Gaussian quantum metrology to exploit quantum enhancement.

3.6 Appendix

Classical Fisher information of Gaussian measurement on Gaussian state

Let us recall that any operator applied on a single-mode can be written by Weyl operator $\hat{V}(\vec{\xi}) = \exp(i\xi_1\hat{x} + i\xi_2\hat{p})$ such as,

$$\hat{O} = \frac{1}{2\pi}\int d^2\xi \mathrm{Tr}[\hat{O}\hat{V}^\dagger(\xi)]\hat{V}(\xi). \tag{3.104}$$

Then, the expectation value of the operator can be easily written as

$$\langle\hat{O}\rangle = \mathrm{Tr}[\hat{\rho}\hat{O}] = \frac{1}{2\pi}\int d^2\xi \mathrm{Tr}[\hat{O}\hat{V}^\dagger(\vec{\xi})]\mathrm{Tr}[\hat{\rho}\hat{V}(\vec{\xi})] \tag{3.105}$$

$$= \frac{1}{2\pi}\int d^2\xi \chi[\hat{O}](-\xi)\chi[\hat{\rho}](\xi). \tag{3.106}$$

where $\chi[\hat{O}](\vec{\xi}) = \mathrm{Tr}[\hat{\rho}\hat{V}(\vec{\xi})]$ is the characteristic function of an operator $\hat{O}$.

The Gaussian POVM is written as

$$\hat{\Pi}_\beta = \frac{1}{\pi}|\beta,\zeta\rangle\langle\beta,\zeta|. \tag{3.107}$$

and the probability of the corresponding POVM element is

$$p(\beta) = \mathrm{Tr}[\hat{\rho}\hat{\Pi}_\beta] = \frac{1}{2\pi^2}\int d^2\xi \chi[|\beta,\zeta\rangle\langle\beta,\zeta|](-\xi)\chi[\rho](\xi). \tag{3.108}$$

We know that the characteristic function of the state $|\beta,\zeta\rangle\langle\beta,\zeta|$ is

$$\chi[|\beta,\zeta\rangle\langle\beta,\zeta|] = \exp[-\frac{1}{2}\vec{\xi}^{\mathrm{T}}\gamma\xi + i\vec{\mu}^{\mathrm{T}}\vec{\xi}], \tag{3.109}$$

where

$$\gamma = \frac{1}{2}\begin{pmatrix} \cosh 2s + \cos\psi \sinh 2s & \sinh 2s \sinh\psi \\ \sinh 2s \sinh\psi & \cosh 2s - \cos\psi \sinh 2s \end{pmatrix}, \tag{3.110}$$

$$\vec{\mu} = \begin{pmatrix} \sqrt{2}|\beta|\cos\theta_c & \sqrt{2}|\beta|\sin\theta_c \end{pmatrix}^{\mathrm{T}} \equiv \begin{pmatrix} \sqrt{2}x & \sqrt{2}y \end{pmatrix}^{\mathrm{T}} \tag{3.111}$$

$$\zeta = se^{i\psi}, \text{and } \beta = |\beta|e^{i\theta_c}. \tag{3.112}$$

Then the probability density can be simplified as

$$\begin{aligned} p(\beta) &= \frac{1}{2\pi^2}\int d^2\xi \exp[-\frac{1}{2}\xi^{\mathrm{T}}\gamma\xi - i\vec{\mu}^{\mathrm{T}}\xi]\exp[-\frac{1}{2}\xi^T\Gamma\xi + i\vec{d}^T\xi] \\ &= \frac{1}{2\pi^2}\int d^2\xi \exp[-\frac{1}{2}\xi^{\mathrm{T}}(\gamma+\Gamma)\xi + i(\vec{d}-\vec{\mu})^{\mathrm{T}}\xi] \\ &= \frac{1}{2\pi^2}\sqrt{\frac{(2\pi)^2}{|\gamma+\Gamma|}}\exp[-\frac{1}{2}(\vec{d}-\vec{\mu})^{\mathrm{T}}(\gamma+\Gamma)^{-1}(\vec{d}-\vec{\mu})] \\ &= \frac{1}{\pi}\sqrt{\frac{1}{|\gamma+\Gamma|}}\exp[-\frac{1}{2}(\vec{d}-\vec{\mu})^{\mathrm{T}}(\gamma+\Gamma)^{-1}(\vec{d}-\vec{\mu})] \\ &= \frac{1}{\pi}\sqrt{\frac{1}{|\gamma+\Gamma|}}\exp\left[-\frac{1}{2}\left(x-\frac{d_1}{\sqrt{2}}, y-\frac{d_2}{\sqrt{2}}\right)\left(\frac{\gamma+\Gamma}{2}\right)^{-1}\right. \\ &\qquad\qquad \left.\left(x-\frac{d_1}{\sqrt{2}}, y-\frac{d_2}{\sqrt{2}}\right)^{\mathrm{T}}\right] \\ &= \frac{1}{\pi}\sqrt{\frac{1}{|\gamma+\Gamma|}}\exp\left[-\frac{1}{2}(\vec{X}-\frac{\vec{d}}{\sqrt{2}})^{\mathrm{T}}\Sigma^{-1}(\vec{X}-\frac{\vec{d}}{\sqrt{2}})\right]. \end{aligned} \tag{3.113}$$

where

$$\Sigma \equiv \frac{\Gamma+\gamma}{2}. \tag{3.114}$$

This implies that Gaussian measurement with a parameter $se^{i\psi}$ on Gaussian states with $\{\Gamma, \vec{d}\}$ results in the Gaussian probability density with covariance matrix and

the first moment vector $\{\Sigma, \vec{d}/\sqrt{2}\}$. Then the Fisher information based on Gaussian measurement is

$$F(\phi) = 2\frac{\partial \vec{d}^{\mathrm{T}}}{\partial \phi}\Sigma^{-1}\frac{\partial \vec{d}}{\partial \phi} + \frac{1}{2}\mathrm{Tr}\left[\Sigma^{-1}\frac{\partial \Sigma}{\partial \phi}\Sigma^{-1}\frac{\partial \Sigma}{\partial \phi}\right] \tag{3.115}$$

$$= 4\frac{\partial \vec{d}^{\mathrm{T}}}{\partial \phi}(\Gamma+\gamma)^{-1}\frac{\partial \vec{d}}{\partial \phi} + \frac{1}{2}\mathrm{Tr}\left[(\Gamma+\gamma)^{-1}\frac{\partial \Gamma}{\partial \phi}(\Gamma+\gamma)^{-1}\frac{\partial \Gamma}{\partial \phi}\right]. \tag{3.116}$$

SLD for displaced squeezed thermal states

Here we derive the SLD for displaced squeezed thermal states. The SLD operator is written as

$$\hat{L}_\phi = 2\sum_{n,m}\frac{p_n - p_m}{p_n + p_m}\langle\psi_m|\partial_\phi\psi_n\rangle|\psi_m\rangle\langle\psi_n|, \tag{3.117}$$

where the summation is taken over n, m for which $p_n + p_m \neq 0$. Using Eq. (3.117) and $|\partial_\phi\psi_n\rangle = -i\hat{a}^\dagger\hat{a}|\psi_n\rangle$ for phase rotation, we can obtain

$$\begin{aligned}
\langle\psi_m|\partial_\phi\psi_n\rangle &= -i\langle m|S^\dagger(\xi)D^\dagger(\alpha)\hat{a}^\dagger\hat{a}\hat{D}(\alpha)\hat{S}(\xi)|n\rangle \\
&= -i\langle m|(\hat{a}^\dagger\cosh r - \hat{a}e^{-i\theta_s}\sinh r + \alpha^*)(\hat{a}\cosh r - \hat{a}^\dagger e^{i\theta_s}\sinh r + \alpha)|n\rangle \\
&= -i\big(-\delta_{m-2,n}\sqrt{m}\sqrt{m-1}e^{i\theta_s}\sinh r\cosh r \\
&+ \delta_{m-1,n}\sqrt{m}\alpha\cosh r - \delta_{m,n-2}\sqrt{n}\sqrt{n-1}e^{-i\theta_s}\sinh r\cosh r \\
&- \delta_{m,n-1}\sqrt{n}\alpha e^{-i\theta_s}\sinh r + \delta_{m,n-1}\sqrt{n}\alpha^*\cosh r \\
&= -\delta_{m-1,n}\sqrt{m}\alpha^* e^{i\theta_s}\sinh r + \delta_{m,n}\cdot(\textit{irrelevant term})\big),
\end{aligned} \tag{3.118}$$

where the last term that is proportional to $\delta_{m,n}$ which is irrelevant in the summation. Substituting $p_n = \bar{n}^n/(1+\bar{n})^{n+1}$ and Eq. (3.118) into Eq. (3.117) and simplifying,

one get

$$\hat{L}_\phi = \hat{R}(\phi)\hat{D}(\alpha)\hat{S}(\xi)(\hat{L}_1 + \hat{L}_2)\hat{S}^\dagger(\xi)\hat{D}^\dagger(\alpha)\hat{R}^\dagger(\phi),$$

where

$$\begin{aligned}
\hat{L}_1 &= \frac{2}{2\bar{n}+1} i \left(\alpha^* \cosh r - \alpha e^{-i\theta_s} \sinh r\right) \hat{a} + \text{h.c.}, \\
\hat{L}_2 &= \frac{(2\bar{n}+1)\sinh 2r}{2\bar{n}^2 + 2\bar{n} + 1} i \hat{a}^{\dagger 2} e^{i\theta_s} + \text{h.c.}.
\end{aligned}$$

For a given operator $\hat{O}$ defined as

$$\begin{aligned}
\hat{O} &= \mathcal{A}\hat{S}(\xi)\hat{D}(\beta)\hat{R}(-\theta_s/2)(\hat{x}\hat{p} + \hat{p}\hat{x})\hat{R}^\dagger(-\theta_s/2)\hat{D}^\dagger(\beta)\hat{S}^\dagger(\xi) \\
&= i\mathcal{A}\left[\hat{a}^{\dagger 2} e^{i\theta_s} - \hat{a}^2 e^{-i\theta_s}\right] + 2i\mathcal{A}\left[\left(\beta \cosh r e^{-i\theta_s} - \beta^* \sinh r\right)\hat{a}\right. \\
&\left. + \left(\beta \sinh r - \beta^* e^{i\theta_s} \cosh r\right)\hat{a}^\dagger\right] + i\mathcal{A}(\beta^{*2} e^{i\theta_s} - \beta^2 e^{-i\theta_s}),
\end{aligned}$$

Assuming that $r \neq 0$, it can easily be shown that when

$$\mathcal{A} = \frac{(2\bar{n}+1)\sinh 2r}{2\bar{n}^2 + 2\bar{n} + 1}, \tag{3.119}$$

$$\theta_b = \theta_s - \theta_c, \tag{3.120}$$

the operator $\hat{L}_1 + \hat{L}_2$ is written as

$$\hat{L}_1 + \hat{L}_2 = \hat{O} - C\mathbb{1}, \tag{3.121}$$

where $C = 2\mathcal{A}|\beta|^2 \sin(\theta_s - 2\theta_c)$. Substituting Eq. (3.121) to Eq. (3.119), one may show that Eq. (3.50) is derived after a little algebra.

SLD for displaced thermal states

The SLD for displaced thermal states can be written as

$$\hat{L}_\phi = \frac{2i}{2\bar{n}+1}\hat{R}(\phi)\hat{D}(\alpha)\left(\alpha^*\hat{a}-\alpha\hat{a}^\dagger\right)\hat{D}^\dagger(\alpha)\hat{R}^\dagger(\phi).$$

This can be further simplified to be

$$\hat{L}_\phi = \frac{2i}{2\bar{n}+1}|\alpha|\left(\hat{a}e^{i(\phi-\theta_c)}-\hat{a}^\dagger e^{-i(\phi-\theta_c)}\right) = \frac{2\sqrt{2}\,|\alpha|}{2\bar{n}+1}\hat{x}_{\theta_c-\phi-\frac{\pi}{2}},$$

where the rotated quadrature operator $\hat{x}_\theta$ is defined as $\hat{x}_\theta = \hat{R}^\dagger(\theta)\hat{x}\hat{R}(\theta)$.

SLD for squeezed thermal states

The SLD for squeezed thermal states can be written as

$$\hat{L}_\phi = \frac{i(2\bar{n}+1)\sinh 2r}{2\bar{n}^2+2\bar{n}+1}\hat{R}(\phi)\hat{S}(\xi)\left(\hat{a}^{\dagger 2}e^{i\theta_s}-\hat{a}^2e^{-i\theta_s}\right)\hat{S}^\dagger(\xi)\hat{R}^\dagger(\phi).$$

This can be further simplified to be

$$\begin{aligned}\hat{L}_\phi &= \frac{i(2\bar{n}+1)\sinh 2r}{2\bar{n}^2+2\bar{n}+1}\hat{R}(\phi-\theta_s/2)\left(\hat{a}^{\dagger 2}-\hat{a}^2\right)\hat{R}^\dagger(\phi-\theta_s/2)\\ &= \frac{(2\bar{n}+1)\sinh 2r}{2\bar{n}^2+2\bar{n}+1}\left(\hat{x}_{\theta_s/2-\phi}\hat{p}_{\theta_s/2-\phi}+\hat{p}_{\theta_s/2-\phi}\hat{x}_{\theta_s/2-\phi}\right).\end{aligned}$$

Proof of optimality of homodyne detection for squeezed vacuum states

Here, we prove that homodyne detection is optimal for squeezed vacuum states explicitly by showing that homodyne detection satisfies Eqs. (3.76), (3.77). First, one can easily verify that Eq. (3.77) is automatically satisfied if the input state is

pure and the POVM measurement setup is composed of rank-one projectors.

Now, we show that $\mathrm{Tr}(\hat{\rho}_\phi \hat{\Pi}_x \hat{L}_\phi)$ is real. Squeezed vacuum states and the SLD operator for the states can be written as

$$\begin{aligned}
\hat{\rho}_\phi &= \hat{R}(\phi)\hat{S}(\xi)|0\rangle\langle 0|\hat{S}^\dagger(\xi)\hat{R}^\dagger(\phi) = \hat{S}(\xi e^{-2i\phi})|0\rangle\langle 0|\hat{S}^\dagger(\xi e^{-2i\phi}) \\
&= |\xi e^{-2i\phi}\rangle\langle \xi e^{-2i\phi}|. \\
\hat{L}_\phi &= i2\sinh 2r\hat{R}(\phi)\hat{S}(\xi)(\hat{a}^{\dagger 2}e^{i\theta_s} - \hat{a}^2 e^{-i\theta_s})\hat{S}^\dagger(\xi)\hat{R}^\dagger(\phi).
\end{aligned}$$

If we assume a homodyne detection with local oscillator angle $\psi/2$, $\hat{\Pi}_x = |x_{\psi/2}\rangle\langle x_{\psi/2}|$, which corresponds to the Gaussian measurement with the parameter $se^{i\psi}$,

$$\begin{aligned}
\mathrm{Tr}\big(\hat{\rho}_\phi \hat{\Pi}_x \hat{L}_\phi\big) &= i2\sinh 2r\mathrm{Tr}\big(|\xi e^{-2i\phi}\rangle\langle \xi e^{-2i\phi}|x_{\psi/2}\rangle\langle x_{\psi/2}| \\
&\qquad \hat{R}(\phi)\hat{S}(\xi)(\hat{a}^{\dagger 2}e^{i\theta_s} - \hat{a}^2 e^{-i\theta_s})\hat{S}^\dagger(\xi)\hat{R}^\dagger(\phi)\big) \\
&= i2\sqrt{2}\sinh 2re^{i\theta_s}\langle \xi e^{-2i\phi}|x_{\psi/2}\rangle\langle x_{\psi/2}|\hat{R}(\phi)\hat{S}(\xi)|2\rangle \\
&= i2\sqrt{2}\sinh 2re^{i(\theta_s-2\phi-\psi)}\langle \xi e^{i(-2\phi-\psi)}|x\rangle\langle x|\hat{S}(\xi e^{i(-2\phi-\psi)})|2\rangle \\
&= i2\sqrt{2}\sinh 2re^{i\chi}\langle re^{i\chi}|x\rangle\langle x|\hat{S}(re^{i\chi})|2\rangle, \\
&= i2\sqrt{2}\sinh 2re^{i\chi}\langle re^{i\chi}|x\rangle\Bigg(\cosh^{-\frac{5}{2}} r\langle x|\exp(-e^{i\chi}\tanh r\frac{\hat{a}^{\dagger 2}}{2})|2\rangle \\
&\qquad + \frac{\sqrt{2}}{2}e^{-i\chi}\tanh r\langle x|\hat{S}(re^{i\chi})|0\rangle\Bigg) \qquad (3.122)
\end{aligned}$$

where $|x_\theta\rangle = \hat{R}^\dagger(\theta)|x\rangle$, $\chi = \theta_s - 2\phi - \psi$. Here,

$$
\begin{aligned}
\langle x|\exp(-e^{i\chi}\tanh r\frac{\hat{a}^{\dagger 2}}{2})|2\rangle &= \sum_{m=0}^{\infty}\langle x|m\rangle\langle m|\exp(-e^{i\chi}\tanh r\frac{\hat{a}^{\dagger 2}}{2})|2\rangle \\
&= \frac{\sqrt{2}}{\pi^{1/4}}e^{-x^2/2}\sum_{m=0}^{\infty}\frac{(-1)^m e^{im\chi}\tanh^m r}{2^{2m+2}m!}H_{2m+2}(x) \\
&= -\frac{\sqrt{2}}{\pi^{1/4}}e^{-x^2/2}\sum_{m=0}^{\infty}(e^{i\chi}\tanh r)^m(m+1)L_{m+1}^{-\frac{1}{2}}(x^2) \\
&= \frac{\sqrt{2}}{\pi^{1/4}}e^{-x^2/2}\frac{e^{\frac{e^{i\chi}\tanh r}{1-e^{i\chi}\tanh r}x^2}}{2(1-e^{i\chi}\tanh r)^{5/2}}(e^{i\chi}\tanh r-1+2x^2)
\end{aligned}
\tag{3.123}
$$

where we have used the derivative of generating function of Laguerre polynomials,

$$
\sum_{m=0}^{\infty}t^m L_m^{(\alpha)} = \frac{1}{(1-t)^{\alpha+1}}e^{-\frac{tx}{1-t}}. \tag{3.124}
$$

Finally, after setting $\cos\chi = \tanh 2r$ and using

$$
\langle x|re^{i\chi}\rangle = \frac{1}{\pi^{1/4}\sqrt{\cosh r - e^{i\chi}\sinh r}}\exp\left[-\frac{x^2}{2}\frac{\cosh r + e^{i\chi}\sinh r}{\cosh r - e^{i\chi}\sinh r}\right], \tag{3.125}
$$

we obtain

$$
\mathrm{Tr}\big(\hat{\rho}_\phi\hat{\Pi}_x\hat{L}_\phi\big) = \frac{\exp(-x^2\cosh 2r)(2x^2\cosh 2r-1)\sqrt{\cosh 2r}}{\sqrt{2\pi}}, \tag{3.126}
$$

which is real. This proves that the homodyne detection with the appropriate local oscillator angle is optimal for squeezed thermal states.

Simplification of the operator $\hat{M}$

Here, we simplify the operator $\hat{M} = \hat{\rho}_1^{-1/2}\sqrt{\hat{\rho}_1^{1/2}\hat{\rho}_0\hat{\rho}_1^{1/2}}\hat{\rho}_1^{-1/2}$ with $\hat{\rho}_0 = e^{-\frac{(\hat{Q}-v_0)^{\mathrm{T}}G_0(\hat{Q}-v_0)}{2}}$ and $\hat{\rho}_1 = e^{-\frac{\hat{Q}^{\mathrm{T}}G_1\hat{Q}}{2}}$. Note that $e^{l^{\mathrm{T}}i\Omega\hat{Q}}e^{-\hat{Q}^{\mathrm{T}}G\hat{Q}/2} \propto e^{-(\hat{Q}-u)^{\mathrm{T}}G(\hat{Q}-u)/2}$ with $u = (e^{-i\Omega G} -$

$1)^{-1}l$, which is frequently used in this section. Simplifying $\hat{\rho}_0$ in the following way,

$$\hat{\rho}_0 = e^{\frac{1}{2}v_0^{\mathrm{T}} i\Omega e^{-i\Omega G_0} v_0} e^{(e^{-i\Omega G_0}v_0 - v_0)^{\mathrm{T}} i\Omega \hat{Q}} e^{-\frac{\hat{Q}^{\mathrm{T}} G_0 \hat{Q}}{2}} \propto e^{l_0^{\mathrm{T}} i\Omega \hat{Q}} e^{-\frac{\hat{Q}^{\mathrm{T}} G_0 \hat{Q}}{2}}$$

with $l_0 = (e^{-i\Omega G_0} - 1)v_0$, one can have

$$\hat{K} = \hat{\rho}_1^{1/2} \hat{\rho}_0 \hat{\rho}_1^{1/2} \propto e^{-\frac{\hat{Q}^{\mathrm{T}} G_1 \hat{Q}}{4}} e^{l_0^{\mathrm{T}} i\Omega \hat{Q}} e^{-\frac{\hat{Q}^{\mathrm{T}} G_0 \hat{Q}}{2}} e^{-\frac{\hat{Q}^{\mathrm{T}} G_1 \hat{Q}}{4}}.$$

Bringing all the displacement operators to the left side, one can further simplify the matrix $\hat{K}$ as

$$\hat{K} \propto e^{k^{\mathrm{T}} i\Omega \hat{Q}} e^{-\frac{\hat{Q}^{\mathrm{T}} G_K \hat{Q}}{2}},$$

where we have defined $k = e^{-i\Omega G_1/2} l_0$ and

$$e^{-\frac{\hat{Q}^{\mathrm{T}} G_K \hat{Q}}{2}} = e^{-\frac{\hat{Q}^{\mathrm{T}} G_1 \hat{Q}}{4}} e^{-\frac{\hat{Q}^{\mathrm{T}} G_0 \hat{Q}}{2}} e^{-\frac{\hat{Q}^{\mathrm{T}} G_1 \hat{Q}}{4}}. \tag{3.127}$$

Defining u_K as $(e^{-i\Omega G_K} - 1)u_K = k$, the operator $\hat{K}$ takes the Gibbs-exponential form, written as

$$\hat{K} \propto e^{-(\hat{Q} - u_K)^{\mathrm{T}} G_K (\hat{Q} - u_K)/2},$$

where u_K is a real vector. The operator $\hat{M} = \hat{\rho}_1^{-1/2} \sqrt{\hat{K}} \hat{\rho}_1^{-1/2}$ can thus be written as

$$\hat{M} \propto e^{\frac{\hat{Q}^{\mathrm{T}} G_1 \hat{Q}}{4}} e^{l_1^{\mathrm{T}} i\Omega \hat{Q}} e^{-\frac{\hat{Q}^{\mathrm{T}} G_K \hat{Q}}{4}} e^{\frac{\hat{Q}^{\mathrm{T}} G_1 \hat{Q}}{4}},$$

where $l_1 = (e^{-i\Omega G_K/2} - 1)u_K$. Again, we bring all the displacement operators to the left side,

$$\hat{M} \propto e^{m^{\mathrm{T}} i\Omega \hat{Q}} e^{-\frac{\hat{Q}^{\mathrm{T}} G_{\mathrm{M}} \hat{Q}}{2}},$$

where $m = e^{i\Omega G_1/2} l_1$ and

$$e^{-\frac{\hat{Q}^{\mathrm{T}} G_{\mathrm{M}} \hat{Q}}{2}} = e^{\frac{\hat{Q}^{\mathrm{T}} G_1 \hat{Q}}{4}} e^{-\frac{\hat{Q}^{\mathrm{T}} G_K \hat{Q}}{4}} e^{\frac{\hat{Q}^{\mathrm{T}} G_1 \hat{Q}}{4}}. \tag{3.128}$$

When $G_{\mathrm{M}} = 0$, corresponding to the case that $G_0 = G_1$, we obtain $\hat{M} \propto e^{m^{\mathrm{T}} i\Omega \hat{Q}}$, where $m = e^{i\Omega G_1/2} l_1$ is a pure imaginary vector. Especially if $G_0 = G_1 = \oplus_{j=1}^{n} g_j \mathbb{1}_2$, we obtain $m = -i[\oplus_{j=1}^{n} \tanh(g_j/2)\mathbb{1}_2]\Omega v_0$. If $G_0 = G_1$ are not diagonal, we introduce a symplectic transformation that diagonalizes the Gibbs matrices, $G_0 = G_1 = (S^{-1})^{\mathrm{T}}[\oplus_{j=1}^{n} g_j \mathbb{1}_2]S^{-1}$, or, equivalently, leading to $e^{-\hat{Q}^{\mathrm{T}} G_0 \hat{Q}/2} = \hat{U}_S e^{-\hat{Q}^{\mathrm{T}}[\oplus_{j=1}^{n} g_j \mathbb{1}_2]\hat{Q}/2} \hat{U}_S^{\dagger}$, where $\hat{U}_S \hat{Q} \hat{U}_S^{\dagger} = S^{-1}\hat{Q}$. As a consequence,

$$\hat{M} \propto \hat{U}_S e^{v_0^{\mathrm{T}}[\oplus_{j=1}^{n} \tanh(g_j/2)\mathbb{1}_2]\hat{Q}} \hat{U}_S^{\dagger} = e^{v_0^{\mathrm{T}}[\oplus_{j=1}^{n} \tanh(g_j/2)\mathbb{1}_2]S^{-1}\hat{Q}},$$

where we have used Eq. (3.65).

When $G_{\mathrm{M}} \neq 0$, the operator $\hat{M}$ can always be written in the Gibbs-exponential form,

$$\hat{M} \propto e^{-(\hat{Q}-u_{\mathrm{M}})^{\mathrm{T}} G_{\mathrm{M}} (\hat{Q}-u_{\mathrm{M}})/2}. \tag{3.129}$$

where $u_{\mathrm{M}} = (e^{-i\Omega G_{\mathrm{M}}} - 1)^{-1} m$. Therefore, $\hat{M}$ can be written as

$$\hat{M} \propto \exp\left[-\frac{1}{2}\hat{Q}^{\mathrm{T}} G_{\mathrm{M}} \hat{Q} - \vec{v}_{\mathrm{M}}^{\mathrm{T}} \hat{Q}\right].$$

Here, $\vec{v}_{\rm M}=0$ if $v_0=0$, $\vec{v}_{\rm M}=G_{\rm M}u_{\rm M}$ if $G_0\neq G_1$, and $G_{\rm M}=0$ and $\vec{v}_{\rm M}=(S^{-1})^{\rm T}[\oplus_{j=1}^{n}\tanh(g_j/2)\mathbb{1}_2]v_0$ if $G_0=G_1$. From Eqs. (3.127) and (3.128), it is clear that $G_{\rm M}$ is the solution of

$$e^{i\Omega G_{\rm M}}e^{i\Omega G_1}e^{i\Omega G_{\rm M}}=e^{i\Omega G_0},$$

and the vector $u_{\rm M}$ is written as

$$\begin{aligned}u_{\rm M}=&(e^{-i\Omega G_{\rm M}}-1)^{-1}e^{i\Omega G_1/2}(e^{-i\Omega G_K/2}-1)(e^{-i\Omega G_K}-1)^{-1}\\&\times e^{-i\Omega G_1/2}(e^{-i\Omega G_0}-1)v_0.\end{aligned}$$

Finally, in order to return to the original problem between two general Gaussian states, $\hat{\rho}_0=e^{-\frac{(\hat{Q}-u_0)^{\rm T}G_0(\hat{Q}-u_0)}{2}}$ and $\hat{\rho}_1=e^{-\frac{(\hat{Q}-u_1)^{\rm T}G_1(\hat{Q}-u_1)}{2}}$, we simply introduce a displacement operator $\hat{D}(u_1)$ with $u_0-u_1=v_0$, so that, by using Eq. (4), we obtain $\hat{M}$ of the original problem written as,

$$\hat{M}\propto\hat{D}(u_1)\exp\left[-\frac{1}{2}\hat{Q}^{\rm T}G_{\rm M}\hat{Q}-\vec{v}_{\rm M}^{\rm T}\hat{Q}\right]\hat{D}^{\dagger}(u_1). \tag{3.130}$$

Full equation for d_1 and d_2.

We simplify Eq. (3.70) for the single-mode case by assuming G_0 and G_1 to be Gibbs matrices of a general single-mode Gaussian state and a thermal state, respectively. Expanding the matrices by Pauli matrices and using

$$\cosh g_1=\frac{2\bar{n}_1+1}{2\bar{n}_1(\bar{n}_1+1)},\quad \sinh g_1=\frac{2\bar{n}_1^2+2\bar{n}_1+1}{2\bar{n}_1(\bar{n}_1+1)},$$

the left hand side of Eq. (3.70) is written as

$$L_0\mathbb{1}_2 + L_1\hat{\sigma}_x + L_2\hat{\sigma}_y,$$

where

$$L_0 = (d_1+d_2)\frac{2\bar{n}_1+1}{2\bar{n}_1(\bar{n}_1+1)}\frac{\sinh 2\sqrt{d_1d_2}}{2\sqrt{d_1d_2}} + \frac{2\bar{n}_1^2+2\bar{n}_1+1}{2\bar{n}_1(\bar{n}_1+1)}\cosh 2\sqrt{d_1d_2}, \quad (3.131)$$

$$L_1 = -i(d_1-d_2)\left(\frac{2\bar{n}_1^2+2\bar{n}_1+1}{2\bar{n}_1(\bar{n}_1+1)}\frac{\sinh 2\sqrt{d_1d_2}}{2\sqrt{d_1d_2}} + \frac{2\bar{n}_1+1}{2\bar{n}_1(\bar{n}_1+1)}\frac{(d_1+d_2)\sinh^2\sqrt{d_1d_2}}{2d_1d_2}\right), \quad (3.132)$$

$$L_2 = \frac{2\bar{n}_1^2+2\bar{n}_1+1}{2\bar{n}_1(\bar{n}_1+1)}\frac{(d_1-d_2)^2-(d_1+d_2)^2\cosh 2\sqrt{d_1d_2}}{4d_1d_2} - \frac{2\bar{n}_1+1}{2\bar{n}_1(\bar{n}_1+1)}\frac{2\sqrt{d_1d_2}(d_1+d_2)\sinh 2\sqrt{d_1d_2}}{4d_1d_2}. \quad (3.133)$$

The right-hand side, on the other hand, is written as

$$R_0\mathbb{1}_2 + R_1\hat{\sigma}_x + R_2\hat{\sigma}_y,$$

where

$$R_0 = \frac{2\bar{n}_0^2+2\bar{n}_0+1}{2\bar{n}_0(\bar{n}_0+1)}, \quad (3.134)$$

$$R_1 = \frac{i(2\bar{n}_0+1)\sinh 2r_0}{2\bar{n}_0(\bar{n}_0+1)}, \quad (3.135)$$

$$R_2 = -\frac{(2\bar{n}_0+1)\cosh 2r_0}{2\bar{n}_0(\bar{n}_0+1)}. \quad (3.136)$$

Equations of (3.131) to enable d_1 and d_2 to be written as functions of $r_0, \bar{n}_0$, and $\bar{n}_1$.

Pure state limit of optimal measurement

Consider a single-mode state with a diagonal covariance matrix of

$$V = \begin{pmatrix} \frac{1}{2}+\varepsilon & 0 \\ 0 & \frac{1}{2}+\varepsilon \end{pmatrix} .$$

Such state is pure in the limit of $\varepsilon \to 0$. The analysis can be trivially extended to a non-diagonal case by adding a squeezing operation SVS^{T}. One can find that

$$e^{i\Omega G} = \frac{W-\mathbb{1}}{W+\mathbb{1}} = \left(\frac{1}{\varepsilon}+1\right)P + \varepsilon Q + O(\varepsilon^2) , \tag{3.137}$$

$$e^{-i\Omega G} = \frac{W+\mathbb{1}}{W-\mathbb{1}} = \left(\frac{1}{\varepsilon}+1\right)Q + \varepsilon P + O(\varepsilon^2) , \tag{3.138}$$

where $W = -2Vi\Omega$ and

$$P = \frac{1}{2}\begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}, \quad Q = \mathbb{1} - P.$$

Note $P^2 = P$ and $Q^2 = Q$, so they are projection operators. The Gibbs matrix of the operator $\hat{M}$ satisfies

$$e^{i\Omega G_1} = e^{-i\Omega G_{\mathrm{M}}} e^{i\Omega G_0} e^{-i\Omega G_{\mathrm{M}}} . \tag{3.139}$$

In the limit where G_1 corresponds to the pure state $|\psi_1\rangle\langle\psi_1|$, we use Eqs. (3.137) to write $e^{i\Omega G_1} \approx \frac{P}{\varepsilon}$. Then a possible solution for $e^{-i\Omega G_{\mathrm{M}}}$ is $e^{-i\Omega G_{\mathrm{M}}} \approx \alpha P$ because the above equation (3.139) becomes $\alpha^2 P e^{i\Omega G_0} P = e^{i\Omega G_1} \approx \frac{P}{\varepsilon}$, which is approximately true for some α. Indeed, for any state $\hat{\rho}_0$ with nonzero overlap with $\hat{\rho}_1$, it is $Pe^{i\Omega G_1}P \propto P$. Therefore, $e^{-i\Omega G_{\mathrm{M}}} \propto P \propto e^{i\Omega G_1}$, namely, $\hat{M} \propto \mathbb{1} - |\psi_1\rangle\langle\psi_1|$, where all approximations made in the above equations refer to the corrections that disappear in the limit of $\varepsilon \to 0$. The operator $\hat{M}$ implies that the measurement with projectors

$\{|\psi_1\rangle\langle\psi_1|, \mathbb{1} - |\psi_1\rangle\langle\psi_1|\}$ is optimal.

The relation between optimal measurements for quantum fidelity and quantum Fisher information

Let $\hat{\rho}_0 = \hat{\rho} + d\hat{\rho}$ and $\hat{\rho}_1 = \hat{\rho}$. For simplicity, we assume $\hat{\rho}$ is a full-rank state, which implies that $\hat{\rho}_0$ and $\hat{\rho}_1$ are full-rank states. Let $\sqrt{\hat{\rho}_1^{1/2}\hat{\rho}_0\hat{\rho}_1^{1/2}} = \hat{\rho} + \hat{X}$, where $X \propto d\hat{\rho}$. Taking the square, we get

$$\hat{\rho}_0^{1/2}\hat{\rho}_1\hat{\rho}_0^{1/2} = \hat{\rho}^2 + \hat{\rho}^{1/2}d\hat{\rho}\hat{\rho}^{1/2} = \hat{\rho}^2 + \hat{\rho}\hat{X} + \hat{X}\hat{\rho},$$

leading to $\hat{\rho}^{1/2}d\hat{\rho}\hat{\rho}^{1/2} = \hat{\rho}\hat{X} + \hat{X}\hat{\rho}$. For $\hat{\rho} = \sum_k p_k|k\rangle\langle k|$ with $\langle k|l\rangle = \delta_{kl}$, one can show

$$\hat{X}_{nm} = \frac{\sqrt{p_n}\sqrt{p_m}}{p_n + p_m} d\hat{\rho}_{nm}.$$

When the states are full rank, the first optimality condition becomes $E_x^{1/2}(1 - \mu_x\hat{\rho}_1^{-1/2}\sqrt{\hat{\rho}_1^{1/2}\hat{\rho}_0\hat{\rho}_1^{1/2}}\hat{\rho}_1^{-1/2}) = 0$. In the limit of small $d\hat{\rho}$,

$$\begin{aligned}\hat{\rho}_1^{-1/2}\sqrt{\hat{\rho}_1^{1/2}\hat{\rho}_0\hat{\rho}_1^{1/2}}\hat{\rho}_1^{-1/2} &= 1 + \hat{\rho}^{-1/2}\hat{X}\hat{\rho}^{-1/2} \\ &= 1 + \sum_{n,m}\frac{d\hat{\rho}_{nm}}{p_n + p_m}|n\rangle\langle m| = 1 + \hat{L}d\theta/2,\end{aligned}$$

where $\hat{L}_\theta d\theta = 2\sum_{n,m} d\hat{\rho}_{nm}/(p_n + p_m)|n\rangle\langle m|$ is the SLD operator, so that the condition becomes

$$\hat{E}_x^{1/2}(1 - \mu_x\hat{\rho}_1^{-1/2}\sqrt{\hat{\rho}_1^{1/2}\hat{\rho}_0\hat{\rho}_1^{1/2}}\hat{\rho}_1^{-1/2}) = \hat{E}_x^{1/2}(1 - \mu_x(1 + \hat{L}_\theta d\theta/2)) = 0.$$

This results in

$$\hat{E}_x^{1/2}(1-\lambda_x\hat{L}_\theta)=0$$

with a constant λ_x, which is equivalent to the optimal condition of Eq. (2.48) for quantum Fisher information.

Now, we turn to the second condition. For two quantum states that are infinitesimally close, Eq. (3.64) can be simplified as

$$\begin{aligned}\mathrm{Tr}[U\hat{\rho}_0^{1/2}\hat{E}_x\hat{\rho}_1^{1/2}] &= \mathrm{Tr}[\sqrt{\hat{\rho}_1^{1/2}\hat{\rho}_0\hat{\rho}_1^{1/2}}\hat{\rho}_1^{-1/2}\hat{E}_x\hat{\rho}_1^{1/2}]\\ &= \mathrm{Tr}[(1+\hat{L}_\theta d\theta/2)\hat{E}_x\hat{\rho}]\in\mathbb{R}.\end{aligned}$$

One can immediately see that this is equivalent to Eq. (2.48).

Infinitesimal limit of G_{M} matrix

Consider the problem of estimating parameter θ. The matrix G_{M} is given by the solution of

$$e^{i\Omega G_{\mathrm{M}}}=e^{-i\Omega G_\theta/2}\sqrt{e^{i\Omega G_\theta/2}e^{i\Omega G_{\theta+d\theta}}e^{i\Omega G_\theta/2}}e^{-i\Omega G_\theta/2}.$$

Since the zeroth order of the two matrices G_θ and $G_{\theta+d\theta}$ is equal in an infinitesimal limit of $d\theta$, the zeroth order of G_{M} is zero. Therefore, one can write $i\Omega G_{\mathrm{M}}=Cd\theta$ for some unknown matrix C and, similarly, $i\Omega G_\theta=A$ and $i\Omega G_{\theta+d\theta}=A+Bd\theta$ for some matrices A and B. From the above equation, it can be shown that C is the

solution of

$$e^{A+Bd\theta} \approx e^{A} + e^{A}Cd\theta + Cd\theta e^{A} + O(d\theta)^2.$$

Using the notation from Ref. [60], one may write $e^{i\Omega G_\theta} = \frac{W_\theta - \mathbb{1}}{W_\theta + \mathbb{1}}$ and expand the matrices W_θ as $W_{\theta+d\theta} = W_A + W_B d\theta$ with $W_\theta = W_A$. Therefore,

$$e^{A+Bd\theta} = e^{i\Omega G_{\theta+d\theta}} = \mathbb{1} - 2\frac{\mathbb{1}}{W_{\theta+d\theta}+\mathbb{1}} = e^{A} + \frac{d\theta}{2}(e^{A}-\mathbb{1})W_B(e^{A}-\mathbb{1}) + O(d\theta)^2$$

and C is the solution of

$$e^{A}C + Ce^{A} = \frac{1}{2}(e^{A}-\mathbb{1})W_B(e^{A}-\mathbb{1}) \ ,$$

or C can be implemented into the discrete Lyapunov equation written as

$$C - W_\theta^{-1} C W_\theta^{-1} = W_\theta^{-1}\frac{\partial W_\theta}{\partial \theta}W_\theta^{-1},$$

for which $(W_\theta+\mathbb{1})C(W_\theta-\mathbb{1}) + (W_\theta-\mathbb{1})C(W_\theta+\mathbb{1}) = 2W_B$ is used. The solution of the Lyapunov equation is

$$C = \sum_{m=0}^{\infty} W_\theta^{-m-1}\frac{\partial W_\theta}{\partial \theta}W_\theta^{-m-1},$$

and thus,

$$G_{\mathrm{M}} = i\Omega\sum_{m=0}^{\infty} W_\theta^{-m-1}\frac{\partial W_\theta}{\partial \theta}W_\theta^{-m-1}d\theta.$$

Especially when $\partial\bar{n}_j/\partial\theta = 0$ and isothermal states, i.e. $\bar{n}_j = \bar{n}$ for all j,

$$C = \sum_{m=0}^{\infty}(-1)^{m+1}W_\theta^{-2m-2}\frac{\partial W_\theta}{\partial\theta} = -\frac{1}{2(2\bar{n}^2+2\bar{n}+1)}\frac{\partial W_\theta}{\partial\theta}, \tag{3.140}$$

where we have used $W_\theta^2 = (2\bar{n}+1)^2\mathbb{1}_{2n}$. Thus,

$$G_{\mathrm{M}} = -\frac{1}{2\bar{n}^2+2\bar{n}+1}\Omega\frac{\partial V_\theta}{\partial\theta}\Omega d\theta.$$

It can also be shown that from the definition of C and $W_\theta = -2V_\theta i\Omega$, G_{M} is also the solution of

$$4V_\theta G_{\mathrm{M}} V_\theta + \Omega G_{\mathrm{M}}\Omega + 2d\theta\frac{\partial V_\theta}{\partial\theta} = 0. \tag{3.141}$$

Writing in the basis, in which V_θ is symplectically diagonalized, one can recover the previous result [89],

$$(G_{\mathrm{M}})_{jk} = \frac{2V_\theta^{\mathrm{s}}\frac{\partial V_\theta^s}{\partial\theta}V_\theta^{\mathrm{s}} - \Omega\frac{\partial V_\theta^s}{\partial\theta}\Omega/2}{\lambda_j^2\lambda_k^2 - 1}d\theta \tag{3.142}$$

where the superscript of s denotes operators being transformed by the symplectic operator S, λ_j's are the symplectic eigenvalues of V_θ, and S is a symplectic matrix that diagonalizes V_θ.

The vector u_{M} for an infinitesimal $d\theta$ is written as

$$\begin{aligned} u_{\mathrm{M}} &= (-i\Omega G_{\mathrm{M}})^{-1}e^{i\Omega G_\theta/2}(e^{-i\Omega G_\theta/2}-1)(e^{-2i\Omega G_\theta}-1)^{-1} \\ &\times e^{-i\Omega G_\theta/2}(e^{-i\Omega G_\theta}-1)\frac{\partial u_\theta}{\partial\theta}d\theta = G_{\mathrm{M}}^{-1}V_\theta^{-1}\frac{\partial u_\theta}{\partial\theta}d\theta/2, \end{aligned}$$

where we have used $e^{i\Omega G_\theta} = \frac{W_\theta - \mathbb{1}}{W_\theta + \mathbb{1}}$. Thus,

$$\vec{v}_{\rm M} = G_{\rm M} u_{\rm M} = V_\theta^{-1} \frac{\partial u_\theta}{\partial \theta} d\theta/2. \tag{3.143}$$

As a final remark, we highlight that Eq. (3.141) with $G_{\rm M}$ and $\vec{v}_{\rm M}$ facilitates the derivation of the quantum Fisher information, being made as

$$\begin{aligned} H(\theta) &= {\rm Tr}[\hat{D}^\dagger(u_\theta)\hat{\rho}_\theta \hat{D}(u_\theta)(\hat{Q}^{\rm T} G_{\rm M} \hat{Q} - 2\vec{v}_{\rm M}^{\rm T}\hat{Q} + \nu)^2]/d\theta^2 \\ &= {\rm Tr}[\hat{\rho}_\theta^0 (\hat{Q}^{\rm T} G_{\rm M} \hat{Q})^2 + 4(\vec{v}_{\rm M}^{\rm T}\hat{Q})^2 + \nu(\hat{Q}^{\rm T} G_{\rm M} \hat{Q}) + \nu^2]/d\theta^2 \\ &= -{\rm Tr}\left[\frac{\partial V_\theta}{\partial \theta} G_{\rm M}\right]/(d\theta) + \frac{\partial u_\theta}{\partial \theta} V_\theta^{-1} \frac{\partial u_\theta}{\partial \theta}, \end{aligned}$$

where $\hat{\rho}_\theta^0 = \hat{D}^\dagger(u_\theta)\hat{\rho}_\theta \hat{D}(u_\theta)$ is a Gaussian state with zero mean and the same covariance matrix as $\hat{\rho}_\theta$. We also have used ${\rm Tr}[\hat{\rho}_\theta^0 \hat{Q}_n \hat{Q}_m \hat{Q}_l \hat{Q}_k] = \sum_{(mlk)} {\rm Tr}[\hat{\rho}_\theta^0 \hat{Q}_n \hat{Q}_m]{\rm Tr}[\hat{\rho}_\theta^0 \hat{Q}_l \hat{Q}_k]$, where (mlk) denotes a cyclic permutation, and ${\rm Tr}[\hat{\rho}_\theta^0 \hat{Q}_n \hat{Q}_m] = V_{nm} + i\Omega_{nm}/2$ [56]. Note that the method we provide above can be straightforwardly applied to multi-parameter cases so as to derive quantum Fisher information matrix.

Chapter 4

Bayesian Error Certification

The contents of the present chapter are largely based on [Y. S. Teo, C. Oh, and H. Jeong, "Bayesian error regions in quantum estimation I: analytical reasonings," New J. Phys. **20,** 093009 (2018)] Ref. [90], [C. Oh, Y. S. Teo, and H. Jeong, "Bayesian error regions in quantum estimation II: region accuracy and adaptive methods," New J. Phys. **20,** 093010 (2018)] Ref. [91], [C. Oh, Y. S. Teo, and H. Jeong, "Probing Bayesian credible regions intrinsically: a feasible error certification for physical systems," Phys. Rev. Lett. **123,** 040602 (2019)] Ref. [92], and [C. Oh, Y. S. Teo, and H. Jeong, "Efficient Bayesian credible-region certification for quantum-state tomography," Phys. Rev. A **100,** 012345 (2019)] Ref. [93].

4.1 Introduction

Precise description of quantum systems is crucial for various applications of quantum information processing. If one wants to build a quantum computer or to construct a quantum network, one has to precisely characterize the input quantum state and quantum gates with the environment and measurement devices. Hence, quantum estimation theory is one of the most important ingredients for quantum information processing. While much study about quantum estimation theory has been focusing on estimators, relatively less attention has been paid on error certification for estimators. Although error certification is already well established in statistic, a serious obstacle emerges when one is interested in a high-dimensional system of

a complicated parameter space. Especially when an unknown parameter of interest is around the boundary of the parameter space, one has to explore the complicated boundary of the space, which may be intractable such as quantum state tomography which deals with the positive-density-matrix space. In this chapter, to circumvent the main obstacle, we derive an analytical expression for Bayesian error regions in asymptotic regime and propose an efficient numerical algorithm that operates in non-asymptotic regime. Before presenting the main results, we begin with introducing a general protocol of quantum estimation and study about Bayesian error certification in the protocol to characterize the estimation error region.

4.2 Bayesian Error Region

Let us consider a situation where we would like to estimate a d-dimensional parameter $\vec{r} = (r_1, ..., r_d)^{\mathrm{T}}$ encoded in quantum systems. The parameter can be one characterizing a quantum state or a physical quantity with which a probe state has interacted. This is a typical situation in quantum state tomography, quantum process tomography, and quantum metrology. To characterize the parameter, we measure the quantum state $\hat{\rho}(\vec{r})$ by a POVM $\{\hat{E}_k\}$ ($\hat{E}_k \geq 0$, $\sum_k \hat{E}_k = \mathbb{1}$) to obtain data $\mathbb{D}$ according to the measurement probabilities $p_k = p_k(\vec{r}) = \mathrm{Tr}[\hat{\rho}(\vec{r})\hat{E}_k]$. In this section, we assume that the measurement outcome x is discrete for simplicity although it is straightforward to generalize to continuous-valued measurement. Based on the dataset $\mathbb{D} = \{n_k\}$ ($\sum_k n_k = N$), n_k denoting the number of clicks for each POVM element $\hat{E}_k$, we infer $\vec{r}$ using standard tools in statistical inference. Particularly, we focus on an important estimator, namely the maximum likelihood estimator $\vec{r}_{\mathrm{ML}}$, which chooses a point that maximizes the likelihood function $L(\mathbb{D}|\vec{r}) = \prod_k p_k^{n_k}$ in the reconstruction space $\mathcal{R}_0$. In typical circumstances, the maximum likelihood esti-

mator is unique in the space of interest. When one conveys the result of estimation, the estimator is not sufficient but it is necessary to provide it with an appropriate error region (or interval for a single-parameter case). While there may be various ways to define the error region, we shall use a so-called Bayesian credible region [94], which is defined as follows: Let us denote by $(d\vec{r})$ the prior probability of the infinitesimal volume of state in $\mathcal{R}_0$, satisfying $\int_{\mathcal{R}_0}(d\vec{r}) = 1$. In Bayesian statistics, the prior probability is a way to introduce the prior knowledge of the experimentalist before conducting the experiment. Then the size $s_{\mathcal{R}}$ of a region $\mathcal{R}$ is

$$s_{\mathcal{R}} = \int_{\mathcal{R}} (d\vec{r}). \tag{4.1}$$

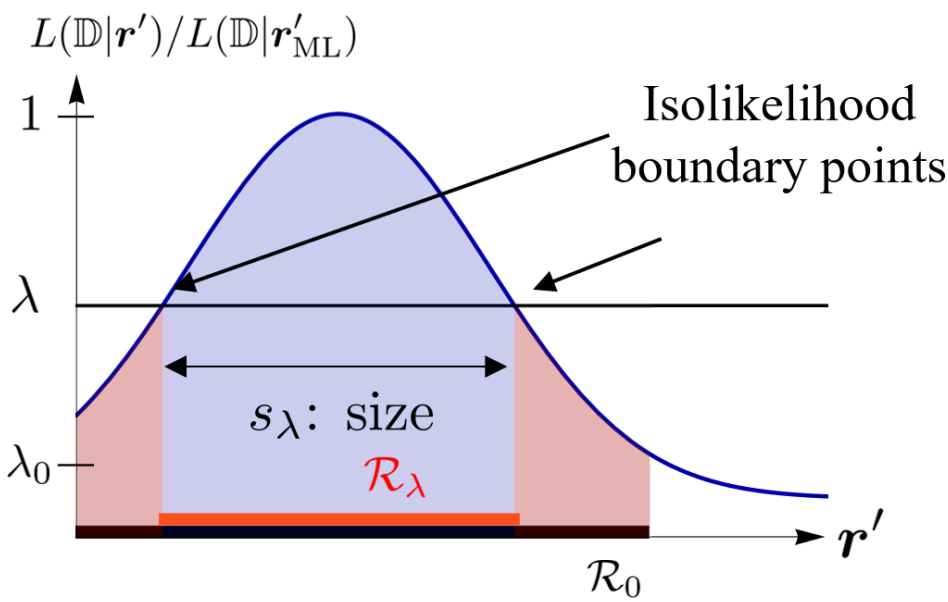


Figure 15: Size and credibility for a single-parameter case. The credible region is parametrized by $0 \leq \lambda \leq 1$, which decides the isolikelihood boundary points. Once the isolikelihood boundary points are determined, the size and credibility can be easily computed as shown in the figure.

The credible region and its credibility characterized by $0 \leq \lambda \leq 1$ are defined as

$$s_\lambda = \int_{\mathcal{R}_0} (d\vec{r}) \chi_\lambda(\vec{r}), \quad c_\lambda = \frac{1}{L(\mathbb{D})} \int_{\mathcal{R}_0} (d\vec{r}) \chi_\lambda(\vec{r}) L(\mathbb{D}|\vec{r}), \tag{4.2}$$

where $\chi_\lambda(\vec{r}) = \eta\left(L(\mathbb{D}|\vec{r}) - \lambda L(\mathbb{D}|\vec{r}_{\mathrm{ML}})\right)$ and $\eta(.)$ is the heaviside theta function. The size and credibility are illustrated in Fig. 15 for a single-parameter case. The credibility of a region represents the probability that the unobserved parameter value lies on the region. There is an important relation between the two quantities such as

$$c_\lambda = \frac{\lambda s_\lambda + \int_\lambda^1 d\lambda' s'_\lambda}{\int_0^1 d\lambda' s'_\lambda}, \tag{4.3}$$

which allows us to calculate the credibility by using the knowledge of the size s_λ, and consequently calculating s_λ is crucial. In a practical circumstance, an experimentalist chooses a credibility (e.g. 95%) in advance and provide a size of the credible region $\mathcal{R}$ after the experiment. Equipped with the λ-parametrized size and credibility, one finds a lambda that gives the prechosen credibility and then obtain the size corresponding to the λ.

There is another important type of Bayesian region which is called a plausible region [95], which is not subjective to the choice of the experimentalist. We say that a point $\vec{r}$ is plausibly the true value if there is evidence such that its normalized posterior probability $L(\mathbb{D}|\vec{r})/L(\mathbb{D})$ is larger than its prior probability $p(\vec{r})$. A plausible region $\mathcal{R}$ is defined as one containing all plausible points of $\vec{r}$ only, which can be characterized by a critical value λ_{crit} [96],

$$\lambda_{\mathrm{crit}} = \int_0^1 d\lambda' s'_\lambda, \tag{4.4}$$

for which $L(\mathbb{D}|\vec{r} \in \partial\mathcal{R}_{\lambda=\lambda_{\mathrm{crit}}}) = L(\mathbb{D})$ To facilitate this understanding, we give a

short instructive proof by noting that the constant $L(\mathbb{D})$ is simply related to the size function s_λ by the definition

$$\begin{aligned} L(\mathbb{D}) &= \int (d\vec{r}')\, L(\mathbb{D}|\vec{r}') = \int (d\vec{r}') \int_0^{L(\mathbb{D}|\vec{r}')} dx' \\ &= L_{\max} \int (d\vec{r}') \int_0^1 d\lambda'\, \eta\left(L(\mathbb{D}|\vec{r}') - \lambda' L_{\max}\right) = L_{\max} \int_0^1 d\lambda'\, s_{\lambda'}, \end{aligned} \tag{4.5}$$

so that the assignment $L(\mathbb{D}|\vec{r} \in \partial\mathcal{R}_{\lambda=\lambda_{\text{crit}}}) \equiv \lambda_{\text{crit}} L_{\max} = L(\mathbb{D})$ gives the expression for λ_{crit}.

Although it is crucial to provide the size and credibility of the credible region together with the estimate, it is often demanding to calculate them for high-dimensional parameter even using numerical methods. Particularly, when the parameter space $\mathcal{R}$ is a space of quantum states, which is the case for quantum state tomography, the calculation of the size and credibility is known as an NP-hard problem because of the complicated boundary of the space $\partial\mathcal{R}$ [19]. In this chapter, we provide an analytical approximation for them and an efficient numerical method for the direct calculation.

We will focus on quantum state tomography case, but the method is general and applicable to any multiparameter estimation. Before we proceed, let us introduce parameterization of density matrices. Density matrices of the D-dimensional Hilbert space can be parameterized by $d = D^2 - 1$ parameters due to the positivity, $\hat{\rho} \geq 0$ and the unit-trace condition $\mathrm{Tr}[\hat{\rho}] = 1$. One way to parametrize density matrices is to introduce d number of $D \times D$ traceless Hermitian matrices $\{\hat{\Pi}_j\}_{j=1}^d$ satisfying the orthogonality condition $\mathrm{Tr}[\hat{\Pi}_j \hat{\Pi}_k] = \delta_{ij}$ for $j,k \geq 1$ with $\hat{\Pi}_0 = \mathbb{1}/D$. The parameters corresponding to a density matrix can be obtained by $r_j = \mathrm{Tr}[\hat{\rho}\hat{\Pi}_j]$. Conversely, the

density matrix is recovered by the relation,

$$\hat{\rho}(\vec{r}) = \hat{\Pi}_0 + \sum_j r_j \hat{\Pi}_j. \tag{4.6}$$

4.3 Analytical approximation

In this section, we derive the analytical approximated expressions for the size and credibility of a credible region $\mathcal{R}$. We assume that the parameter space $\mathcal{R}_0$ of $\vec{r}$ is a convex space, which is a typical situation in quantum estimation experiments. Especially when the parameter space is the space of all legitimate quantum states $\hat{\rho}$ of a given dimension, i.e., $\hat{\rho} \geq 0$ and $\mathrm{Tr}[\hat{\rho}] = 1$, the parameter space is obviously convex. The motivation of deriving analytical approximation of the size and credibility is that the computation of the size and credibility is known to be an NP-hard problem [19] because of the complicated influence from the parameter space boundary $\partial\mathcal{R}_0$. In order to derive approximated expressions for the size and credibility, we also assume that the number of copies N is large enough for the Gaussian approximation of likelihood function being valid. Finally, although the choice of a prior distribution is essential in Bayesian statistics, we take the uniform prior distribution to reveal interesting properties of the size and credibility avoiding technical complications. The influence of the choice of a prior distribution becomes negligible when the number of copies N is sufficiently large.

We present the results for three different circumstances. The first case is that the credible region is inside of the parameter space. The second case is that the credible region is truncated by the parameter space with the estimator being inside of the region. The final case is that the estimator is on the boundary and the credible region is again truncated. Although there may be another cases than the above three

cases, in the limit of large N, these three cases will be the typical situations that we encounter in practice.

4.3.1 Case 1: Interior-point theory for a full likelihood

Case 1 exhibits the situations where the estimator is inside of the parameter space $\mathcal{R}_0$ and the number of copies N is sufficiently large that the region is not truncated by the boundary of the parameter space as described in Fig. 16. For a d-dimensional parameter $\vec{r}$, if $\vec{r} \notin \partial\mathcal{R}_0$, then for a given data $\mathbb{D}$ collected with large number of copies of N, we approximate the likelihood around the maximum likelihood estimator as

$$L(\mathbb{D}|\vec{r}) \simeq L_{\max} \exp\left(-\frac{1}{2}\vec{\Delta}^{\mathrm{T}}(\vec{r}) F_{\mathrm{ML}} \vec{\Delta}(\vec{r})\right), \tag{4.7}$$

which is a Gaussian function [97] centered at the estimator $\vec{r}_{\mathrm{ML}}$ ($\vec{\Delta}(\vec{r}) \equiv \vec{r} - \vec{r}_{\mathrm{ML}}$) that has a covariance equal to the inverse of d-dimensional Fisher information matrix,

$$F_{ij}(\vec{r}) = \sum_k \frac{N}{p_k} \frac{\partial p_k}{\partial r_i} \frac{\partial p_k}{\partial r_j}, \tag{4.8}$$

$$\vec{\Delta}(\vec{r}) = \vec{r} - \vec{r}_{\mathrm{ML}}, \tag{4.9}$$

at $\vec{r}_{\mathrm{ML}}$, i.e., $F_{\mathrm{ML}} \equiv F(\vec{r}_{\mathrm{ML}})$. By the virtue of the Gaussian shape of the likelihood function, we may regard the credible region $\mathcal{R} = \mathcal{R}_\lambda$ as a hyperellipsoid obeying the equation,

$$\vec{\Delta}^{\mathrm{T}}(\vec{r}) F_{\mathrm{ML}} \vec{\Delta}(\vec{r}) \leq -2\ln\lambda. \tag{4.10}$$

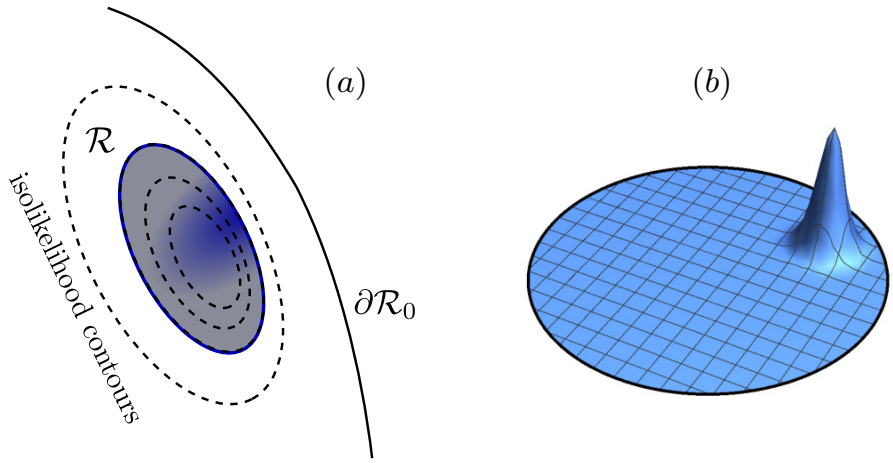


Figure 16: Credible region contained in $\mathcal{R}_0$ (Case 1). (a) Case 1 arises when the estimate lies in $\mathcal{R}_0$ and it is sufficiently far from the boundary $\partial\mathcal{R}_0$. The latter condition is satisfied if $N \gg 1$ when the estimate is inside of $\mathcal{R}$. (b) The likelihood function vanishes at the boundary.

In this case, it is straightforward to find the expression of s_λ and c_λ, which are given as

$$s_\lambda = \frac{V_d}{V_{\mathcal{R}_0}}(-2\ln\lambda)^{d/2}/\sqrt{|F_{\rm ML}|}, \tag{4.11}$$

$$c_\lambda = 1 - \frac{\Gamma(d/2, -\ln\lambda)}{(d/2-1)!}, \tag{4.12}$$

where $V_d = \pi^{d/2}/(d/2)!$ is the volume of the $(d-1)$-sphere of unit radius, and $\Gamma(a,y)$ is the order-a upper incomplete Gamma function of y (See Appendix for the derivation). Note that the size function diverges as $s_\lambda \sim (-\ln\lambda)^{d/2}$ as λ goes to zero. This divergence arises due to the Gaussian approximation of a likelihood function ignoring the boundary of the parameter space (See Sec. 4.3.4 for details).

Under the Gaussian approximation in Eq. (4.7), we can easily obtain

$$\lambda_{\rm crit} = \sqrt{|2\pi F_{\rm ML}^{-1}|}/V_{\mathcal{R}_0}\,, \tag{4.13}$$

and so the plausible region's size and credibility are given by

$$s_{\lambda_{\text{crit}}} = \frac{V_d}{V_{\mathcal{R}_0}} \left[-\ln\left(\frac{|2\pi F_{\text{ML}}^{-1}|}{V_{\mathcal{R}_0}^2} \right) \right]^{d/2} |F_{\text{ML}}|^{-1/2},$$
$$c_{\lambda_{\text{crit}}} \approx 1 - \frac{(d/2)^{d/2-1}}{(d/2-1)!} \frac{(\ln N)^{d/2-1}}{N^{d/2}}, \tag{4.14}$$

We note here that for the plausible region, the scaling behaviors of $s_{\lambda_{\text{crit}}}$ and $c_{\lambda_{\text{crit}}}$ with N are more complicated. For i.i.d. copies, we have $s_{\lambda_{\text{crit}}} \sim (\ln N + \cdots)^{d/2}/N^{d/2}$ and $1 - c_{\lambda_{\text{crit}}} \sim (\ln N)^{d/2-1}/N^{d/2}$, where the appearance of logarithmic scaling comes from picking the largest credible region that contains all plausible parameters.

4.3.2 Case 2: Interior-point theory for a truncated likelihood

Case 2 arises when the estimate is close to the boundary of the parameter space and the number of copies N is not sufficiently large, which leads to the truncation of the likelihood function, which is depicted in Fig. 17. It frequently occurs in quantum

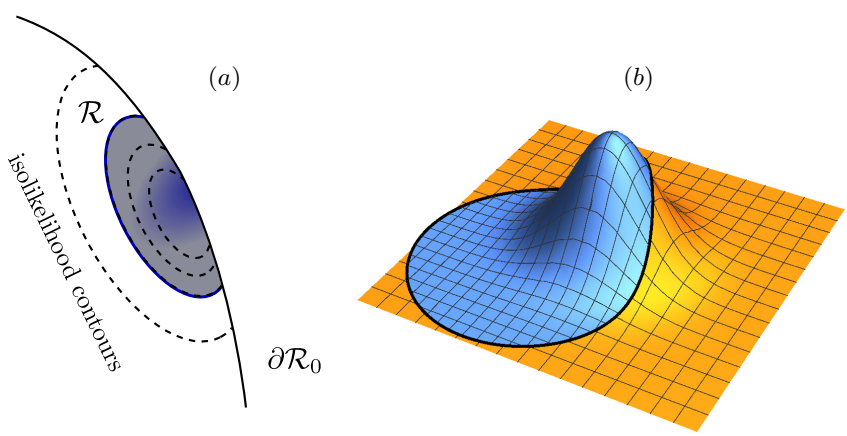


Figure 17: Credible region truncated by the boundary $\partial\mathcal{R}_0$ (Case 2). (a) Case 2 arises when the estimate is close to the boundary of the space $\mathcal{R}$ and the number of copies N is not sufficiently large. (b) The likelihood function is truncated by the boundary, which complicates the analysis of the region $\mathcal{R}$.

state tomography since a typical aim is to generate a pure state ideally or a state sufficiently close to a pure state, which is on the boundary of the quantum state space or adjacent to it. This case is not as simple as the previous case because of the truncation of the likelihood function by the boundary of the parameter space. Even when we use the Gaussian approximation of the likelihood function, we need to know the boundary precisely, which may be computationally hard, for example, for a quantum state space. In order to resolve the difficulty caused by the complicated boundary $\partial\mathcal{R}_0$, we will assume that the boundary is smooth enough to be considered as a hyperplane P as illustrated in Fig. 18. This assumption is valid for the Bloch sphere corresponding to a qubit case. This hyperplane P has a normal vector $\vec{n}$ that is perpendicular to the isolikelihood contour at $\vec{r}_P$, which is the point having the largest likelihood on $\partial\mathcal{R}_0 \cup \mathcal{R}$ (see Appendix for the detail of finding $\vec{r}_P$).

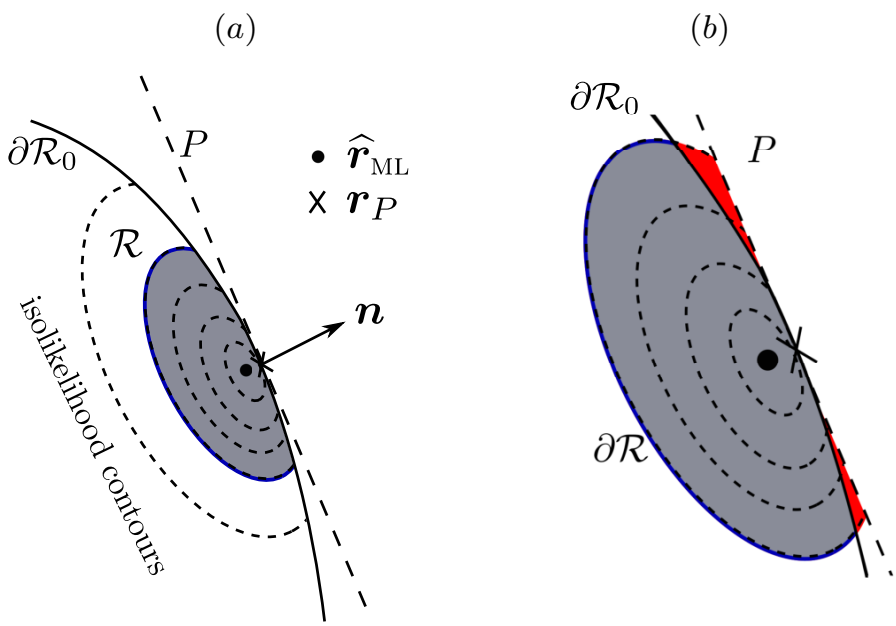


Figure 18: We approximate the boundary $\partial\mathcal{R}_0$ as a tangential hyperplane P characterized by a normal vector $\vec{n}$. (b) The red region represents the error caused by the approximation. The approximation is valid when the boundary is smooth enough that the red region is small.

After obtaining $L_{\max}^{(\partial\mathcal{R}_0)} = L_{\max}\exp\left(-\vec{\Delta}^{\mathrm{T}}(\vec{r}_P)F_{\mathrm{ML}}\vec{\Delta}(\vec{r}_P)/2\right)$, it is possible to show that the estimated fraction γ of the hyperellipsoid truncation is given in terms of the regularized incomplete beta function $I_y(a,b)$ as

$$\begin{aligned}\gamma &= 1 - I_{\frac{1-l}{2}}\left(\frac{d+1}{2},\frac{d+1}{2}\right),\\ l &= \min\left\{\sqrt{\frac{\ln\lambda_{\mathrm{int}}}{\ln\lambda}},1\right\},\lambda_{\mathrm{int}} = \frac{L_{\max}^{(\partial\mathcal{R}_0)}}{L_{\max}},\end{aligned} \tag{4.15}$$

with which we arrive at the generalized interior-point statement $s_\lambda \approx \gamma V_{d,\lambda}/V_{\mathcal{R}_0}$. For $\lambda = \lambda_{\mathrm{int}}$, $\gamma = 1$ characterizes the untruncated size expression in Eq. (4.11). The approximate credibility has no simple closed form but may be computed with the relation in Eq. (4.3) efficiently.

Details of the derivation of (4.15) is given in Appendix. More relevantly, let us briefly discuss the volume estimate characterized by the fraction in (4.15) in broad terms. For this, we emphasize that $\partial\mathcal{R}_0$ can be a highly sophisticated surface with corners and edges. For instance, if $\mathcal{R}_0$ is the space of quantum states of Hilbert-space dimension $D = 2$—the qubit space—, then $\partial\mathcal{R}_0$ that is enforced by the operator positivity constraint is a 2-sphere. However if $D > 2$, $\partial\mathcal{R}_0$ is generally a complicated surface with corners and edges, for the convex space is "neither a polytope nor a smooth body." [98] For such boundaries, the approximated volume fraction offered by (4.15) is an overestimate of the actual fraction for any finite N due to the convex nature of $\mathcal{R}_0$. If however $\vec{r}_{\mathrm{ML}}$ lies on a smooth $\partial\mathcal{R}\cap\partial\mathcal{R}_0$ to which we may approximate the local boundary with a hyperplane, then in the limit of large N, this overestimate approaches the exact answer, which applies, for instance, to the qubit space.

It is easy to see that this methodology gives the asymptotically exact, not an over-

estimated volume fraction in single-parameter estimation ($d = 1$), as the $\partial\mathcal{R} \cap \partial\mathcal{R}_0$ intersects P at exactly the point r_P. We note, however, that the likelihood near r_P is exponential in r. The corresponding quantities s_λ, c_λ and λ_{crit} also admit analytical expressions

$$\begin{aligned} s_\lambda &= V_{1,\lambda} + \eta(\lambda_{\text{int}} - \lambda)\frac{\ln\lambda - \ln\lambda_{\text{int}}}{V_{\mathcal{R}_0}|\Delta(r_P)|}\,, \\ c_\lambda &= \frac{|\Delta(r_P)|\sqrt{2F_{\text{ML}}}\left[\sqrt{\pi} - \Gamma(1/2, -\ln\lambda)\right] + \eta(\lambda_{\text{int}} - \lambda)(\lambda - \lambda_{\text{int}})}{\sqrt{2\pi F_{\text{ML}}}|\Delta(r_P)| - \lambda_{\text{int}}}\,, \\ \lambda_{\text{crit}} &= \frac{\sqrt{2\pi}}{V_{\mathcal{R}_0}\sqrt{F_{\text{ML}}}} - \frac{\lambda_{\text{int}}}{V_{\mathcal{R}_0}|\Delta(r_P)|}\,, \end{aligned} \tag{4.16}$$

which can be derived by evaluating the one-dimensional version of the integral in (4.44). The limiting case in which $\lambda_{\text{int}} \to 0$ can be confirmed right away.

4.3.3 Case 3: Boundary-point theory

Case 3 represents the case when the ML estimator is on the boundary of the parameter space, which inevitably leads to the truncation of the likelihood function by the boundary. It also frequently arises in quantum state tomography with the same reason as Case 2.

With the statistical conviction that the true parameter $\vec{r}$ is close to the boundary-point ML estimator $\vec{r}_{\text{ML}}$, we may again expand $\ln L(\mathbb{D}|\vec{r})$ to second order,

$$\begin{aligned} \ln L(\mathbb{D}|\vec{r}) &\approx \ln L_{\max} + \vec{\Delta}(\vec{r})^{\mathrm{T}}\vec{g}_{\text{ML}} - \frac{1}{2}\vec{\Delta}(\vec{r})^{\mathrm{T}}F_{\text{ML}}\vec{\Delta}(\vec{r})\,, \\ \vec{g}_{\text{ML}} &= \vec{\partial}_{\text{ML}}\ln L(\mathbb{D}|\vec{r}_{\text{ML}})\,, \end{aligned} \tag{4.17}$$

where now evidently the first order does not vanish since $\vec{r}_{\text{ML}}$ is on the boundary and $L_{\max}$, the maximum likelihood value for $\mathcal{R}_0$, is less than the exterior maximal

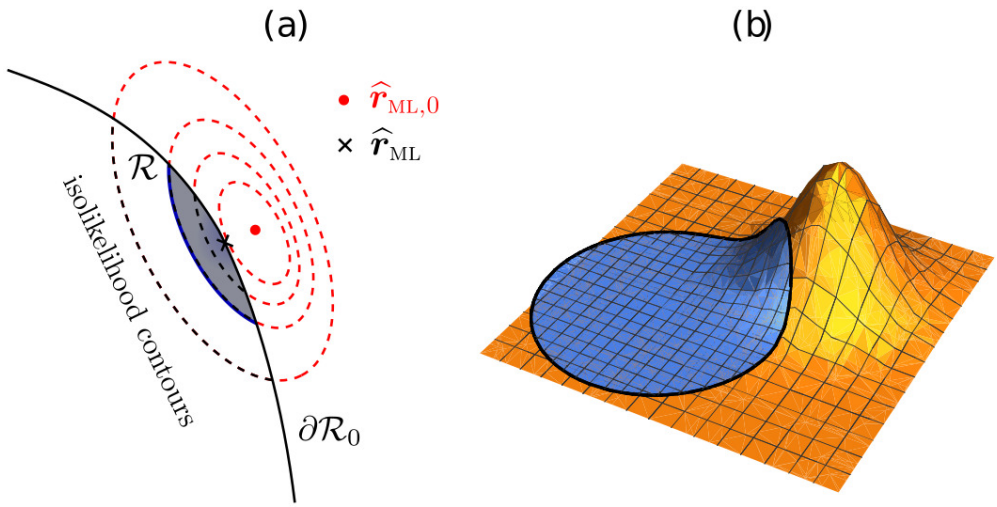


Figure 19: Credible region truncated by the boundary $\partial\mathcal{R}_0$ (Case 3). (a) Case 3 arises when the ML estimate is on the boundary, regardless of the number of copy N. (b) The likelihood function is truncated by the boundary $\partial\mathcal{R}_0$.

value $L_{\max,\mathrm{G}} = L_{\max} \exp\left(\vec{g}_{\mathrm{ML}}^{\mathrm{T}} F_{\mathrm{ML}}^{-1} \vec{g}_{\mathrm{ML}}/2\right)$ for the approximated Gaussian function. Similar to Case 2, we may introduce a hyperplane P' that contains $\vec{r}_{\mathrm{ML}}$ and has a normal $\vec{n}' = \vec{g}_{\mathrm{ML}}$ that is orthogonal to the Gaussian isocontour intersecting $\vec{r}_{\mathrm{ML}}$. The volume $V_{\mathcal{R}}$ of $\mathcal{R}$ can then be (over)estimated with the shaded volume presented in Fig. 20. For smooth boundaries, this estimate once more becomes asymptotically exact.

Interestingly, we point out the role changes for some relevant quantities: $\vec{r}_{\mathrm{ML}}$ now takes the place of $\vec{r}_P$ as the boundary point in the hyperplane and $L_{\max,\mathrm{G}}$ is now replacing $L_{\max}$ to be the largest possible likelihood. We may next define $\lambda_{\mathrm{eff}} = \lambda L_{\max}/L_{\max,\mathrm{G}} < 1$ to be the effective "λ" that characterizes the approximated Gaussian likelihood with respect to the actual one. Finally, after realizing that the estimated volume for $V_{\mathcal{R}}$ falls on the opposite side of the hyperplane in contrast with

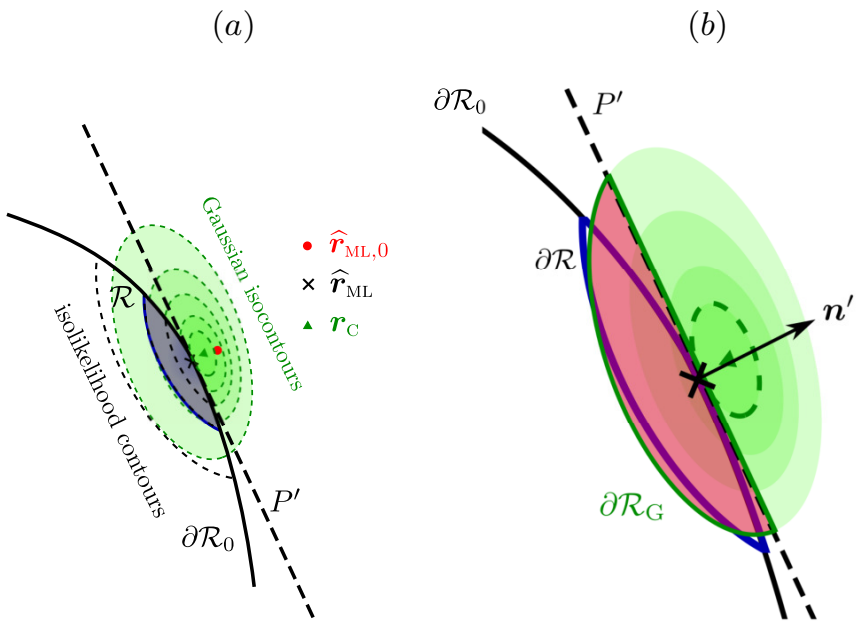


Figure 20: (a) We apply a similar approximation as case 2. The boundary is assumed to be smooth so that a tangential hyperplane replaces the boundary for truncation. (b) The red region is now considered as the credible region which contains some points outside of the original credible region.

that in Case 2, we can write down the fraction

$$\gamma' = I_{\frac{1-l'}{2}}\left(\frac{d+1}{2}, \frac{d+1}{2}\right), \quad l' = \sqrt{\frac{\ln \lambda_{\mathrm{bd}}}{\ln \lambda_{\mathrm{eff}}}} \leq 1, \quad \lambda_{\mathrm{bd}} = \frac{L_{\max}}{L_{\max,\mathrm{G}}}, \tag{4.18}$$

of the total hyperellipsoidal volume that contributes to the approximate size estimate $s_\lambda \approx \gamma' V_{d,\lambda}/V_{\mathcal{R}_0}$.

The asymptotically exact region quantities for $d = 1$ can be obtained by taking the aforementioned role changes into account. This suggests the replacements in (4.15) (*from Case 2 to Case 3*) $|\Delta(r_P)| \to g_{\mathrm{ML}}$, $\lambda_{\mathrm{int}} \to \lambda_{\mathrm{eff}}/\lambda$ and $\lambda \to \lambda_{\mathrm{eff}}$, which immediately gives rise to

$$s_\lambda = -\frac{\ln \lambda}{V_{\mathcal{R}_0}\, g_{\mathrm{ML}}}, \qquad c_\lambda = 1 - \lambda, \quad \lambda_{\mathrm{crit}} = \frac{1}{V_{\mathcal{R}_0}\, g_{\mathrm{ML}}}, \tag{4.19}$$

with the appropriate sign changes due to the opposite "side" of the truncation to Case 2.

4.3.4 Remarks on logarithmic divergence and $V_{\mathcal{R}_0}$

In all the Bayesian-region property formulas developed [(4.11), (4.14), (4.15), (4.18)] as a means to provide an asymptotic size and credibility certification for the ML estimator $\vec{r}_{\mathrm{ML}}$, the size formulas exhibit logarithmic divergences—$s_\lambda \sim (-\ln\lambda)^{d/2}$. This feature stems from the Gaussian approximations in (4.7) and (4.17) that ignores the parameter-space boundary $\partial\mathcal{R}_0 \backslash (\partial\mathcal{R} \cap \partial\mathcal{R}_0)$ that falls on "the other side" of the joint one (if there is any). These approximations are strictly valid for the likelihood portion sufficiently near the maximum. For extremely small λ values or high credibilities, the asymptotic size formulas either give highly conservative (much larger) estimates for s_λ, or gradually exceeds the unit physical upper bound.

This reinforces the importance of measuring a sufficiently large number of copies N such that most portion of the likelihood is approximately part of a Gaussian function. Put differently, there exists the sufficient condition

$$N \gg N_{\min} \quad \text{where} \quad \vec{\Delta}^{\mathrm{T}}(\vec{r}_P) F_{\mathrm{ML}} \vec{\Delta}(\vec{r}_P)\Big|_{N=N_{\min}} = -2\ln\lambda \tag{4.20}$$

given a particularly interesting range of λ. This is geometrically equivalent to keeping the tails of the likelihood from penetrating the boundary $\partial\mathcal{R}_0 \neq \partial\mathcal{R} \cap \partial\mathcal{R}_0$ as much as possible, so that the logarithmic divergence has no visible effect on the size estimation.

Furthermore, all operational formulas invoke the knowledge of the volume $V_{\mathcal{R}_0}$ of $\mathcal{R}_0$ under the uniform-prior assertion. For parameter estimation settings with simple convex boundary constraints this can be found very easily. For instance, $V_{\mathcal{R}_0}$ for an

a priori uniformly distributed phase $a \leq \theta \leq b$ is $b-a$. In the case of quantum-state characterization $V_{\mathcal{R}_0}$ is much more complicated, but known to have closed forms for specialized priors [99, 100]. Just as an example we shall take the prior to be the uniform distribution over the continuous space $\mathcal{R}_0 = \mathcal{M}_D$ of D-dimensional complex positive matrices of unit trace that represent quantum states $\hat{\rho}$, or the *Lebesgue prior* for this space. For this prior, the volume for the $(d = D^2 - 1)$-dimensional state parameter $\vec{r}$ has the closed form [100]

$$V_{\mathcal{M}_D} = \frac{\pi^{D(D-1)/2}}{(D^2-1)} \prod_{j=1}^{D-1} j!\,. \tag{4.21}$$

4.3.5 Examples in quantum-state tomography

In quantum-state tomography of a single-qubit $(D = 2)$, states in the space $\mathcal{R}_0 = \mathcal{M}_2$ of statistical operators can be represented as the 2×2 positive semidefinite matrix

$$\hat{\rho} = \begin{pmatrix} r_1 & r_2 - ir_3 \\ r_2 + ir_3 & 1 - r_1 \end{pmatrix} \tag{4.22}$$

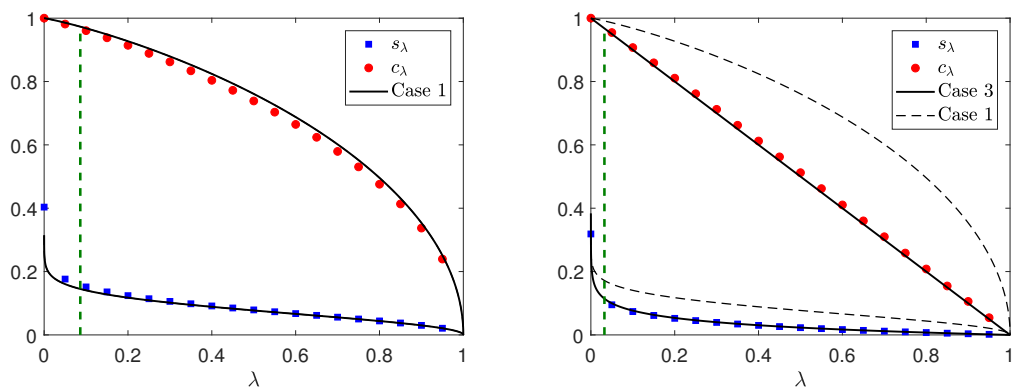
in terms of the $(d = 3)$-dimensional state parameter $\vec{r}$. The qubit space also has the nice property that the boundary $\partial\mathcal{M}_2$ is smooth—it is the surface of a 2-sphere. This implies that $\partial\mathcal{M}_2$ is smooth and can eventually be described by a hyperplane for sufficiently large N. We shall see that the expressions in (4.15) and (4.18) indeed exactly describe the actual size and credibility in this limit.

To verify our theoretical results, we may consider three different classes of qubit states. For the numerical computation of s_λ and c_λ, one may first generate a set of qubit states for the integrations by performing uniform rejection sampling. In accordance with the Lebesgue measure, the parametrization in (4.22) allows a uni-

form sampling on the parameter ranges $0 \leq r_1 \leq 1$ and $-1 \leq r_2, r_3 \leq 1$ depending on the class of qubit states, where the range of r_1 trivially maintains the unit-trace constraint. From this set of random operators, rejection sampling is then carried out by simply eliminating randomly generated operators this way that are not positive semidefinite. These matrices may be numerically filtered out by verifying efficiently that their Cholesky decompositions do not exist [101]. In what follows, the yield percentage from uniform rejection sampling, that is the percentage ratio of the number of positive operators out of the total number of sampled Hermitian operators, is calculated explicitly for each of the three classes.

4.3.5.1 One-parameter qubit $(d = 1)$

Suppose that $r_2 = r_3 = 0$, so that only the single parameter $r = r_1$ needs to be estimated. The POVM considered shall then be the simple $(M = 2)$-outcome projective



(a) (b)

Figure 21: Single-parameter qubit estimation. (a) For a one-dimensional qubit in a mixed state specified by $r = 0.99$, $N = 30$ is sufficiently large for boundary effects of $\mathcal{M}_2$ to vanish, which explains the accuracy of the interior-point expressions in Eq.(4.11) and (4.12). The plausible region, of 0.966 credibility, is defined with $\lambda_{\text{crit}} = 0.08$ (dashed line). (b) In the case where $r_{\text{ML}} = 1$ is in $\partial\mathcal{R} \cap \partial\mathcal{R}_0$, while $N = 30$ avoids the tail-boundary effects at $r = 0$, the part at $r = 1$ modifies the behaviors of s_λ and c_λ according to (4.18). Here, the plausible region, of 0.967 credibility, is constructed with $\lambda_{\text{crit}} = 0.03$.

measurement onto the eigenstates of $\hat{\sigma}_z = |0\rangle\langle 0| - |1\rangle\langle 1|$ that directly probes r,

$$p_1 = \langle 0|\hat{\rho}|0\rangle = r, \qquad p_2 = \langle 1|\hat{\rho}|1\rangle = 1 - r. \tag{4.23}$$

The value of $V_{\mathcal{M}_2^{(d=1)}}$ is simply equal to one, the Lebesgue length of the interval $0 \leq r \leq 1$. As the Lebesgue prior is defined for the entire $\mathcal{M}_2^{(d=1)}$, we have $L(\mathbb{D}|r=0) = L(\mathbb{D}|r=1) = 0$ such that only Case 1 and 3 apply. Rejection sampling is certainly not necessary for such a simple class of states. Figure 21 studies the behaviors of theoretical results for these two cases.

4.3.5.2 Two-parameter qubit $(d = 2)$

Suppose we know that only $r_3 = 0$, then $\hat{\rho}$ lies in the plane $(r_1 - 1/2)^2 + r_2^2 \leq 1/4$. The volume $V_{\mathcal{M}_2^{(d=2)}}$ of this two-parameter subspace $\mathcal{M}_2^{(d=2)}$ can then be easily calculated to be

$$V_{\mathcal{M}_2^{(d=2)}} = \int_{\left(r_1' - \frac{1}{2}\right)^2 + r_2'^2 \leq \frac{1}{4}} dr_1'\, dr_2' = \frac{\pi}{4}, \tag{4.24}$$

and the yield percentage through uniform rejection sampling for these states is therefore equal to 39.27%. The POVM employed is the $M = 4$ "crosshair" measurement consisting of projections onto the eigenstates of both Pauli operators $\hat{\sigma}_z$ and $\hat{\sigma}_x = |+\rangle\langle +| - |-\rangle\langle -|$:

$$\begin{aligned} p_1 &= \frac{1}{2}\langle 0|\hat{\rho}|0\rangle = \frac{r_1}{2}, \qquad p_3 = \frac{1}{2}\langle +|\hat{\rho}|+\rangle = \frac{1}{2}(1 + 2r_2) \\ p_2 &= \frac{1}{2}\langle 1|\hat{\rho}|1\rangle = \frac{1 - r_1}{2}, \quad p_4 = \frac{1}{2}\langle -|\hat{\rho}|-\rangle = \frac{1}{2}(1 - 2r_2). \end{aligned} \tag{4.25}$$

Figure 22 illustrates the validity of our theory.

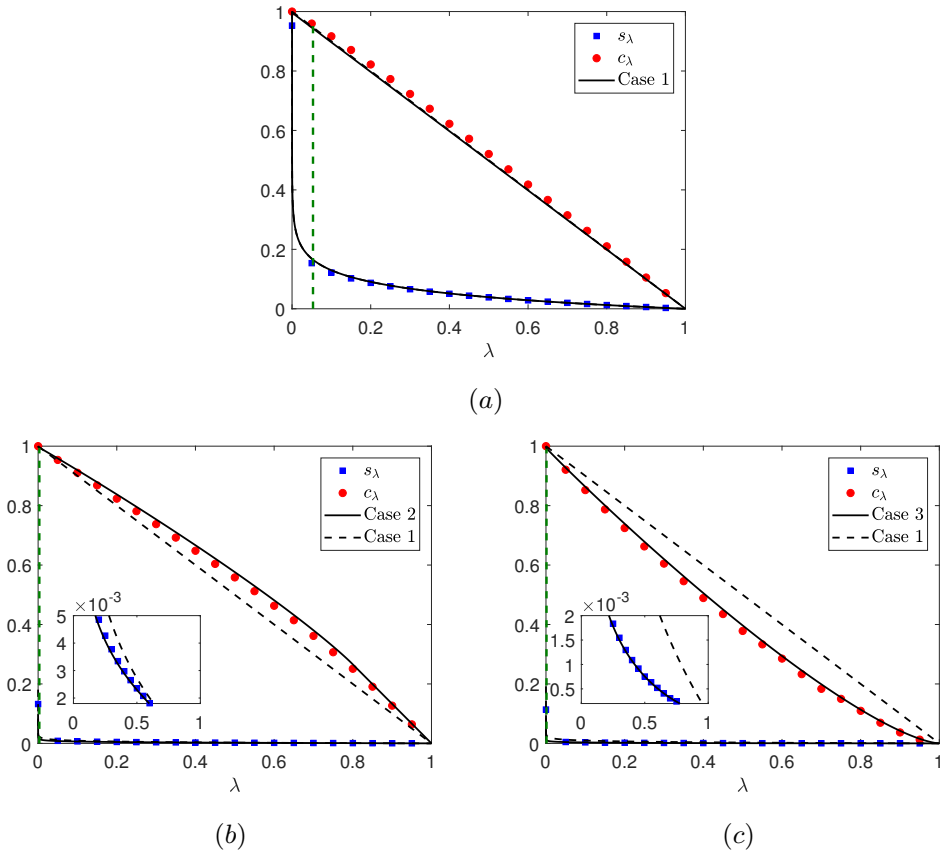


Figure 22: Two-parameter qubit estimation. (a) Tomography is carried out on a two-dimensional qubit which quantum state is represented by $\vec{r} = (0.8\ 0.1)^{\mathrm{T}}$ inside the Bloch ball. The interior ML estimator $\vec{r}_{\mathrm{ML}}$ for $N = 50$ is far enough from the boundary so that the results of Case 1 apply. The plausible region of 0.957 credibility is defined by $\lambda_{\mathrm{crit}} \approx 0.05$. (b) For a different state $\vec{r} = (0.8\ 0.4)^{\mathrm{T}}$, $\vec{r}_{\mathrm{ML}}$ for $N = 500$ is near $\partial\mathcal{R} \cap \partial\mathcal{R}_0$ and the generalized solutions for Case 2 clearly resolve the curvature modifications on s_λ (see also the inset for a blown up plot of s_λ) and c_λ. Here $\lambda_{\mathrm{crit}} \approx 0.0031$ gives a plausible region of 0.994 credibility. (c) Similarly, whenever Case 3 happens, the modifications result in $\lambda_{\mathrm{crit}} \approx 0.0014$ for a plausible region of 0.99 credibility with a given dataset.

4.3.5.3 Three-parameter qubit $(d = 3)$

For full qubit tomography, we require a minimum set of $M = 2^2 = 4$-outcome informationally complete (IC) POVM to completely characterize the qubit quantum state. We consider the popular tetrahedron POVM composed of the following four symmetrically oriented measurement outcomes (symmetric IC POVM or

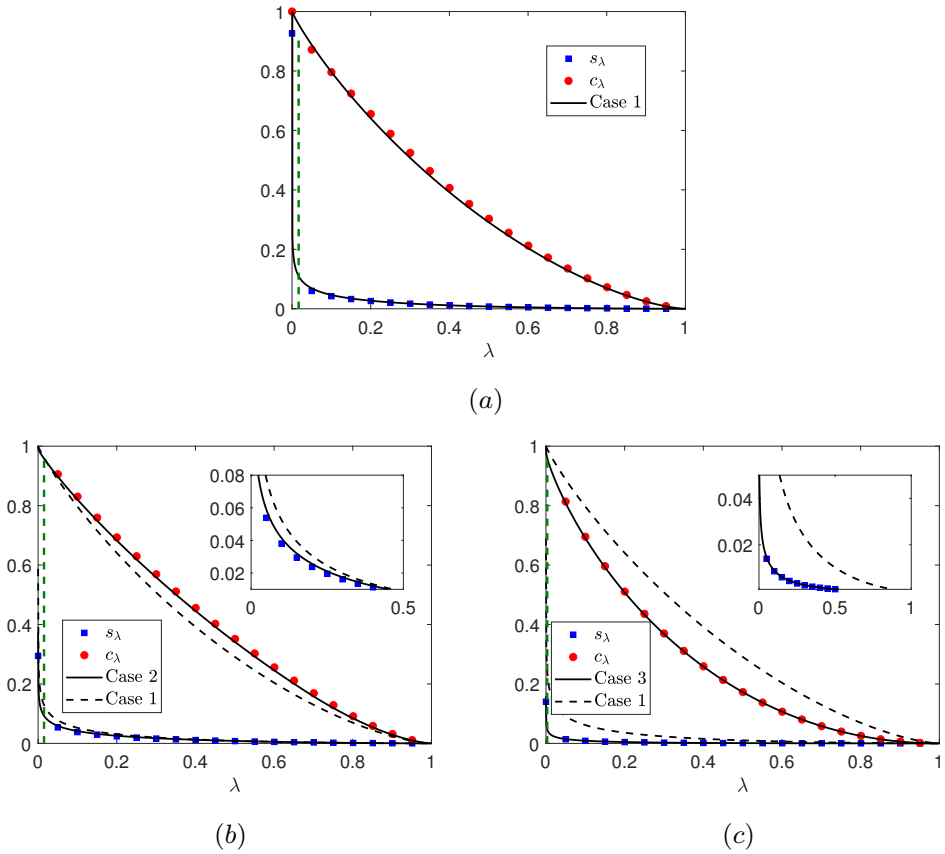


Figure 23: Full qubit estimation. Credible-region quantities are plotted for tomography on the complete qubit characterized by $\vec{r} = (0.8, 0.4, 0.1)$ using the tetrahedron measurement by measuring data made up of $N = 90$ copies. (a) In the optimistic Case 1, the plausible region, of 0.927 credibility, is defined by $\lambda_{\text{crit}} \approx 0.017$. (b) With the same N, boundary effects begin to influence the characteristics of both region size and credibility when $\vec{r}_{\text{ML}}$ is near $\partial\mathcal{R} \cap \partial\mathcal{M}_2$ as in Case 2, giving a plausible region of 0.963 credibility at $\lambda_{\text{crit}} \approx 0.015$ for a particular dataset. (c) Case 3 happens rather frequently as well, with an example dataset that gives a plausible region of 0.964 credibility at $\lambda_{\text{crit}} \approx 0.0033$.

SIC POVM)

$$\vec{a}_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \vec{a}_2 = \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}, \vec{a}_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}, \vec{a}_4 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}. \quad (4.26)$$

This qubit POVM as well as its extensions to higher dimensions is known to consti-

tute an optimal class of measurements in quantum information under certain conditions [102, 103, 104]. The volume of the $\mathcal{M}_2$ under the Lebesgue prior can be shown to be $\pi/6$ either by setting $D = 2$ in (4.21) or simply calculating the spherical volume

$$V_{\mathcal{M}_2} = \int_{\left(r_1' - \frac{1}{2}\right)^2 + r_2'^2 + r_3'^2 \leq \frac{1}{4}} dr_1' \, dr_2' \, dr_3' = \frac{4}{3}\pi \left(\frac{1}{2}\right)^3 = \frac{\pi}{6}. \tag{4.27}$$

The yield percentage for $\mathcal{M}_2$ is 13.09%. The analyses of all three cases are described in Fig. 23.

4.3.5.4 Qutrit

The qutrit is the next simplest quantum system of dimension $D = 3$ which state

$$\hat{\rho} = \begin{pmatrix} r_1 & r_3 + ir_4 & r_5 + ir_6 \\ r_3 - ir_4 & r_2 & r_7 + ir_8 \\ r_5 - ir_6 & r_7 - ir_8 & 1 - r_1 - r_2 \end{pmatrix} \tag{4.28}$$

can be completely characterized by the $(d = 3^2 - 1 = 8)$-dimensional state parameter $\vec{r}$. Therefore the minimum number of POVM outcomes needed to estimate $\vec{r}$ is $M = 9$. The volume of the qutrit space, according to (4.21), is $V_{\mathcal{M}_3} = \pi^3/20160$. To compute s_λ and c_λ over $\mathcal{M}_3$, we may again perform uniform rejection sampling over the ranges $0 \leq r_1, r_2 \leq 1$ and $-1 \leq r_3, \ldots, r_8 \leq 1$. This time, we see that the yield percentage for $\mathcal{M}_3$ is significantly lower than that for $\mathcal{M}_2$—2.4×10^{-3} % to be more precise for the uniform Lebesgue prior. Although it is possible to sample diagonal entries of $\hat{\rho}$ such that $\mathrm{Tr}[\hat{\rho}] = 1$ (*i.e.* sampling on *any* unit simplex) without sample wastage by renormalizing exponentially distributed random real numbers [105, 106], inevitably as D grows, the method of rejection sampling for

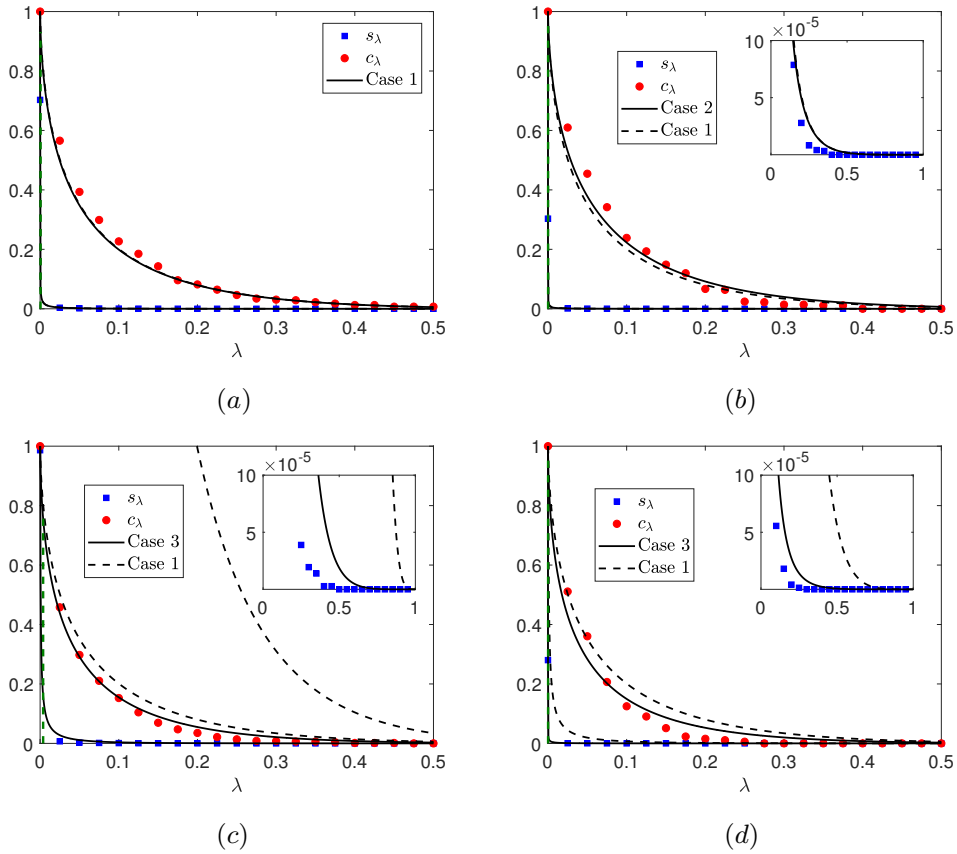


Figure 24: Qutrit Bayesian regions constructed with a $(M = 90)$-outcome POVM. (a) Case 1 ($N = 150$) and (b) Case 2 ($N = 180$) are studied with the maximally-mixed true state $\hat{\rho} = \mathbb{1}/3$. (c,d) Case 3 refers to the true pure state described by the equal superposition $(|0\rangle + |1\rangle + |2\rangle)/\sqrt{3}$ of three orthonormal kets. The 3rd case is presented with an ML estimator of (c) rank-1 ($N = 30$) and that for (d) rank-2 ($N = 90$). All insets blow up the scale for s_λ. Panels (c) and (d) show that the (overestimated) size approximations still fare much better than the optimistic expressions in (4.11). Improvements on s_λ estimates with asymptotic truncations become more conspicuous especially when (c) logarithmic divergence dominates in the low-N regime, in which truncations can reduce a significant amount of Gaussian-approximation artifacts.

off-diagonal parameters rapidly becomes an inefficient and obsolete option for generating adequate parameter samples.

The qubit system possesses a dimension D small enough such that the average error $\mathrm{E}[|\vec{r} - \vec{r}_{\mathrm{ML}}|]$ is small and the Gaussian approximations in (4.7) and (4.17) are valid even when N is not very large. Quantum systems of larger D, starting with the

qutrit, generally requires a correspondingly larger N to achieve similar tomographic precisions [107, 108]. For very large N values, the likelihood function becomes extremely narrow since its curvature is asymptotically governed by $F_{\mathrm{ML}} \sim N$. As a result, the size s_λ is tricky to calculate numerically with sophisticated Monte Carlo methods [105, 106]. For the purpose of demonstrating the performance of our results, we may slightly circumvent this problem by considering an overcomplete POVM ($M > 9$) while maintaining a reasonable N value, which similarly reduces the average error [107] for the Gaussian approximations to hold.

Figure 24 showcases qutrit tomography for all the various cases discussed in the previous section. For qutrits, the size corrections are generally overestimates because of the complicated $\partial\mathcal{M}_3$.

4.3.6 Remarks

In this section, we provided an asymptotic theory of Bayesian error regions for general convex parameter spaces that cover a wide range of applications in quantum information, including quantum-state tomography, whenever a uniform prior is used to describe the unknown true parameter. This allows one to perform asymptotic error certification for uniform priors that circumvents NP-hard Monte Carlo computations. The theory provides asymptotically analytical formulas for the region size and credibility in cases where the ML estimator is an interior point [Eq. (4.11), (4.14), and (4.15)], as well as the case where the ML estimator is on the boundary of the parameter space [Eq. (4.18)]. These expressions approach the exact answers whenever the joint boundary of both the region and full parameter space is smooth such as quantum-state tomography of qubits. Otherwise they generally give conservative overestimates for the region size as this is related to the way region truncations are handled by the theory. When applied to examples in quantum-state to-

mography, these asymptotic expressions give extremely accurate estimates in spite of the sophisticated state-space boundaries. The theoretical framework presented here can in principle be generalized to any other prior so long as analytical integrals for Gaussian likelihoods and the volume of the parameter space are known for that prior. This, however, has to be done on a case-by-case basis at the moment. In high-dimensional quantum-state tomography, as shown by the qutrit case, the assumption of smooth-boundary begins to fail, so that we need to construct an efficient numerical method. In the following section, we shall propose an effective numerical method that works for the error certification of high-dimensional parameter estimation.

4.4 Efficient Monte-Carlo Method

In this section, we provide an efficient numerical method to calculate the size and credibility of the credible region $\mathcal{R}$. A traditional numerical method to calculate the region is to sample points in the parameter space and filter the points which are outside of the region of interest, which we call Monte-Carlo (MC) filtering. This kinds of MC sampling becomes highly inefficient as the dimension of the parameter space gets larger. If the dimension of the parameter space is large, the ratio of the size of the region of interest to the entire parameter space becomes exponentially small so that the required number of samples to obtain a reasonable number of points inside of the region of interest is exponentially large in terms of the dimension.

4.4.1 In-region sampling

In this section, we introduce a numerical method to compute both size s_λ and credibility c_λ without MC filtering. The intuition behind our method is to realize that if one inspects the average of some quantity q_λ over the region $\mathcal{R}_\lambda$ through λ, then the ratio of change with λ bears information about the behavior of the size s_λ with λ. It is formally described by the following lemma:

Lemma 1. *(Region-average computation (RAC) lemma) For any prior $(d\vec{r}')$ and the number of copies N, the size s_λ (up to a multiplicative factor), and hence the credibility c_λ, are all inferable from the $\mathcal{R}_\lambda$-average quantity $u_\lambda = \overline{\ln L(\mathbb{D}|\vec{r}') - \ln(\lambda L_{max})}^{\mathcal{R}_\lambda}$.*

We prove this lemma by taking derivative of $u_\lambda S_\lambda$ in λ. Noting that $\partial S_\lambda/\partial\lambda = \int_{\mathcal{R}_0}(d\vec{r}')\delta(L-\lambda L_{\max})$, we obtain the following differential equation

$$\frac{\partial y_\lambda}{\partial\lambda} = -\frac{y_\lambda}{\lambda u_\lambda}. \tag{4.29}$$

The differential equation characterizes the full evolution of y_λ with a boundary condition $s_{\lambda=0} = 1$. Finally, the equation can be easily solved by iterating $y_{\lambda_{j+1}} = y_{\lambda_j}/(\lambda_j u_{\lambda_j})$ following Euler's method [109], so that we obtain c_λ using Eq. (4.3).

Now, our task of computing the size and credibility is to generate a sufficient number of random samples to calculate the $\mathcal{R}_\lambda$-average quantity u_λ. There exist many MC schemes to compute u_λ, many of which use Markov-chain algorithms [105, 110]. *Hit-and-run* sampling [111, 112, 113, 114] is one such extensively-studied scheme. Hit-and-run algorithm begins with a construction of a convex set $\mathcal{B} \supseteq \mathcal{R}_\lambda$. For large $N \gg 1$ and some λ, two general cases exist as shown in Fig. 25. In case A, we define $\mathcal{B}$ as the hyperellipsoid $\mathcal{E}_\lambda$ centered at $\vec{r}_c = \vec{r}_{\rm ML}$ that profiles the Gaussian L when $\vec{r}_{\rm ML}$ is an interior point. In case B, where $\vec{r}_{\rm ML}$ is on the

boundary of the space $\mathcal{R}_0$, we set $\mathcal{B}$ as the truncated hyperellipsoid $\mathcal{E}'_\lambda$ centered at $\vec{r}_c = \vec{r}_{\rm ML} + F_{\rm ML}^{-1}\vec{g}_{\rm ML}$, where $F_{\rm ML}$ is the Fisher information matrix evaluated at $\vec{r}_{\rm ML}$ and $\vec{g}_{\rm ML} = \partial \ln L/\partial \vec{r}'|_{\vec{r}'=\vec{r}_{\rm ML}}$. Note that there exists an intermediate case between case A and B as illustrated in the previous section. Although the hyperellipsoid approximates the credible region well and contains the region for large N, one may choose a more conservative hyperellipsoid that is bigger than described above to be more accurate, which does not change the time-duration of the algorithm much. Next, we pick a point in the convex set $\mathcal{B}$, which can be straightforwardly done by picking the ML estimator $\vec{r}_{\rm ML}$. We then draw a random line segment passing the selected point in the convex body and pick a random point on the segment. We keep the point if it is in the region $\mathcal{R}_\lambda$, the condition of which is that the likelihood function at the point is larger than λ and it corresponds to a physical quantity (for example, if it corresponds to a positive semi-definite density matrix in the case of

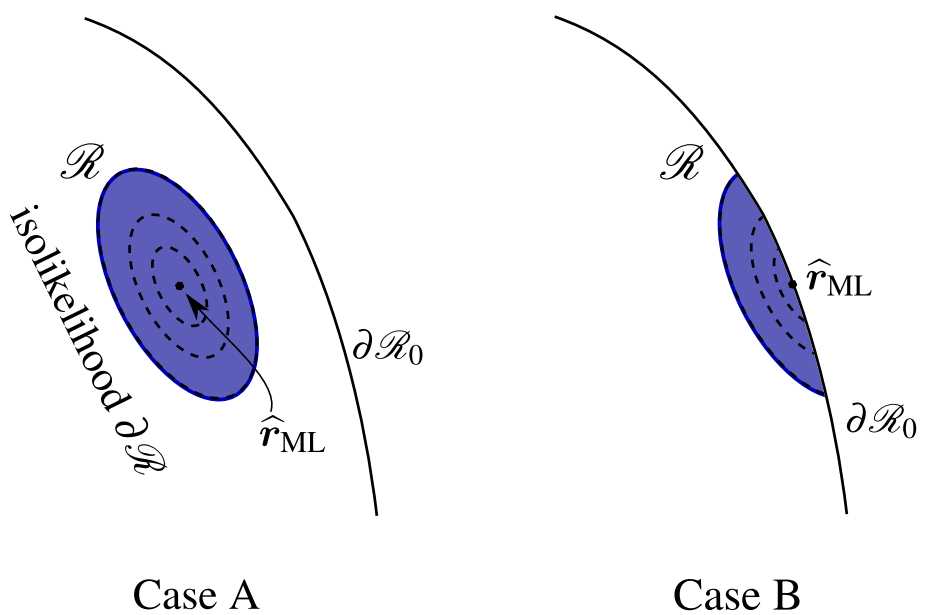


Figure 25: Typical two cases arise in quantum state tomography. Case A represents the Bayesian region inside of the parameter space, while Case B that truncated by the boundary of the parameter space.

quantum state tomography). If it is not in the region, we generate a random point on the segment between $\vec{r}_{\mathrm{ML}}$ and the previous point, which is repeated until the point lies on $\mathcal{R}_\lambda$. After picking a point in $\mathcal{R}_\lambda$, we keep this point and repeat the routine with choosing the previously selected point as a new starting point until we obtain a sufficient number of samples. Here, a line segment of a random direction can be obtained by generating a random vector v, elements of which are generated by a normal distribution. Also, when selecting a random point on the line segment, the probability distribution of a random point is set to be equal to the truncated marginal distribution.

4.4.2 Region capacity

Inspired by the in-region sampling method, we propose a concept of region capacity, which is an average distance of the region from the ML estimator $\vec{r}_{\mathrm{ML}}$; it measures how close is points in $\mathcal{R}$ from the ML estimator on average. Formally, the $\mathcal{R}$-average is defined as

$$S_{\mathcal{D},\lambda} \equiv \overline{\mathcal{D}(\vec{r}',\vec{r}_{\mathrm{ML}})}^{\mathcal{R}_\lambda} = \int_{\mathcal{R}_\lambda} (d\vec{r}')\, \mathcal{D}(\vec{r}',\vec{r}_{\mathrm{ML}}) \Big/ \int_{\mathcal{R}_\lambda} (d\vec{r}') \tag{4.30}$$

where $\mathcal{D}$ is any distance function that one can choose.

One can argue that if the metric is an l_p-norm of $p > 0$, $S_{\mathcal{D},\lambda}$ monotonically decreases with λ when $N \gg 1$ for an appropriate $(d\vec{r}')$. To see this we begin with $\mathcal{D} \equiv \mathcal{D}_p(\vec{r}',\vec{r}_{\mathrm{ML}}) = \left(\sum_j |r'_j - \vec{r}_{\mathrm{ML},j}|^p\right)^{1/p}$. According to Fig. 26, after the substitu-

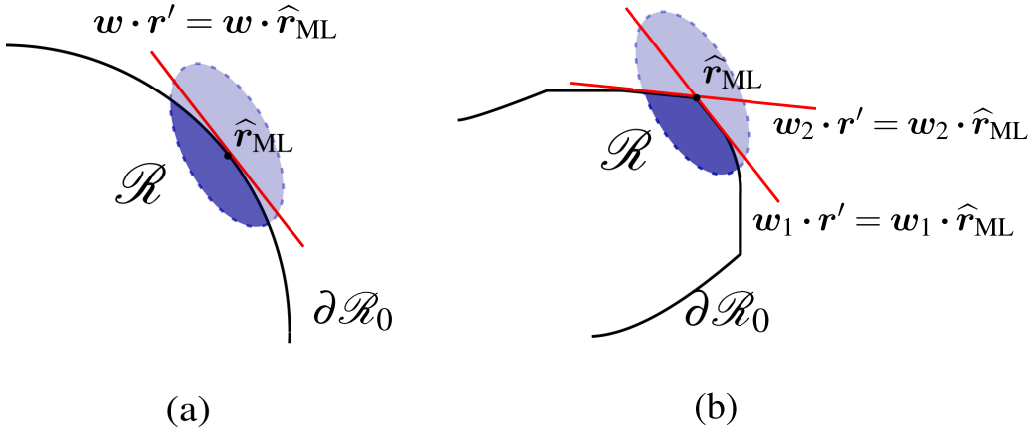


Figure 26: Barring unforeseen pathological examples, we shall assume that the $\mathscr{R}_0$ for any physical system possesses a boundary $\partial\mathscr{R}_0$ that is either (a) a smooth surface, or (b) has corners and edges. For the latter, a corner at which an ML estimator might reside can be well approximated by multiple hyperplanes if $N \gg 1$.

tion $\vec{r}'' = \vec{r}' - \vec{r}_{\rm ML}$, we have for the more complicated Case B,

$$S_{\mathcal{D}_p,\lambda} \to \frac{\int (d\vec{r}'')\, \mathcal{D}_p\, \eta(1 - \vec{r}''^{\rm T} F_{\rm ML} \vec{r}''/(-2\ln\lambda)) \prod_j \eta(\vec{w}_j^{\rm T}\vec{r}'')}{\int (d\vec{r}'') \eta(1 - \vec{r}''^{\rm T} F_{\rm ML} \vec{r}''/(-2\ln\lambda)) \prod_j \eta(\vec{w}_j^{\rm T}\vec{r}'')}$$
$$\sim \sqrt{-2\ln\lambda} \quad \left[\text{if } (d\alpha\vec{r}'') = g(\alpha)\,(d\vec{r}'')\right]. \tag{4.31}$$

The same conclusion for Case A follows by definition, and remains unchanged also for $\mathcal{D}_p(\vec{r}', \vec{r}_{\rm ML}) = \sum_j |r'_j - \vec{r}_{{\rm ML},j}|^p$ since $S_{\mathcal{D}_p,\lambda} \sim (-\ln\lambda)^{p/2}$ is also monotonic in λ. These imply that $S_{\mathcal{D}_p,\lambda}$ induced by any l_p-norm behaves as a proper capacity measure in the limit $N \gg 1$ under a sufficient class of priors that includes the uniform primitive prior. The new practice for Bayesian CR certification is then to report the three-tuple $\left(\vec{r}_{\rm ML}, C_{\lambda_0}(= 0.95 \text{ say}), S_{\mathcal{D}_p,\lambda_0}\right)$ for some $p > 0$.

4.4.2.1 Analytical error certification with region capacity

It turns out that the approximated extensions of all $\int_{\mathcal{R}_0}$ integrals to the whole $\vec{r}'$ space free *all* $\mathcal{R}$-average quantities from any geometrical dependence on $\mathcal{R}_0$, unlike S_λ that asymptotically depends on $\mathcal{R}_0$'s volume [90]. We may then use this observation to acquire asymptotic formulas for $S_{\mathcal{D}_p,\lambda}$ and u_λ to perform approximate analytical error certifications. To this end, we regard $S_2 \equiv S_{\mathcal{D}_2}$ induced by the squared l_2-norm ($p = 2$), $\mathcal{D} \equiv |\vec{r}' - \vec{r}_{\mathrm{ML}}|^2$, as the prototypical metric-induced capacity measure for $\mathcal{R}_\lambda$. Let us first discuss the case in which $\vec{r}_{\mathrm{ML}}$ is an interior point of $\mathcal{R}_\lambda$ (Case A). Since $\mathcal{R}_\lambda = \mathcal{E}_\lambda$, finding S_2 becomes the business of doing a hyperellipsoidal average of $\mathcal{D}$. This gets us to

$$S_{2\mathrm{A},\lambda} = \mathrm{Tr}[F_{\mathrm{ML}}^{-1}]\frac{(-\ln\lambda)}{d/2+1}, \quad u_{\mathrm{A},\lambda} = -\frac{2}{d+2}\ln\lambda. \tag{4.32}$$

The logarithmic divergences in λ, a derivation byproduct from Gaussian approximation of L and relaxation of $\partial\mathcal{R}_0$, pose no ill consequence so long as N is sufficiently large such that $\mathcal{R}_\lambda \subset \mathcal{R}_0$ for all λ values that give desirably large $C_\lambda < 1$.

The situation becomes more complicated for Case B, which demands geometrical knowledge about $\partial\mathcal{R}_0$ for an exact calculation of S_2 (see Fig. 27). This tempts us to use a first-order approximation by expanding the likelihood L about $\vec{r}_{\mathrm{ML}}$ to a Gaussian function of hyperellipsoidal-$\mathcal{E}'_\lambda$ profile centered at $\vec{r}_{\mathrm{c}}$, and next introducing a hyperplane containing $\vec{r}_{\mathrm{ML}}$ that is tangent to its isoGaussian (constant-Gaussian-value) contour. S_2 is then a hyperellipsoidal-cap (formed by the hyperplane and the hyperellipsoid from the Gaussian expansion of L) average. All related technical

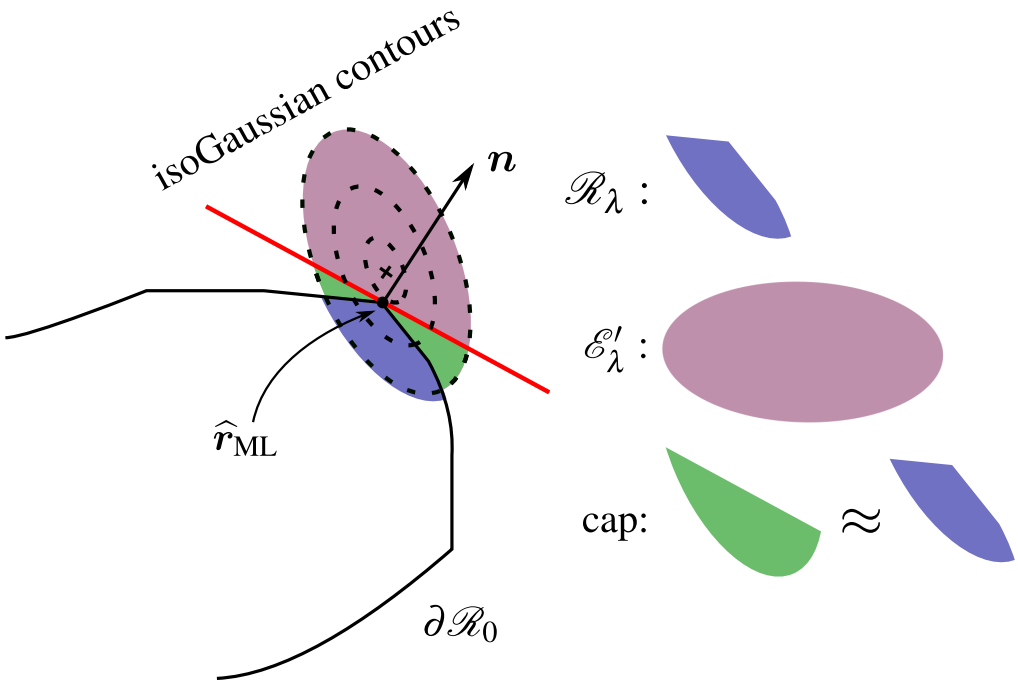


Figure 27: After expanding the likelihood L about $\vec{r}_{\rm ML}$ to a Gaussian function centered at $\vec{r}_{\rm c}$ (cross) with its own isoGaussian contours, a hyperplane (red solid line) is introduced in a manner that its normal $\vec{n}$ is orthogonal to the isoGaussian curve at $\vec{r}_{\rm ML}$ to form a cap.

calculations are provided in Appendix, and we simply state the final formulas:

$$\begin{aligned} S_{\rm 2B,\lambda} &= 2\,{\rm Tr}[\vec{M}]/\mathcal{N}_{d,l,1}\,, \\ u_{\rm B,\lambda} &= [-\ln\lambda' + {\rm Tr}[\vec{g}_{\rm ML}\vec{m}^{\rm T} - F_{\rm ML}\vec{M}]/\mathcal{N}_{d,l,1}] \\ &\quad \times \ln(\lambda L_{\rm max})/\ln(\lambda' L_{\rm max})\,, \end{aligned} \tag{4.33}$$

involving $V_d = \pi^{d/2}/(d/2)!$, $l = \sqrt{\ln(\lambda/\lambda')/(-\ln\lambda')}$, $\mathcal{N}_{d,l,x} = V_d\,{\rm I}_{(1-l)/2}((d+x)/2,(d+x)/2)$ depending on the incomplete beta function ${\rm I}_\cdot(\cdot,\cdot)$, and

$$\begin{aligned} \vec{m} &= \left[-\frac{V_{d-1}}{l(d+1)}\left(1-l^2\right)^{(d+1)/2} + \mathcal{N}_{d,l,1}\right] F_{\rm ML}^{-1}\vec{g}_{\rm ML}\,, \\ \vec{M} &= \frac{-\ln\lambda'}{d+2}\,\mathcal{N}_{d,l,3}\,F_{\rm ML}^{-1} + \frac{1}{2}\,\vec{m}\,\vec{g}_{\rm ML}^{\rm T}F_{\rm ML}^{-1}\,. \end{aligned} \tag{4.34}$$

It is easy to see that Eqs. (4.33) and (4.34) include Case A by recognizing that the "effective λ" (λ') approaches λ ($\vec{g}_{\rm ML}=\vec{0}$), so that $l\to 0$ gives $\mathcal{N}_{d,0,x}=V_d$ and $\vec{M}=(-\ln\lambda)\,F_{\rm ML}^{-1}/(d+2)$.

4.4.2.2 Discussions for quantum-state tomography

All results presented thus far apply to arbitrary physical systems. Here, we specifically investigate quantum-state tomography, thereby endowing explicit forms to all important quantities that are pertinent to Bayesian CR error certification.

For an unknown quantum state $\hat{\rho}$ of Hilbert-space dimension D, every data-copy measurement in a tomography experiment is usually mutually independent, so that the log-likelihood $\ln L=\sum_{j=1}^{M} n_j \ln p_j$ catalogs the relative frequency data $\sum_{j=1}^{M} n_j = N$ of all M measurement outcomes $\hat{\Pi}_j\geq 0$ ($\sum_j \hat{\Pi}_j=\mathbb{1}$), each with the Born probability $p_j=\mathrm{tr}[\hat{\rho}\hat{\Pi}_j]$. We can express $\hat{\rho}$ and $\hat{\Pi}_j$ in terms of the Hermitian basis $\{\mathbb{1}/\sqrt{D},\hat{\Omega}_j\}_{j=1}^{D^2-1}$ such that $\mathrm{tr}[\hat{\Omega}_j]=0$ and $\mathrm{tr}[\hat{\Omega}_j\hat{\Omega}_k]=\delta_{j,k}$, so that we may denote the ($d=D^2-1$)-dimensional $\vec{r}=\mathrm{tr}[\hat{\rho}\,\vec{\Omega}]$ and $\vec{q}_j=\mathrm{tr}[\hat{\Pi}_j\,\vec{\Omega}]$. This leads to $F_{\rm ML}=N\sum_{j=1}^{M}\vec{q}_j\vec{q}_j^{\rm T}/p_{{\rm ML},j}$ ($N\gg 1$) and $\vec{g}_{\rm ML}=\sum_{j=1}^{M} n_j\vec{q}_j/p_{{\rm ML},j}$ for the ML state estimator $\hat{\rho}_{\rm ML}$ of ML probabilities $p_{{\rm ML},j}=\mathrm{tr}[\hat{\rho}_{\rm ML}\hat{\Pi}_j]$. In concrete terms, for Case A, $\hat{\rho}_{\rm ML}$ is full rank, such that the CR $\mathcal{R}_{\vec{\lambda}}\approx\mathcal{E}_{\vec{\lambda}}$; whereas for Case B, $\rho_{\rm ML}$ is rank-deficient and $\mathcal{R}_{\vec{\lambda}}\approx\mathcal{R}_0\cap\mathcal{E}'_{\vec{\lambda}}$ is therefore approximately a truncated $\mathcal{E}'_{\vec{\lambda}}$ (covariance profile of the Gaussian expansion of L about $\vec{r}_{\rm ML}$) by the quantum-state space $\mathcal{R}_0$—the convex set of unit-trace positive operators. The uniform $(d\vec{r}')$ is assumed.

To compare with the closed-form approximations in Eqs. (4.32) and (4.33), we pick the l_2-norm to measure the region capacity $S_{\rm HS}\equiv S_2$ of $\mathcal{R}$, which is equivalent to the Hilbert-Schmidt (HS) distance for quantum states. We emphasize that for sufficiently large N, all arguments leading to the monotonicity of $S_{\mathcal{D},\vec{\lambda}}$ still applies for

Case B as $\vec{g}_{\mathrm{ML}} \to \vec{0}$. Figures 28 and 29 showcase our in-region sampling theory. The matches in both Case A and B between theory and hit-and-run sampling are very good for moderate D, but are expected to have some discrepancies for more complex systems due to the more pronounced corners in $\partial\mathcal{R}_0$ [98]. Instead, accelerated hit-and-run can be used [114].

4.4.3 Numerical Complexity estimation of hit-and-run algorithm

After suppressing dependences on logarithmic factors and error parameters, it was argued that the number of hit-and-run steps needed to gather enough sample points and form an ensemble described by $p(\vec{r})$ in hit-and-run is $O\left(d^2 R_{\mathrm{out}}^2/R_{\mathrm{in}}^2\right) = O\left(D^4 R_{\mathrm{out}}^2/R_{\mathrm{in}}^2\right)$ [115, 116] in the limit $D \gg 2$, where R_{out} is the radius of the smallest outer sphere that contains $\mathcal{R}_\lambda$ and R_{in} is that of the largest inner sphere that can be inscribed in $\mathcal{R}_\lambda$. Together with the floating-point-operations complexity $O(D^3)$ in

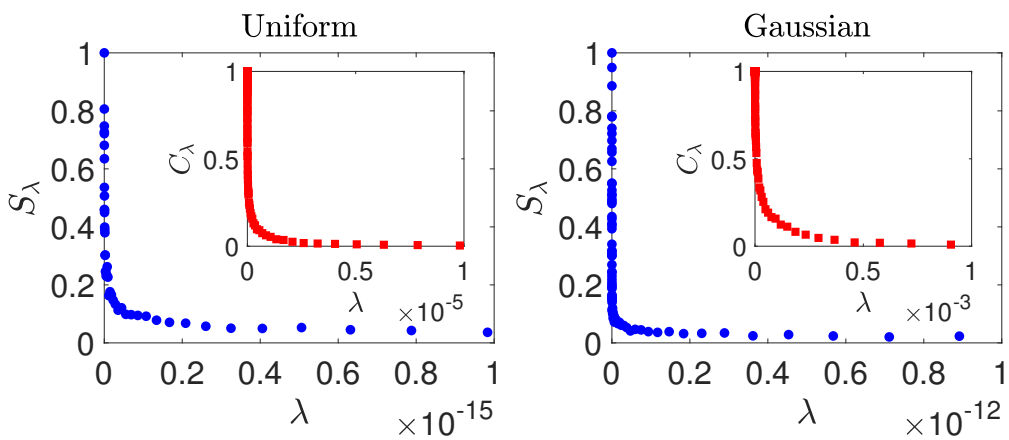


Figure 28: Plots for S_λ and C_λ generated from the in-region sampling technique on three-qubit systems ($D = 8$), with a rank-1 $\vec{r}_{\mathrm{ML}}$, $M = 512$ square-root measurement outcomes and $N/M = 5000$. The rapidly decreasing S_λ is a signature of typically small regions of such datasets, which cannot be handled with MC-filtering. The results for C_λ obtained from the sampled u_λ generated with 200 recursive steps of Euler's method to solve Eq. (4.29) for S_λ. The flexibility of in-region sampling is demonstrated by presenting graphs sampled according to both the uniform and Gaussian distributions.

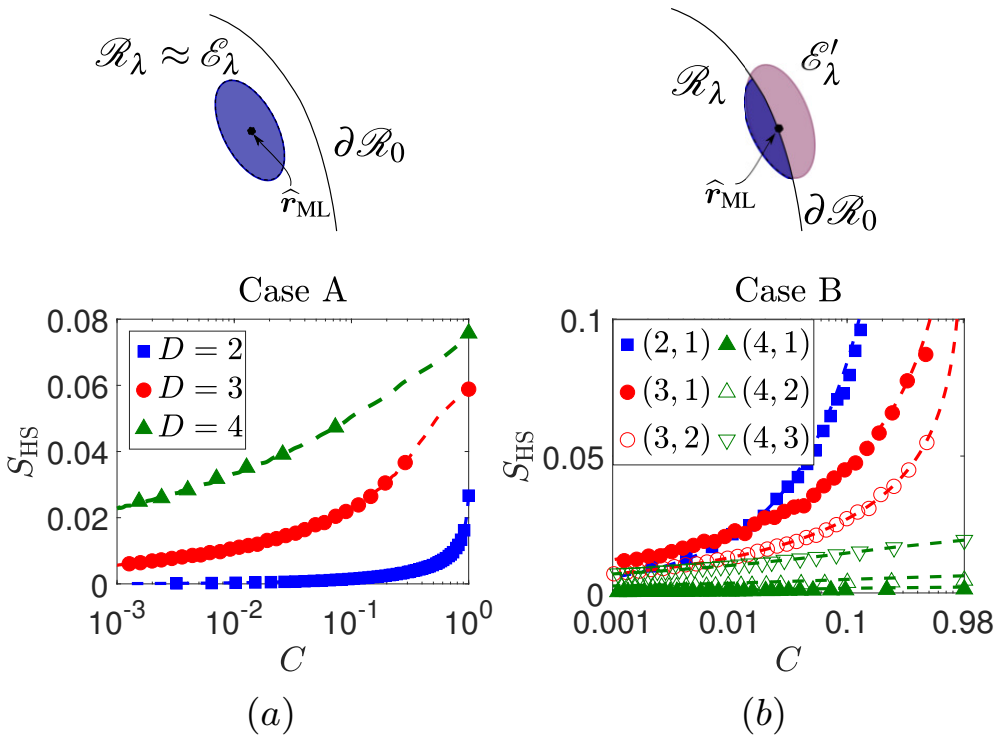


Figure 29: The (magnified) per-$\mathbb{D}$ graphs of $S_{\rm HS}$ versus C for (a) Case A and (b) Case B for various D, with $M = D^3$ random outcomes and $N/M = 500$. The two-tuples in the legend of (b) represent $(D, \text{rank}\{\vec{r}_{\rm ML}\})$. The respective dashed curves passing through the markers are calculated using Eqs. (4.32) and (4.33). The magnification factors (top to bottom, left to right in legend) for Case A are 10, 50 and 150, and those for Case B are 100, 200, 150, 10, 20 and 50.

a typical Cholesky decomposition algorithm [117], we have an estimate for the complexity $\text{cmpl} = O(D^7 R_{\rm out}^2/R_{\rm in}^2)$ for the entire hit-and-run scheme.

The treatment of Case A is straightforward as we have the complete information about $\mathscr{R}_\lambda \approx \mathscr{E}_\lambda$ in the large-N limit. If we denote $\sigma_>$ and $\sigma_<$ to respectively be the largest and smallest eigenvalue of $\widetilde{F_{\rm ML}}^{-1/2}$, then the corresponding outer and inner radii are $R_{\rm out} = \sigma_>$ and $R_{\rm in} = \sigma_<$ [see Fig. 30(a)], so that $\text{cmpl}_{\rm A} = O\Big(D^7 \text{cond}\Big(F_{\rm ML}^{-1}\Big)\Big)$ involves the conditional number $\text{cond}\Big(F_{\rm ML}^{-1}\Big) = \sigma_>^2/\sigma_<^2$.

The analysis for Case B requires extra care given the complicated state-space boundary $\partial\mathscr{R}_0$. While complete and precise details of $\mathscr{R}_0$ are absent so far, from [98], we

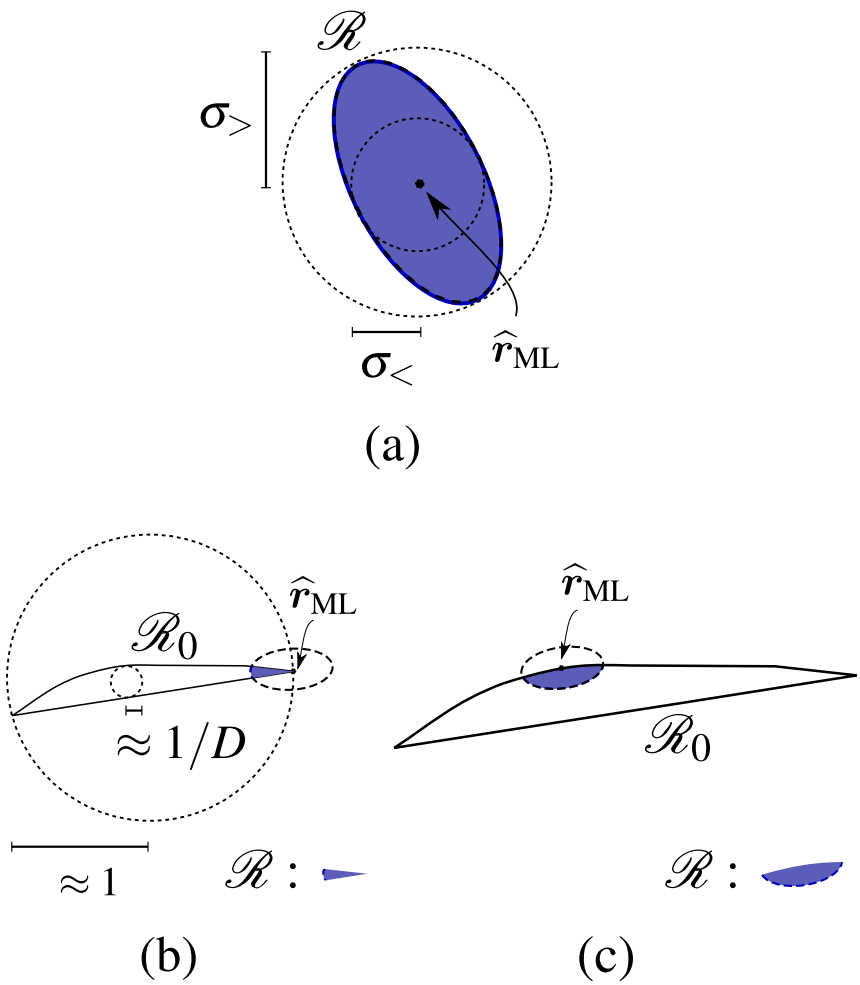
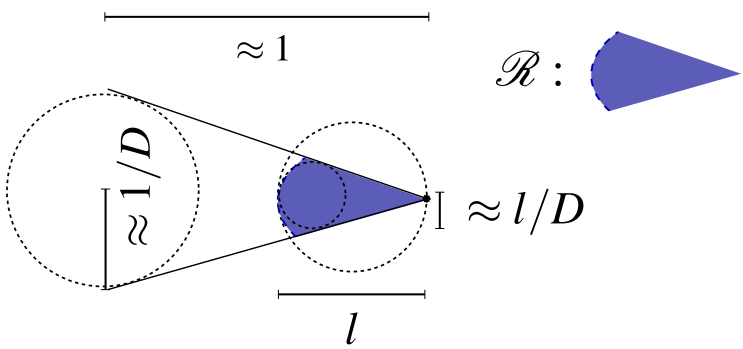


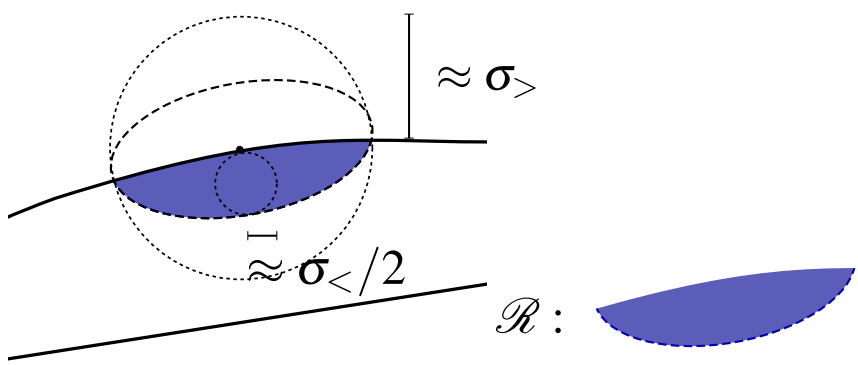
Figure 30: Schematic diagrams for the geometrical relationship between the CR $\mathscr{R}$ and the quantum state space $\mathscr{R}_0$. The situation for (a) Case A is completely known and so complexity estimation for hit-and-run is a simple matter. To acquire conservative complexity estimates for Case B, two special types of such CR may exist: either the CR (b) lies on an extremely sharp corner of $\mathscr{R}_0$ in at least one of its dimension (Type I) in whichever orientation, or (c) on one of its edges that is almost flat (Type II) in all its dimensions, with the longest $\mathscr{E}'_\lambda$-axis oriented along the flat surface.

know that in the Euclidean space, the largest inner sphere inscribable in $\mathscr{R}_0$ has a radius that approaches $1/D$ for $D \gg 2$, and that the smallest outer sphere that contains $\mathscr{R}_0$ has a radius going to 1 in the same dimension limit. The overall shape of $\mathscr{R}_0$ is therefore a "squashed" convex body for large D, such that *at least one* of its dimensions drops appreciably to zero. To estimate the complexity for Case B, we

consider CRs of two tractable types: a Type I CR is located at an extremely sharp corner of $\mathcal{R}_0$ that is made from at least one of its rapidly shrinking dimensions, as shown in Fig. 30(b), whereas a Type II CR is situated at an extremely flat boundary of $\mathcal{R}_0$ where all of its dimensions remain approximately constant within the CR as in Fig. 30(c). For a conservative estimate of cmpl, we consider an $\mathcal{R}$ such that



(a)



(b)

Figure 31: Schematic diagrams for (a) Type I and (b) Type II Bayesian regions. Type I regions have complexities that are strongly influenced by the cornered geometry (greatly exaggerated for visual aid), whereas Type II regions have complexities that strongly depends on the eigenvalue aspect ratios of F_{ML}. All other intermediate CR types give rise to complexities affected by the geometries of both $\partial\mathcal{R}_0$ and F_{ML}.

the longest axis of $\mathcal{E}'_\lambda$ is aligned with the flat surface. All other types of Case-B CRs may be viewed as intermediate situations of these two and have no analytical complexity estimates known to us. The data-copy number $N \gg 1$ is assumed to be sufficiently large such that $\vec{g}_{\mathrm{ML}} \approx \vec{0}$ and $\vec{r}_{\mathrm{c}} \approx \vec{r}_{\mathrm{ML}}$.

To estimate cmpl for a Type I CR, we assume that the corner is extremely sharp in one particular dimension such that the curvature of $\partial\mathcal{R}_0$ extending out from $\vec{r}_{\mathrm{ML}}$ is almost flat. Then following Fig. 31(a), the concept of similar figures give $R_{\mathrm{out}}/R_{\mathrm{in}} \approx D$, which is independent of F_{ML} for extremely sharp corners, and $\mathrm{cmpl}_{\mathrm{B,I}} = O\big(D^9\big)$. The complexity for Type II CRs may be estimated with the help of Fig. 31(b), where $R_{\mathrm{out}}/R_{\mathrm{in}} \approx 2\,\mathrm{cond}\Big(F_{\mathrm{ML}}^{-1/2}\Big)$ is now independent of $\partial\mathcal{R}_0$ due to its extremely mild edge features, leading us to $\mathrm{cmpl}_{\mathrm{B,II}} = O\Big(D^7\,\mathrm{cond}\Big(F_{\mathrm{ML}}^{-1}\Big)\Big) = \mathrm{cmpl}_{\mathrm{A}}$.

4.4.4 Remarks

In realistic multi-dimensional parameter estimation problems, sufficiently large dataset almost exclusively results in extremely small Bayesian credible regions relative to the entire parameter space. The conventional practice of first doing Monte Carlo to sample the parameter space followed by sample filtering almost always fails to accurately construct such small error regions. Our technique of in-region sampling introduced in this dissertation is capable of constructing any such small regions efficiently with perfect yield. In-region sampling is equivalent to computing region-averages that is efficient with a wide range of numerical methods. The region-average perspective of in-region sampling allows us to operationally formulate an alternative concept of region capacity through averaging any l_p distance norm between two credible-region points, for which, in the special case $p = 2$, closed-form approximation formulas to facilitate ultrafast analytical Bayesian error estimations with sufficiently large datasets are readily available. Either way, efficient Bayesian

error certifications can now be carried out on physical systems of varying complexity. For exceedingly large quantum systems where Monte Carlo computations start to become visibly taxing, these asymptotic formulas can serve as large-scale approximate certifiers at least for high credibility values.

4.5 Conclusion

In this section, we have investigated Bayesian error region in quantum parameter estimation. We have derived asymptotic analytical expressions for size and credibility of Bayesian error regions by using Gaussian approximation and assuming a smooth boundary of the parameter space. While the analytical expressions fit very well to numerical simulation for low-dimensional system such as a qubit, it starts to exhibit a discrepancy as the dimension of quantum systems of interest gets larger. In order to make it possible to calculate the size and credibility even for high-dimensional quantum systems, we proposed an efficient numerical method by the so-called in-region sampling. The algorithm is inspired by an observation that the behavior of the average of a function defined in the parameter space of interest gives the size of Bayesian region through a differential equation, which can be solved by Euler's method.

The in-region sampling provides a different measure for closeness of Bayesian region which is called region capacity. It is defined as the average distance from the maximum likelihood estimator to points in the region. Since it is based on a metric, it is expected to be informative to assess the estimation for various applications of quantum information processing.

4.6 Appendix

Derivation of analytical expression of size and credibility in Case 1

We start with the definition of size and credibility of Eq. (4.2) and the Gaussian approximation of the likelihood function in Eq. (4.7) in Case 1 to calculate the size of the credible region. We proceed by using the well-known integral representation

$$\eta(x) = \int \frac{dt}{2\pi i} \frac{e^{ixt}}{t - i\varepsilon}\bigg|_{\varepsilon=0} \tag{4.35}$$

of the Heaviside step function and the recognition that $\eta(L(\mathbb{d}|\vec{r}) - \lambda L_{\max}) = \eta(\ln L(\mathbb{D}|\vec{r} - \ln(\lambda L_{\max}))$ to write

$$\begin{aligned} s_\lambda &= \int_{\mathcal{R}_0} (d\vec{r}')\chi_\lambda(\vec{r}') \\ &= \int \frac{dt}{2\pi i} \frac{e^{-it\ln(\lambda L_{\max})}}{t - i\varepsilon} \int_{\mathcal{R}_0} e^{it\ln L(\mathbb{D}|\vec{r}')}\bigg|_{\varepsilon=0}. \end{aligned} \tag{4.36}$$

Now, using the Gaussian approximation of the likelihood function,

$$\begin{aligned} \int_{\mathcal{R}_0} e^{it\ln L(\mathbb{D}|\vec{r}')} &\approx \frac{e^{it\ln L_{\max}}}{V_{\mathcal{R}_0}} \int \left(\prod_j dr'_j\right) e^{-\frac{it}{2}(\vec{r}' - \vec{r}_{\rm ML})\cdot F_{\rm ML}\cdot(\vec{r}' - \hat{\vec{r}}_{\rm ML})} \\ &= \frac{e^{it\ln L_{\max}}}{V_{\mathcal{R}_0}} \left(\frac{2\pi}{it}\right)^{d/2} |F_{\rm ML}|^{-1/2}. \end{aligned} \tag{4.37}$$

The integral in t can then be completed with another identity,

$$\frac{1}{a^n} = \frac{1}{(n-1)!} \int_0^\infty dy y^{n-1} e^{-ay}, \tag{4.38}$$

$$
\begin{aligned}
s_\lambda &= \frac{(2\pi)^{d/2}}{V_{\mathcal{R}_0}} |F_{\mathrm{ML}}|^{-1/2} \int \frac{dt}{2\pi i} \frac{e^{-it\ln\lambda}}{(it)^{d/2}(t-i\varepsilon)} \bigg|_{\varepsilon=0} \\
&= \frac{(2\pi)^{d/2}}{V_{\mathcal{R}_0}(d/2-1)!} |F_{\mathrm{ML}}|^{-1/2} \int_0^\infty dy y^{d/2-1} \int \frac{dt}{2\pi i} \frac{e^{-it(\ln\lambda+y)}}{t-i\varepsilon} \bigg|_{\varepsilon=0} \\
&= \frac{V_d}{V_{\mathcal{R}_0}} (-2\ln\lambda)^{d/2} |F|^{-1/2}. \qquad (4.39)
\end{aligned}
$$

The credibility may be calculated either with Eq. (4.2) or Eq. (4.3). The former can be done with a similar trick as above. The latter is more straightforward with the following identities:

$$
\int_\lambda^1 d\lambda' (-\ln\lambda')^\alpha = \alpha! - \Gamma(\alpha+1, -\ln\lambda), \qquad (4.40)
$$

$$
\Gamma(\alpha+1, y) = \alpha\Gamma(\alpha, y) + y^\alpha e^{-y} \qquad (4.41)
$$

for the upper incomplete Gamma function.

The estimation of r_P

As $\hat{\vec{r}}$ is close to $\partial\mathcal{R} \cap \partial\mathcal{R}_0$, the column $\vec{r}_P$ can be estimated by first generating a set $\left\{\vec{r}_j^{(\mathrm{bd})}\right\}_{j=1}^L$ of L boundary parameter columns, which can be done by generating many random d-dimensional columns $\vec{\varepsilon}_j$ of small magnitudes and defining $\vec{r}_j^{(\mathrm{bd})} = \mathcal{M}(\hat{\vec{r}} + \vec{\varepsilon}_j)$, where $\mathcal{M}$ is a map that brings any column that lies outside of $\mathcal{R}_0$ to $\partial\mathcal{R}_0$ (the probability of generating a random boundary point without the action of $\mathcal{M}$ is effectively zero). Then $\vec{r}_P$ may be taken to be the boundary point that gives the maximum likelihood value $L_{\max}^{(\partial\mathcal{R}_0)}$.

As an example, we suppose that in state tomography, $\hat{\vec{r}}$ is the $(d = D^2 - 1)$-dimensional real parameter column that uniquely represents the D-dimensional ML quantum state $\hat{\rho}$ that lies close to $\partial\mathcal{R} \cap \partial\mathcal{R}_0$. Then a set of random columns $\vec{\varepsilon}_j$, distributed according to the standard Gaussian distribution for instance, is added to $\hat{\vec{r}}$ one

at a time and the resulting columns $\hat{\vec{r}} + \vec{\varepsilon}_j \to H_j$ are transformed into the corresponding Hermitian operators $H_j = H_j^\dagger$. We discard those H_js that are full-rank positive operators and move on to others that are nonpositive, and apply the map $\mathcal{M}(\cdot) = \mathcal{N}[\,\cdot + \sigma_{\min}(\cdot)1]$ to H_j, which adds a multiple of the identity equal to the minimum eigenvalue $\sigma_{\min}$ and trace-normalize the resulting operator. This turns the nonpositive H_js into boundary states $\rho_j^{(\mathrm{bd})} \to \vec{r}_j^{\,(\mathrm{bd})}$ that is near $\hat{\vec{r}}$ if $\vec{\varepsilon}_j$ is small enough.

The derivation of (4.15)

With the Gaussian likelihood in (4.7) centered at $\vec{r}_{\mathrm{ML}}$, let us denote the full hyperellipsoid defined by the isolikelihood contour at some value of λ as $\mathcal{E}_\lambda$. If $\mathcal{R} = \mathcal{R}_\lambda$ is truncated, then the region $\widetilde{\mathcal{R}}_\lambda \supseteq \mathcal{R}_\lambda$ that is bounded $\partial\mathcal{E}_\lambda \cap \partial P$ is an overestimate of $\mathcal{R}_\lambda$. The task here is to calculate the volume $V_{\widetilde{\mathcal{R}}_\lambda}$ of this region.

The hyperellipsoidal surface $\partial\mathcal{E}_\lambda$ for any λ is described by the equation

$$(\vec{r} - \vec{r}_{\mathrm{ML}}) \cdot F'_{\mathrm{ML}} \cdot (\vec{r} - \vec{r}_{\mathrm{ML}}) = 1 \tag{4.42}$$

with $F'_{\mathrm{ML}} = F_{\mathrm{ML}} / (-2\ln\lambda)$, or in terms of its more convenient diagonal-basis representation found with the spectral decomposition $F'_{\mathrm{ML}} = \vec{O} D \vec{O}^{\mathrm{T}}$,

$$(\vec{r}' - \vec{r}'_{\mathrm{ML}}) \cdot D \cdot (\vec{r}' - \vec{r}'_{\mathrm{ML}}) = 1\,, \tag{4.43}$$

where $\vec{a}' = \vec{O}^{\mathrm{T}} \cdot \vec{a}$, where the diagonal entries D_j of D are reciprocals of squares of the λ-hyperellipsoidal axes lengths. In the primed coordinates, the hyperplane P, which contains $\vec{r}'_P$, the ML estimator over $\partial\mathcal{R}_\emptyset$, and the normal $\vec{n}' \propto D \cdot (\vec{r}'_P - \vec{r}'_{\mathrm{ML}})$, satisfies the equation $\vec{n}' \cdot \vec{r}' = \vec{n}' \cdot \vec{r}'_P$. One easy trick to calculate $V_{\widetilde{\mathcal{R}}_\lambda}$ would then be

to first start with the integral definition

$$V_{\widetilde{\mathcal{R}}_\lambda} = V_{\mathcal{R}_0} \int (d\vec{r}')\, \eta\left(1 - (\vec{r}' - \vec{r}'_{\mathrm{ML}}) \cdot D \cdot (\vec{r}' - \vec{r}'_{\mathrm{ML}})\right) \eta\left(\vec{n}' \cdot (\vec{r}'_P - \vec{r}')\right), \qquad (4.44)$$

and next perform the change of variables $\vec{r}' \to \vec{r}'' = D^{1/2} \cdot (\vec{r}' - \vec{r}'_{\mathrm{ML}})$ to express this same volume

$$V_{\widetilde{\mathcal{R}}_\lambda} = \frac{V_{\mathcal{R}_0}}{\sqrt{|D|}} \int_{S_{d-1}} (d\vec{r}'')\, \eta\left(\vec{n}' \cdot \left(\vec{r}'_P - \vec{r}'_{\mathrm{ML}}\right) - \vec{n}' \cdot D^{-1/2} \cdot \vec{r}''\right) \qquad (4.45)$$

as a multiple of the volume of intersection between a corresponding unit $(d-1)$-hypersphere S_{d-1} and a transformed hyperplane P' described by the equation $\vec{n}' \cdot D^{-1/2} \cdot \vec{r}'' = \vec{n}' \cdot (\vec{r}'_P - \vec{r}'_{\mathrm{ML}})$ in the $\vec{r}''$ reference frame.

For the primitive prior and the earlier definition of $\vec{n}'$, this intersection volume has a known analytical answer, which depends on the shortest distance

$$\begin{aligned} l &= l_0 = \frac{|\vec{n}' \cdot (\vec{r}'_P - \vec{r}'_{\mathrm{ML}})|}{||D^{-1/2} \cdot \vec{n}'||} \\ &= \sqrt{\frac{(\vec{r}_P - \vec{r}_{\mathrm{ML}}) \cdot F_{\mathrm{ML}} \cdot (\vec{r}_P - \vec{r}_{\mathrm{ML}})}{-2\ln\lambda}} = \sqrt{\frac{\ln\lambda_{\mathrm{int}}}{\ln\lambda}} \end{aligned} \qquad (4.46)$$

between the center of the hypersphere and P'. It follows that the magnitude of l_0 increases with λ. At the critical value $\lambda = \lambda_{\mathrm{int}}$, we have $l_0 = 1$, which tells us that at this critical value $\partial\mathcal{E}_{\lambda \geq \lambda_{\mathrm{int}}} \cap \partial P = \emptyset$. Beyond $\lambda > \lambda_{\mathrm{int}}$ we must have the shortest distance $l = 1$ set to unity since this would imply that $V_{\widetilde{\mathcal{R}}_0} = V_{\mathcal{E}_\lambda} = \gamma V_{d,\lambda}$. It can then be shown, for instance either refer to [118] or Appendix D of [91], that $V_{\widetilde{\mathcal{R}}_\lambda} = \gamma V_{\mathcal{E}_\lambda} = \gamma V_{d,\lambda}$, where $\gamma = 1 - I_{\frac{1-l}{2}}\left(\frac{d+1}{2}, \frac{d+1}{2}\right)$.

Approximation of various $\mathcal{D}$ measures

The HS distance measure $\mathcal{D}_{\rm HS}(\vec{r}',\vec{r}_{\rm ML})$ takes the very simple quadratic form in $\mathcal{D}_{\rm HS} = (\vec{r}' - \vec{r}_{\rm ML})^2$ under any circumstance, whereas the trace-class distance $\mathcal{D}_{\rm tr}$ has no easy functional form in terms of $\vec{r}'$ for $D > 2$. Nevertheless in the limits $N \gg 1$ *and* $D \gg 2$, based on the principles of random matrix theory, it is deduced that the asymptotic expression

$$S_{\rm tr} \approx \frac{8\sqrt{D\,S_{\rm HS}}}{3\pi} \tag{4.47}$$

relating the final $\mathcal{R}$-averages $S_{\rm HS}$ and $S_{\rm tr}$ is approximately valid for both Case A and B, which incidentally takes the same form found in [119] that was calculated for statistical-fluctuation studies.

We start with making an approximate connection between $S_{\rm tr}$ and $S_{\rm HS}$ by examining the Hermitian operator $\Delta\hat{\rho}' = \hat{\rho}' - \hat{\rho}_{\rm ML}$ ($\rho' \in \mathcal{R}$). In Case A, the distribution of $\Delta\hat{\rho}'$ in $\mathcal{R}$ has zero mean, $\overline{\Delta\hat{\rho}'}^{\mathcal{R}} = 0$. This is also approximately true for the Case B situation when N is sufficiently large such that $\mathcal{R}$ is small. Furthermore, the space of $\Delta\hat{\rho}'$ is essentially a bounded set of Hermitian random operators. Here, we shall make the assumption that each matrix entry $\Delta\hat{\rho}'_{jk}$ in the computational basis is an independent random complex number. Under this condition, the $\Delta\hat{\rho}'$s form what is now known as a *Wigner ensemble* [120, 121, 122, 123] with the second moment equal to $\overline{|\Delta\hat{\rho}'_{jk}|^2}^{\mathcal{R}} = {\rm tr}[(\Delta\rho')^2] = S_{\rm HS}$. Moreover, they are known to have an i.i.d. eigenvalue spectrum that follows the Wigner semicircle law

$$\sigma\left(\Delta\rho'/\sqrt{D}\right) \sim \frac{1}{2\pi S_{\rm HS}}\sqrt{4\,S_{\rm HS} - x^2} \qquad \text{for } -2\sqrt{S_{\rm HS}} \le x \le 2\sqrt{S_{\rm HS}} \tag{4.48}$$

in the large-D limit. The trace-class distance $\mathcal{D}_{\mathrm{tr}}$ can thus be calculated with the integral

$$\mathcal{D}_{\mathrm{tr}} \approx \frac{\sqrt{D}}{2\pi S_{\mathrm{HS}}} \int_{-2\sqrt{S_{\mathrm{HS}}}}^{2\sqrt{S_{\mathrm{HS}}}} dx\,|x|\,\sqrt{4S_{\mathrm{HS}} - x^2} = \frac{8\sqrt{D S_{\mathrm{HS}}}}{3\pi}\,, \tag{4.49}$$

so that we end up with (4.47).

For Case B, that $\overline{\Delta\hat{\rho}'}^{\mathcal{R}} = 0$ is obvious, but as we have no means of analytically estimate $\overline{\Delta\hat{\rho}'}^{\mathcal{R}}$, we make a further approximation that as long as $\mathcal{R}$ is sufficiently small, the offset to $\overline{\Delta\hat{\rho}'}^{\mathcal{R}}$ will proportionately be small, so that (4.47) remains a reasonable asymptotic approximation.

The Bures distance measure $\mathcal{D}_{\mathrm{B}}$ also has no tractable functional form in $\vec{r}'$ for general D. To find the asymptotic link with $\vec{r}'$ this time, it is technically more convenient to inspect the behavior of $\mathcal{F}$ around $\hat{\rho}_{\mathrm{ML}} \leftrightarrow \vec{r}_{\mathrm{ML}}$ as $N \gg 1$.

We will show by using a Taylor expansion about $\hat{\rho}_{\mathrm{ML}}$ that we have

$$\mathcal{F}_{\mathrm{A}} \approx 1 - \frac{1}{2}(\vec{r}' - \vec{r}_{\mathrm{ML}})^{\mathrm{T}}\, Q_D\, (\vec{r}' - \vec{r}_{\mathrm{ML}}) \tag{4.50}$$

for Case A and

$$\begin{aligned}\mathcal{F}_{\mathrm{B}} \approx{}& 1 + (\vec{r}' - \vec{r}_{\mathrm{ML}})^{\mathrm{T}} \mathrm{tr}[P_r \vec{\Omega}] \\ &+ \frac{1}{2}(\vec{r}' - \vec{r}_{\mathrm{ML}})^{\mathrm{T}} \left(\frac{1}{2}\mathrm{tr}[P_r \vec{\Omega}]\mathrm{tr}[P_r \vec{\Omega}^{\mathrm{T}}] - Q_r\right) (\vec{r}' - \vec{r}_{\mathrm{ML}})\end{aligned} \tag{4.51}$$

for Case B, where P_r is the projector onto the support of $\hat{\rho}_{\mathrm{ML}}$ having the rank-deficient spectral decomposition $\hat{\rho}_{\mathrm{ML}} = \sum_{j=1}^{r} |\lambda_j\rangle\lambda_j\langle\lambda_j|$, and

$$Q_r = \sum_{j=1}^{r}\sum_{k=1}^{r} \frac{\langle\lambda_j|\vec{\Omega}|\lambda_k\rangle\langle\lambda_k|\vec{\Omega}^{\mathrm{T}}|\lambda_j\rangle}{\lambda_j + \lambda_k}\,. \tag{4.52}$$

A Taylor expansion of $\mathcal{F}$ about $\vec{r}_{\rm ML}$, or

$$\begin{aligned}\mathcal{F} \approx 1 + (\vec{r}' - \vec{r}_{\rm ML})^{\rm T} \frac{\partial \mathcal{F}_{\rm ML}}{\partial \vec{r}_{\rm ML}} \\ + \frac{1}{2}(\vec{r}' - \vec{r}_{\rm ML})^{\rm T} \frac{\partial}{\partial \vec{r}_{\rm ML}} \frac{\partial \mathcal{F}_{\rm ML}}{\partial \vec{r}_{\rm ML}} (\vec{r}' - \vec{r}_{\rm ML}) ,\end{aligned} \tag{4.53}$$

reveals the large-N characteristics that is needed for analysis. The structure of quantum fidelity, however, demands the operator variation of $\sqrt{A}$ for a positive (semidefinite) A. An integral representation of $\sqrt{A}$ exists [124] and can be written as

$$\sqrt{A} = \lim_{\varepsilon \to 0^+} \int_0^\infty \frac{dt}{\pi\sqrt{t}} \frac{A}{t + A + \varepsilon} , \tag{4.54}$$

where the limit is understood to be applied at the very end of all calculations so that Eq. (4.54) is valid even for A with zero eigenvalues.

The first-order variation of $\mathrm{tr}[\sqrt{A}]^2$ produces

$$\begin{aligned}\delta \,\mathrm{tr}[\sqrt{A}]^2 &= 2\,\mathrm{tr}[\sqrt{A}] \lim_{\varepsilon \to 0} \int_0^\infty \frac{dt}{\pi\sqrt{t}} \,\mathrm{tr}[\delta \frac{A}{t + A + \varepsilon}] \\ &= 2\,\mathrm{tr}[\sqrt{A}] \lim_{\varepsilon \to 0} \int_0^\infty \frac{dt}{\pi\sqrt{t}} \,\mathrm{tr}\Big\{ \delta A \frac{1}{t + A + \varepsilon} \\ &\qquad - \frac{A}{t + A + \varepsilon} \delta A \frac{1}{t + A + \varepsilon} \Big\} \\ &= 2\,\mathrm{tr}[\sqrt{A}] \lim_{\varepsilon \to 0} \mathrm{tr}[\delta A \frac{A + 2\varepsilon}{2\,(A + \varepsilon)^{3/2}}] .\end{aligned} \tag{4.55}$$

In terms of $\mathcal{F}$, we substitute $A \equiv \hat{\rho}_{\rm ML}^{1/2} \hat{\rho}' \hat{\rho}_{\rm ML}^{1/2}$, and evaluate the above result with $\hat{\rho}' = \hat{\rho}_{\rm ML}$, or $A \to A_{\rm ML} = \hat{\rho}_{\rm ML}^2$, then with $\delta A_{\rm ML} = \hat{\rho}_{\rm ML}^{1/2} \delta\vec{r}_{\rm ML} \cdot \vec{\Omega} \hat{\rho}_{\rm ML}^{1/2}$,

$$\frac{\partial \mathcal{F}_{\rm ML}}{\partial \vec{r}_{\rm ML}} = 2 \lim_{\varepsilon \to 0} \mathrm{tr}[\hat{\rho}_{\rm ML} \frac{\hat{\rho}_{\rm ML}^2 + 2\varepsilon}{2\,(\hat{\rho}_{\rm ML}^2 + \varepsilon)^{3/2}} \vec{\Omega}] , \tag{4.56}$$

where we remind the Reader that tr acts on operators only, not on the vectorial

character. For Case B in which $\hat{\rho}_{\mathrm{ML}} = \sum_{j=1}^{r} |\lambda_j\rangle\lambda_j\langle\lambda_j|$ is rank-deficient, we get, after taking the trace,

$$\frac{\partial \mathcal{F}_{\mathrm{ML}}}{\partial \vec{r}_{\mathrm{ML}}} = \mathrm{tr}[P_r \vec{\Omega}]\,, \tag{4.57}$$

where $P_r = \sum_{j=1}^{r} |\lambda_j\rangle\langle\lambda_j|$. It is then trivial to realize that this first-order derivative is zero for Case A. Qualitatively, this confirms the fact that when $\vec{r}_{\mathrm{ML}}$ is an interior point, $\mathcal{F}$ has a local maximum at this point as it should, while a boundary estimator evaluates to a nonzero $\mathcal{F}$ slope.

Upon denoting $\vec{W}_{\mathrm{ML}} = \hat{\rho}_{\mathrm{ML}}^{1/2}\vec{\Omega}\hat{\rho}_{\mathrm{ML}}^{1/2}$, the second-order variation follows from the second line of (4.55):

$$\begin{aligned}\delta\frac{\partial\,\mathrm{tr}[\sqrt{A}]^2}{\partial \vec{r}'} = 2\,\delta\Bigg[\mathrm{tr}[\sqrt{A}]\lim_{\varepsilon\to 0}\int_0^\infty \frac{dt}{\pi\sqrt{t}} \\ \times \mathrm{tr}[\vec{W}_{\mathrm{ML}}\frac{t+\varepsilon}{(t+A+\varepsilon)^2}]\Bigg]\,.\end{aligned} \tag{4.58}$$

A product-rule dissociation of (4.58) comprises a $\delta\,\mathrm{tr}[\sqrt{A}]$ and

$$\begin{aligned}&\lim_{\varepsilon\to 0}\int_0^\infty \frac{dt}{\pi\sqrt{t}}\,\mathrm{tr}[\vec{W}_{\mathrm{ML}}\delta\frac{t+\varepsilon}{(t+A+\varepsilon)^2}] \\ &= -\lim_{\varepsilon\to 0}\int_0^\infty \frac{dt}{\pi\sqrt{t}}\,\mathrm{tr}\Bigg\{\vec{W}_{\mathrm{ML}}\left[\frac{t+\varepsilon}{t+A+\varepsilon}\,\delta A\,\frac{1}{(t+A+\varepsilon)^2}\right. \\ &\qquad\left. + \frac{t+\varepsilon}{(t+A+\varepsilon)^2}\,\delta A\,\frac{1}{t+A+\varepsilon}\right]\Bigg\}\,.\end{aligned} \tag{4.59}$$

After evaluating the variation at $\hat{\rho}' = \hat{\rho}_{\mathrm{ML}}$ and further undoing all integrations with the help of its spectral decomposition, Case B yields

$$\begin{aligned}\frac{\partial}{\partial \vec{r}_{\mathrm{ML}}}\frac{\partial \mathcal{F}_{\mathrm{ML}}}{\partial \vec{r}_{\mathrm{ML}}} = &\frac{1}{2}\mathrm{tr}[P_r\vec{\Omega}]\mathrm{tr}[P_r\vec{\Omega}^{\mathrm{T}}] \\ &- \sum_{j=1}^{r}\sum_{k=1}^{r}\frac{\langle\lambda_j|\vec{\Omega}|\lambda_k\rangle\langle\lambda_k|\vec{\Omega}^{\mathrm{T}}|\lambda_j\rangle}{\lambda_j+\lambda_k}\,.\end{aligned} \tag{4.60}$$

The counterpart expression for Case A is immediate, of course.

Analytical expression of region capacity for Case A

The approximated expression for region capacity reduces the necessary ingredients for large-N (or D) analytical estimations of $S_{\mathcal{D},\lambda}$ to just the scalar $\int_{\mathcal{R}_\lambda}(d\vec{r}')$, column $\int_{\mathcal{R}_\lambda}(d\vec{r}')\,\vec{\Delta}'_{\rm ML}$ and dyadic $\int_{\mathcal{R}_\lambda}(d\vec{r}')\,\vec{\Delta}'_{\rm ML}\,\vec{\Delta}'^{\rm T}_{\rm ML}$, where $\vec{\Delta}'_{\rm ML}=\vec{r}'-\vec{r}_{\rm ML}$.

When $\mathcal{R}_\lambda \approx \mathcal{E}_\lambda$, these three integrals takes on simple analytical forms. We start with

$$\int_{\mathcal{R}_\lambda}(d\vec{r}') = \int_{\mathcal{R}_0}(d\vec{r}')\,\eta(1-\vec{\Delta}'^{\rm T}_{\rm ML}F_{\rm ML}\vec{\Delta}'_{\rm ML}/(-2\log\lambda)) \tag{4.61}$$

and transform $\vec{r}'\to\vec{r}''=D^{1/2}\,O^{\rm T}\vec{\Delta}'_{\rm ML}$ to the translated diagonal coordinate variables of $F_{\rm ML}/(-2\log\lambda)=ODO^{\rm T}$, so that in the large-$N$ limit and uniform primitive prior, we may relax the boundary of $\mathcal{R}_0$ and write

$$\begin{aligned}\int_{\mathcal{R}_\lambda}(d,\vec{r}') &\to \frac{|D^{-1/2}|}{V_{\mathcal{R}_0}}\int(d\vec{r}'')_{\rm unif}\,\eta(1-\vec{r}''^2)\\ &= \frac{V_d}{V_{\mathcal{R}_0}}\,(-2\log\lambda)^{d/2}\,|F_{\rm ML}|^{-1/2},\end{aligned} \tag{4.62}$$

which is a function of the volume $V_d=\pi^{d/2}/(d/2)!$ of the d-dimensional unit hyperball, the inverse of $F_{\rm ML}$ that characterizes $\mathcal{E}_\lambda$ together with the logarithm of λ.

In this case, the integral column is zero since the integrand after variable transformation becomes odd in $\vec{r}''$, and we are thus left with

$$\int_{\mathcal{R}_\lambda}(d\vec{r}')\,\vec{\Delta}'_{\rm ML}\,\vec{\Delta}'^{\rm T}_{\rm ML} \to \frac{|D^{-1/2}|}{V_{\mathcal{R}_0}}\,OD^{-1/2}ID^{-1/2}O^{\rm T}, \tag{4.63}$$

and

$$
\begin{aligned}
I &= \int (d\vec{r}'')_{\text{unif}}\, \eta(1-\vec{r}''^2)\, \vec{r}''\, \vec{r}''^{\mathrm{T}} \\
&= \int_{\text{unit sphere}} (d\vec{r}'')\, \vec{r}''\, \vec{r}''^{\mathrm{T}} \\
&= \int_0^1 dr''\, r''^{d+1} \int (d\{\text{solid angle}\})\, \vec{e}''\, \vec{e}''^{\mathrm{T}} = \frac{V_d}{d+2} 1\,,
\end{aligned}
\tag{4.64}
$$

where the last equality is explained by the orthogonally invariant of the $(d-1)$-dimensional solid-angle measure over the unit columns $\vec{e}$, and so

$$
\int_{\mathcal{R}_\lambda} (d\vec{r}')\, \vec{\Delta}'_{\text{ML}}\, \vec{\Delta}'^{\mathrm{T}}_{\text{ML}} \to \frac{V_d}{V_{\mathcal{R}_0}} (-2\log\lambda)^{d/2+1} |F_{\text{ML}}|^{-1/2}\, F_{\text{ML}}^{-1}\,. \tag{4.65}
$$

With all these components, the relevant asymptotic formulas concerning all three distance measures

$$
\begin{aligned}
S^{(\text{A})}_{\text{HS},\lambda} &\approx \text{Tr}[F_{\text{ML}}^{-1}] \frac{-\log\lambda}{d/2+1}\,, \\
S^{(\text{A})}_{\text{tr},\lambda} &: \quad \text{as in (4.47)}\,, \\
S^{(\text{A})}_{\text{B},\lambda} &\approx \text{Tr}[F_{\text{ML}}^{-1} Q_D] \frac{-\log\lambda}{d+2}\,.
\end{aligned}
\tag{4.66}
$$

Here Tr now addresses the dyadic character, as opposed to tr, and we witness the manifestation of logarithmic divergences from both the relaxation of $\partial\mathcal{R}_0$ and Gaussian approximation of L.

Next, to analytically calculate u_λ using $f(L) = \log L$ with which C_λ can be found,

we note that due to the Gaussian form of L,

$$u_\lambda = -\log\lambda - \frac{\int_{\mathcal{R}_\lambda} (d\vec{r}')\, \vec{\Delta}'^{\mathrm{T}}_{\mathrm{ML}} F_{\mathrm{ML}} \vec{\Delta}'_{\mathrm{ML}}}{2\int_{\mathcal{R}_\lambda} (d\vec{r}')} \tag{4.67}$$

is a dyadic trace function of $\int_{\mathcal{R}_\lambda} (d\vec{r}')\, \vec{\Delta}'_{\mathrm{ML}} \vec{\Delta}'^{\mathrm{T}}_{\mathrm{ML}}$, so that we may use the right-hand side of (4.65) and put down

$$u_{\mathrm{A},\lambda} = -\frac{2}{d+2}\log\lambda \tag{4.68}$$

after some basic trace and logarithmic manipulations. It is clear that $d/(d+2) \leq u_{\mathrm{A},\lambda} \leq 1$ is bounded.

Analytical expression of region capacity for Case B

In Case B, although the geometry of $\mathcal{R}_\lambda \approx \mathcal{E}'_\lambda \cap \mathcal{R}_0$ is now much trickier to deal with, the central limit theorem allows us to approximate $\mathcal{R}_\lambda$ by a regular analytical region.

As shown in Fig. 27, one can introduce a hyperplane $\mathcal{P}$, described by $\vec{n}\cdot(\vec{r}' - \vec{r}_{\mathrm{ML}}) = 0$ ($\vec{n} \propto \vec{g}_{\mathrm{ML}}$) that is tangent to the level curve of the Gaussian function at $\vec{r}_{\mathrm{ML}}$. The hyperspherical cap formed by $\mathcal{P}$ and $\mathcal{E}'_\lambda$ hence asymptotically contains $\mathcal{R}_\lambda$, where we have essentially modeled the highly nontrivial $\partial\mathcal{R}_\lambda \cap \partial\mathcal{R}_0$ as $\mathcal{P}$. This model im-

plies the estimated assignment

$$\begin{aligned}
&\int_{\mathcal{R}_\lambda} (d\vec{r}') \cdots \\
&= \frac{1}{V_{\mathcal{R}_0}} \int (d\vec{r}')_{\text{unif}}\, \eta(1 - (\vec{r}' - \vec{r}_c)^{\mathrm{T}} F_{\text{ML}} (\vec{r}' - \vec{r}_c)/(-2\log\lambda')) \\
&\qquad \times \eta(\vec{n} \cdot (\vec{r}_{\text{ML}} - \vec{r}')) \cdots .
\end{aligned} \tag{4.69}$$

The change of variable $\vec{r}' \to \vec{r}'' = D'^{1/2} O'^{\mathrm{T}} (\vec{r}' - \vec{r}_c)$ with respect to the diagonal coordinates of $F_{\text{ML}}/(-2\log\lambda') = O' D' O'^{\mathrm{T}}$ leads to

$$\int_{\mathcal{R}_\lambda} (d\vec{r}')\, q_\lambda(\vec{r}') \approx \frac{|D'^{-1/2}|}{V_{\mathcal{R}_0}} \int (d\vec{r}'')_{\text{cap}}\, q_\lambda(O' D'^{-1/2} \vec{r}'') \tag{4.70}$$

for any function q, which is parametrized by the cap element $(d\vec{r}'')_{\text{cap}} = (d\vec{r}'')_{\text{unif}}\, \eta(1 - \vec{r}''^2)\, \eta(a - \vec{b}^{\mathrm{T}} \vec{r}'')$, $a = \vec{n}^{\mathrm{T}}(\vec{r}_{\text{ML}} - \vec{r}_c)$ and $\vec{b} = D'^{-1/2} O'^{\mathrm{T}} \vec{n}$. One can check that

$$\begin{aligned}
l \equiv \frac{a}{|\vec{b}|} &= \frac{\vec{g}_{\text{ML}} \cdot (\vec{r}_{\text{ML}} - \vec{r}_c)}{|D'^{-1/2} O'^{\mathrm{T}} \vec{g}_{\text{ML}}|} \\
&= \sqrt{\frac{(\vec{r}_{\text{ML}} - \vec{r}_c)^{\mathrm{T}} F_{\text{ML}} (\vec{r}_{\text{ML}} - \vec{r}_c)}{(-2\log\lambda')}} \leq 1 .
\end{aligned} \tag{4.71}$$

In other words, we have

$$\overline{q_\lambda(\vec{r}')}^{\mathcal{R}_\lambda} \approx \frac{\int (d\vec{r}'')_{\text{cap}}\, q_\lambda(O' D'^{-1/2} \vec{r}'')}{\int (d\vec{r}'')_{\text{cap}}} , \tag{4.72}$$

and that for any q_λ belonging to either one of the three distance measures or $\log L - \log(\lambda L_{\max})$, the building blocks of $\overline{q_\lambda(\vec{r}')}^{\mathcal{R}_\lambda}$ are only $\int (d\vec{r}'')_{\text{cap}}$, $\int (d\vec{r}'')_{\text{cap}}\, \vec{r}''$ and $\int (d\vec{r}'')_{\text{cap}}\, \vec{r}''\, \vec{r}''^{\mathrm{T}}$. These integrations can be all carried out [93].

In combining all results gathered, we denote $\mathcal{N}_{d,l,x} = V_d\, \mathrm{I}_{(1-l)/2}((d+x)/2, (d+$

$x)/2)$, which depends on the incomplete Euler's beta function $\mathrm{I}_{\cdot}(\cdot,\cdot)$, and organize two new auxiliary quantities

$$\begin{aligned}\vec{m} &= \left[-\frac{V_{d-1}}{l(d+1)}\left(1-l^2\right)^{(d+1)/2}+\mathcal{N}_{d,l,1}\right] {F_{\mathrm{ML}}}^{-1}\vec{g}_{\mathrm{ML}}\,,\\ M &= \frac{-\log\lambda'}{d+2}\,\mathcal{N}_{d,l,3}\,{F_{\mathrm{ML}}}^{-1}+\frac{1}{2}\,\vec{m}\,\vec{g}_{\mathrm{ML}}^{\mathrm{T}}{F_{\mathrm{ML}}}^{-1}\,.\end{aligned} \tag{4.73}$$

This helps to clean the respective formulas

$$\begin{aligned}S_{\mathrm{HS},\lambda}^{(\mathrm{B})} &\approx \frac{\mathrm{Tr}[2M]}{\mathcal{N}_{d,l,1}}\,,\\ S_{\mathrm{tr},\lambda}^{(\mathrm{B})} &: \text{ as in (4.47)}\,,\\ S_{\mathrm{B},\lambda}^{(\mathrm{B})} &\approx \frac{\mathrm{tr}[P_r\vec{\Omega}^{\mathrm{T}}\vec{m}]+\mathrm{Tr}[M\,Q_r]}{\mathcal{N}_{d,l,1}} \approx \frac{\mathrm{tr}[P_r\vec{\Omega}^{\mathrm{T}}\vec{m}]}{\mathcal{N}_{d,l,1}}\,,\end{aligned} \tag{4.74}$$

for the distance-induced capacity functions and

$$\begin{aligned}u_{\mathrm{B},\lambda} = &[-\log\lambda'+\mathrm{Tr}[\vec{g}_{\mathrm{ML}}\vec{m}^{\mathrm{T}}-F_{\mathrm{ML}}M]/\mathcal{N}_{d,l,1}]\\ &\times\log(\lambda L_{\max})/\log(\lambda' L_{\max})\,.\end{aligned} \tag{4.75}$$

We caution the Reader once more regarding the actions of tr and Tr at the right-hand side of $S_{\mathrm{B},\lambda}^{(\mathrm{B})}$ in (4.74).

For consistency, we end this section by noting that Eqs. (4.74) and (4.75) cover Eqs. (4.66) and (4.67) because Case A implies that $\lambda'=\lambda$ ($\vec{g}_{\mathrm{ML}}=\vec{0}=\vec{m}$), such that $l=0$ then gives $\mathcal{N}_{d,0,x}=V_d$ and $M=(-\log\lambda)\,F_{\mathrm{ML}}^{-1}/(d+2)$.

Chapter 5

Conclusion

Quantum estimation theory will play a core role in science and technology as nano- and micro-scale systems become important. Considering feasibility of an estimation procedure, in this dissertation, we have studied two different topics in quantum estimation theory: quantum metrology using Gaussian states and Bayesian error certification. Since quantum estimation theory is obliged to be practically applicable, we have studied Gaussian states, which are relatively less demanding to generate and manipulate in laboratory but have a power to render quantum advantages in various applications of quantum information processing. Optical interferometry has been and will be extensively used for estimation of small displacements. Replacing classical resources used in existing interferometers with quantum counterparts, one will be able to broaden the scale of systems that can be handled and observe undiscovered features in a new field.

In order to constitute quantum enhanced optical interferometer, we have analyzed Mach-Zehnder interferometer with squeezed states. We have demonstrated the quantum enhancement obtained by nonclassical effects of squeezed states and its robustness. Considering practically accessible measurement schemes, we found that homodyne detection is a promising measurement to employ quantum enhancement under a practical loss mechanism.

Motivated by the robustness of homodyne detection against photon-loss, we have extended the detection to Gaussian measurement, a special case of which is homodyne detection. By comparing the ultimate sensitivity of general Gaussian states

and the sensitivity achieved by Gaussian measurement, we concluded that Gaussian measurement is not optimal if we consider photon-loss. Thus, non-Gaussian measurement is necessary even for characterizing Gaussian operations with Gaussian input probes for optimal precision.

Furthermore, optimal non-Gaussian measurements have been identified in general Gaussian quantum metrology and classified for different important tasks by exemplifying particular circumstances such as phase, squeezing, displacement, and loss estimation. Unfortunately, the experimental implementation of optimal measurement for phase and squeezing estimation is still unknown, so that it will be an important mission to figure it out.

In the second part, we study calculation of Bayesian error certification. In estimation procedure, it is necessary to provide an error interval or error region together with the estimator to present the reliability of the estimator. As the dimension of quantum systems under consideration grows, precise calculation of error regions becomes intractable even in numerical methods. In order to resolve the difficulty, we derived an approximated analytical expression in an asymptotic regime using Gaussian approximation. We confirmed that the expression matches very well with simulation when the dimension of Hilbert space is not large. One interesting point is that the expression works even with a reasonable number of copies. However, one problem of the approximation we found is that it gets inaccurate when the estimator is around the boundary of the parameter space and the dimension of the space is large. This difficulty arises due to the complicated boundary of the high-dimensional parameter space. To overcome the problem, we proposed an efficient numerical method for calculating the size and credibility of the region. The proposed method is called an in-region sampling and it is highly efficient even when the dimension of the parameter space is large because it does not sample in the

whole space. Our method enables one to calculate size and credibility of Bayesian error regions for high-dimensional systems. Finally, we introduced a concept of region capacity which measures the average distance from the estimator to the region, which can measure the size of error region in a different manner. As characterization and manipulation of high-dimensional systems is crucial for from quantum communication to quantum computing, our efficient numerical algorithm is expected to play an important role.

Bibliography

[1] B. P. Abbott et al., "Observation of gravitational waves from a binary black hole merger," *Physical Review Letters*, vol. 116, no. 6, p. 061102, 2016.

[2] C. M. Caves, "Quantum-mechanical noise in an interferometer," *Physical Review D*, vol. 23, no. 8, p. 1693, 1981.

[3] V. Giovannetti, S. Lloyd, and L. Maccone, "Quantum-enhanced measurements: beating the standard quantum limit," *Science*, vol. 306, no. 5700, pp. 1330–1336, 2004.

[4] V. Giovannetti, S. Lloyd, and L. Maccone, "Advances in quantum metrology," *Nature Photonics*, vol. 5, no. 4, p. 222, 2011.

[5] S. Pirandola, B. R. Bardhan, T. Gehring, C. Weedbrook, and S. Lloyd, "Advances in photonic quantum sensing," *Nature Photonics*, vol. 12, no. 12, pp. 724–733, 2018.

[6] A. Streltsov, G. Adesso, and M. B. Plenio, "Colloquium: Quantum coherence as a resource," *Reviews of Modern Physics*, vol. 89, no. 4, p. 041003, 2017.

[7] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, "Quantum entanglement," *Reviews of Modern Physics*, vol. 81, no. 2, p. 865, 2009.

[8] R. Demkowicz-Dobrzański, M. Jarzyna, and J. Kołodyński, "Quantum limits in optical interferometry," in *Progress in Optics*, vol. 60, pp. 345–435, Elsevier, 2015.

[9] J. Aasi et al., "Enhanced sensitivity of the ligo gravitational wave detector by using squeezed states of light," *Nature Photonics*, vol. 7, no. 8, p. 613, 2013.

[10] J. B. Fixler, G. Foster, J. McGuirk, and M. Kasevich, "Atom interferometer measurement of the newtonian constant of gravity," *Science*, vol. 315, no. 5808, pp. 74–77, 2007.

[11] A. D. Cronin, J. Schmiedmayer, and D. E. Pritchard, "Optics and interferometry with atoms and molecules," *Reviews of Modern Physics*, vol. 81, no. 3, p. 1051, 2009.

[12] J. J. Bollinger, W. M. Itano, D. J. Wineland, and D. J. Heinzen, "Optimal frequency measurements with maximally correlated states," *Physical Review A*, vol. 54, no. 6, p. R4649, 1996.

[13] H. Nha and H. Carmichael, "Proposed test of quantum nonlocality for continuous variables," *Physical Review Letters*, vol. 93, no. 2, p. 020401, 2004.

[14] J. Laurat, G. Keller, J. A. Oliveira-Huguenin, C. Fabre, T. Coudreau, A. Serafini, G. Adesso, and F. Illuminati, "Entanglement of two-mode Gaussian states: characterization and experimental production and manipulation," *Journal of Optics B: Quantum and Semiclassical Optics*, vol. 7, no. 12, p. S577, 2005.

[15] N. C. Menicucci, P. Van Loock, M. Gu, C. Weedbrook, T. C. Ralph, and M. A. Nielsen, "Universal quantum computation with continuous-variable cluster states," *Physical Review Letters*, vol. 97, no. 11, p. 110501, 2006.

[16] E. T. Jaynes and O. Kempthorne, "Confidence intervals vs bayesian intervals," in *Foundations of probability theory, statistical inference, and statistical theories of science*, pp. 175–257, Springer, 1976.

[17] P. M. Lee, *Bayesian statistics*. Arnold Publication, 1997.

[18] J. M. Bernardo and A. F. Smith, *Bayesian theory*, vol. 405. John Wiley & Sons, 2009.

[19] D. Suess, Ł. Rudnicki, D. Gross et al., "Error regions in quantum state tomography: computational complexity caused by geometry of quantum states," *New Journal of Physics*, vol. 19, no. 9, p. 093013, 2017.

[20] R. J. Glauber, "Coherent and incoherent states of the radiation field," *Physical Review*, vol. 131, no. 6, p. 2766, 1963.

[21] K. E. Cahill and R. J. Glauber, "Ordered expansions in boson amplitude operators," *Physical Review*, vol. 177, no. 5, p. 1857, 1969.

[22] A. Ferraro, S. Olivares, and M. G. Paris, "Gaussian states in continuous variable quantum information," *arXiv preprint quant-ph/0503237*, 2005.

[23] X.-B. Wang, T. Hiroshima, A. Tomita, and M. Hayashi, "Quantum information with Gaussian states," *Physics reports*, vol. 448, no. 1-4, pp. 1–111, 2007.

[24] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, "Gaussian quantum information," *Reviews of Modern Physics*, vol. 84, no. 2, p. 621, 2012.

[25] G. Adesso, S. Ragy, and A. R. Lee, "Continuous variable quantum information: Gaussian states and beyond," *Open Systems & Information Dynamics*, vol. 21, no. 01n02, p. 1440001, 2014.

[26] A. Serafini, *Quantum Continuous Variables: A Primer of Theoretical Methods*. CRC Press, 2017.

[27] J. Williamson, "On the algebraic problem concerning the normal forms of linear dynamical systems," *American journal of mathematics*, vol. 58, no. 1, pp. 141–163, 1936.

[28] B. Dutta, N. Mukunda, R. Simon, et al., "The real symplectic groups in quantum mechanics and optics," *Pramana*, vol. 45, no. 6, pp. 471–497, 1995.

[29] H. Cramér, *Mathematical methods of statistics*, vol. 9. Princeton university press, 1999.

[30] R. A. Fisher, "Theory of statistical estimation," in *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 22, pp. 700–725, Cambridge University Press, 1925.

[31] C. W. Helstrom, "Quantum detection and estimation theory," *Journal of Statistical Physics*, vol. 1, no. 2, pp. 231–252, 1969.

[32] S. Boixo, S. T. Flammia, C. M. Caves, and J. M. Geremia, "Generalized limits for single-parameter quantum estimation," *Physical Review Letters*, vol. 98, no. 9, p. 090401, 2007.

[33] S. L. Braunstein and C. M. Caves, "Statistical distance and the geometry of quantum states," *Physical Review Letters*, vol. 72, no. 22, p. 3439, 1994.

[34] C. Oh, S.-Y. Lee, H. Nha, and H. Jeong, "Practical resources and measurements for lossy optical quantum metrology," *Physical Review A*, vol. 96, no. 6, p. 062304, 2017.

[35] C. Oh, C. Lee, C. Rockstuhl, H. Jeong, J. Kim, H. Nha, and S.-Y. Lee, "Optimal Gaussian measurements for phase estimation in single-mode Gaussian metrology," *npj Quantum Information*, vol. 5, no. 1, p. 10, 2019.

[36] C. Oh, C. Lee, L. Banchi, S.-Y. Lee, C. Rockstuhl, and H. Jeong, "Optimal measurements for quantum fidelity between Gaussian states and its relevance to quantum metrology," *Physical Review A*, vol. 100, no. 1, p. 012323, 2019.

[37] W. Heisenberg, "Über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik," in *Original Scientific Papers Wissenschaftliche Originalarbeiten*, pp. 478–504, Springer, 1985.

[38] H. Lee, P. Kok, and J. P. Dowling, "A quantum rosetta stone for interferometry," *Journal of Modern Optics*, vol. 49, no. 14-15, pp. 2325–2338, 2002.

[39] J. P. Dowling, "Quantum optical metrology–the lowdown on high-N00N states," *Contemporary physics*, vol. 49, no. 2, pp. 125–143, 2008.

[40] M. Holland and K. Burnett, "Interferometric detection of optical phase shifts at the Heisenberg limit," *Physical Review Letters*, vol. 71, no. 9, p. 1355, 1993.

[41] P. M. Anisimov, G. M. Raterman, A. Chiruvelli, W. N. Plick, S. D. Huver, H. Lee, and J. P. Dowling, "Quantum metrology with two-mode squeezed vacuum: parity detection beats the Heisenberg limit," *Physical Review Letters*, vol. 104, no. 10, p. 103602, 2010.

[42] J. Joo, W. J. Munro, and T. P. Spiller, "Quantum metrology with entangled coherent states," *Physical Review Letters*, vol. 107, no. 8, p. 083601, 2011.

[43] J. Joo, K. Park, H. Jeong, W. J. Munro, K. Nemoto, and T. P. Spiller, "Quantum metrology for nonlinear phase shifts with entangled coherent states," *Physical Review A*, vol. 86, no. 4, p. 043828, 2012.

[44] P. Walther, J.-W. Pan, M. Aspelmeyer, R. Ursin, S. Gasparoni, and A. Zeilinger, "De broglie wavelength of a non-local four-photon state," *Nature*, vol. 429, no. 6988, p. 158, 2004.

[45] M. W. Mitchell, J. S. Lundeen, and A. M. Steinberg, "Super-resolving phase measurements with a multiphoton entangled state," *Nature*, vol. 429, no. 6988, p. 161, 2004.

[46] I. Afek, O. Ambar, and Y. Silberberg, "High-noon states by mixing quantum and classical light," *Science*, vol. 328, no. 5980, pp. 879–881, 2010.

[47] A. A. Michelson and E. W. Morley, "On the relative motion of the earth and of the luminiferous ether," *Sidereal Messenger, vol. 6, pp. 306-310*, vol. 6, pp. 306–310, 1887.

[48] B. Yurke, S. L. McCall, and J. R. Klauder, "SU (2) and SU (1, 1) interferometers," *Physical Review A*, vol. 33, no. 6, p. 4033, 1986.

[49] A. Divochiy, F. Marsili, D. Bitauld, A. Gaggero, R. Leoni, F. Mattioli, A. Korneev, V. Seleznev, N. Kaurova, O. Minaeva et al., "Superconducting nanowire photon-number-resolving detector at telecommunication wavelengths," *Nature Photonics*, vol. 2, no. 5, p. 302, 2008.

[50] C. Cahall, K. L. Nicolich, N. T. Islam, G. P. Lafyatis, A. J. Miller, D. J. Gauthier, and J. Kim, "Multi-photon detection using a conventional superconducting nanowire single-photon detector," *Optica*, vol. 4, no. 12, pp. 1534–1535, 2017.

[51] A. I. Lvovsky and M. G. Raymer, "Continuous-variable optical quantum-state tomography," *Reviews of Modern Physics*, vol. 81, no. 1, p. 299, 2009.

[52] L.-M. Duan, G. Giedke, J. I. Cirac, and P. Zoller, "Inseparability criterion for continuous variable systems," *Physical Review Letters*, vol. 84, no. 12, p. 2722, 2000.

[53] P. Giorda, M. Allegra, and M. G. Paris, "Quantum discord for Gaussian states with non-Gaussian measurements," *Physical Review A*, vol. 86, no. 5, p. 052328, 2012.

[54] S. Pirandola, G. Spedalieri, S. L. Braunstein, N. J. Cerf, and S. Lloyd, "Optimality of Gaussian discord," *Physical Review Letters*, vol. 113, no. 14, p. 140405, 2014.

[55] L. Pezzé and A. Smerzi, "Mach-Zehnder interferometry at the Heisenberg limit with coherent and squeezed-vacuum light," *Physical Review Letters*, vol. 100, no. 7, p. 073601, 2008.

[56] C. Gardiner, P. Zoller, and P. Zoller, *Quantum noise: a handbook of Markovian and non-Markovian quantum stochastic methods with applications to quantum optics*, vol. 56. Springer Science & Business Media, 2004.

[57] B. Escher, R. de Matos Filho, and L. Davidovich, "General framework for estimating the ultimate precision limit in noisy quantum-enhanced metrology," *Nature Physics*, vol. 7, no. 5, p. 406, 2011.

[58] M. O. Scully and M. S. Zubairy, *Quantum optics*, 1999.

[59] D. F. Walls and G. J. Milburn, *Quantum optics*. Springer Science & Business Media, 2007.

[60] L. Banchi, S. L. Braunstein, and S. Pirandola, "Quantum fidelity for arbitrary Gaussian states," *Physical Review Letters*, vol. 115, no. 26, p. 260501, 2015.

[61] C. C. Gerry and J. Mimih, "The parity operator in quantum optical metrology," *Contemporary Physics*, vol. 51, no. 6, pp. 497–511, 2010.

[62] K. P. Seshadreesan, P. M. Anisimov, H. Lee, and J. P. Dowling, "Parity detection achieves the Heisenberg limit in interferometry with coherent mixed with squeezed vacuum light," *New Journal of Physics*, vol. 13, no. 8, p. 083026, 2011.

[63] K. P. Seshadreesan, S. Kim, J. P. Dowling, and H. Lee, "Phase estimation at the quantum cramér-rao bound via parity detection," *Physical Review A*, vol. 87, no. 4, p. 043833, 2013.

[64] W. P. Schleich, *Quantum optics in phase space*. John Wiley & Sons, 2011.

[65] H. Vahlbruch, M. Mehmet, K. Danzmann, and R. Schnabel, "Detection of 15 dB squeezed states of light and their application for the absolute calibration of photoelectric quantum efficiency," *Physical Review Letters*, vol. 117, no. 11, p. 110801, 2016.

[66] M. D. Lang and C. M. Caves, "Optimal quantum-enhanced interferometry using a laser power source," *Physical Review Letters*, vol. 111, no. 17, p. 173601, 2013.

[67] M. Jarzyna and R. Demkowicz-Dobrzański, "Quantum interferometry with and without an external phase reference," *Physical Review A*, vol. 85, no. 1, p. 011801, 2012.

[68] R. Zamir, "A proof of the Fisher information inequality via a data processing argument," *IEEE Transactions on Information Theory*, vol. 44, no. 3, pp. 1246–1250, 1998.

[69] C. Ferrie, "Data-processing inequalities for quantum metrology," *Physical Review A*, vol. 90, no. 1, p. 014101, 2014.

[70] M. G. Genoni, S. Mancini, and A. Serafini, "General-dyne unravelling of a thermal master equation," *Russian Journal of Mathematical Physics*, vol. 21, no. 3, pp. 329–336, 2014.

[71] M. G. Paris, "Quantum estimation for quantum technology," *International Journal of Quantum Information*, vol. 7, no. supp01, pp. 125–137, 2009.

[72] J.-P. Blaizot and G. Ripka, *Quantum theory of finite systems*, vol. 3. MIT press Cambridge, MA, 1986.

[73] R. Nichols, P. Liuzzo-Scorpo, P. A. Knott, and G. Adesso, "Multiparameter Gaussian quantum metrology," *Physical Review A*, vol. 98, no. 1, p. 012114, 2018.

[74] A. Monras and M. G. Paris, "Optimal quantum estimation of loss in bosonic channels," *Physical Review Letters*, vol. 98, no. 16, p. 160401, 2007.

[75] A. Bhattacharyya, "On a measure of divergence between two statistical populations defined by their probability distributions," *Bull. Calcutta Math. Soc.*, vol. 35, pp. 99–109, 1943.

[76] A. Uhlmann, "The "transition probability" in the state space of a-algebra," *Reports on Mathematical Physics*, vol. 9, no. 2, pp. 273–279, 1976.

[77] C. A. Fuchs and C. M. Caves, "Mathematical techniques for quantum communication theory," *Open Systems & Information Dynamics*, vol. 3, no. 3, pp. 345–356, 1995.

[78] J. Twamley, "Bures and statistical distance for squeezed thermal states," *Journal of Physics A: Mathematical and General*, vol. 29, no. 13, p. 3723, 1996.

[79] H. Scutaru, "Fidelity for displaced squeezed thermal states and the oscillator semigroup," *Journal of Physics A: Mathematical and General*, vol. 31, no. 15, p. 3659, 1998.

[80] G.-S. Paraoanu and H. Scutaru, "Fidelity for multimode thermal squeezed states," *Physical Review A*, vol. 61, no. 2, p. 022306, 2000.

[81] P. Marian, T. A. Marian, and H. Scutaru, "Bures distance as a measure of entanglement for two-mode squeezed thermal states," *Physical Review A*, vol. 68, no. 6, p. 062309, 2003.

[82] H. Nha and H. Carmichael, "Distinguishing two single-mode Gaussian states by homodyne detection: an information-theoretic approach," *Physical Review A*, vol. 71, no. 3, p. 032336, 2005.

[83] S. Olivares, M. G. Paris, and U. L. Andersen, "Cloning of Gaussian states by linear optics," *Physical Review A*, vol. 73, no. 6, p. 062330, 2006.

[84] P. Marian and T. A. Marian, "Bures distance as a measure of entanglement for symmetric two-mode Gaussian states," *Physical Review A*, vol. 77, no. 6, p. 062319, 2008.

[85] P. Marian and T. A. Marian, "Uhlmann fidelity between two-mode Gaussian states," *Physical Review A*, vol. 86, no. 2, p. 022340, 2012.

[86] G. Spedalieri, C. Weedbrook, and S. Pirandola, "A limit formula for the quantum fidelity," *Journal of Physics A: Mathematical and Theoretical*, vol. 46, no. 2, p. 025304, 2012.

[87] A. S. Holevo, "Entropy gain and the Choi-Jamiolkowski correspondence for infinite-dimensional quantum evolutions," *Theoretical and Mathematical Physics*, vol. 166, no. 1, pp. 123–138, 2011.

[88] O. Pinel, P. Jian, N. Treps, C. Fabre, and D. Braun, "Quantum parameter estimation using general single-mode Gaussian states," *Physical Review A*, vol. 88, no. 4, p. 040102, 2013.

[89] Z. Jiang, "Quantum Fisher information for states in exponential form," *Physical Review A*, vol. 89, no. 3, p. 032128, 2014.

[90] Y. S. Teo, C. Oh, and H. Jeong, "Bayesian error regions in quantum estimation I: analytical reasonings," *New Journal of Physics*, vol. 20, no. 9, p. 093009, 2018.

[91] C. Oh, Y. S. Teo, and H. Jeong, "Bayesian error regions in quantum estimation II: region accuracy and adaptive methods," *New Journal of Physics*, vol. 20, no. 9, p. 093010, 2018.

[92] C. Oh, Y. S. Teo, and H. Jeong, "Probing bayesian credible regions intrinsically: a feasible error certification for physical systems," *Physical Review Letters*, vol. 123, no. 4, p. 040602, 2019.

[93] C. Oh, Y. S. Teo, and H. Jeong, "Efficient bayesian credible-region certification for quantum-state tomography," *Physical Review A*, vol. 100, no. 1, p. 012345, 2019.

[94] J. Shang, H. K. Ng, A. Sehrawat, X. Li, and B.-G. Englert, "Optimal error regions for quantum state estimation," *New Journal of Physics*, vol. 15, no. 12, p. 123026, 2013.

[95] M. Evans, *Measuring statistical evidence using relative belief*. Chapman and Hall/CRC, 2015.

[96] X. Li, J. Shang, H. K. Ng, and B.-G. Englert, "Optimal error intervals for properties of the quantum state," *Physical Review A*, vol. 94, no. 6, p. 062112, 2016.

[97] J. Řeháček, D. Mogilevtsev, and Z. Hradil, "Tomography for quantum diagnostics," *New Journal of Physics*, vol. 10, no. 4, p. 043022, 2008.

[98] I. Bengtsson, S. Weis, and K. Życzkowski, "Geometric methods in physics," in *XXX Workshop 2011 Trends in Mathematics*, 2013.

[99] K. Zyczkowski and H.-J. Sommers, "Hilbert–Schmidt volume of the set of mixed quantum states," *Journal of Physics A: Mathematical and General*, vol. 36, no. 39, p. 10115, 2003.

[100] A. Andai, "Volume of the quantum mechanical state space," *Journal of Physics A: Mathematical and General*, vol. 39, no. 44, p. 13641, 2006.

[101] S. M. Rump, "Verification of positive definiteness," *BIT Numerical Mathematics*, vol. 46, no. 2, pp. 433–452, 2006.

[102] B.-G. Englert, D. Kaszlikowski, H. K. Ng, W. K. Chua, J. Řeháček, and J. Anders, "Efficient and robust quantum key distribution with minimal state tomography," *arXiv preprint quant-ph/0412075*, 2004.

[103] T. Durt, C. Kurtsiefer, A. Lamas-Linares, and A. Ling, "Wigner tomography of two-qubit states and quantum cryptography," *Physical Review A*, vol. 78, no. 4, p. 042338, 2008.

[104] N. Bent, H. Qassim, A. Tahir, D. Sych, G. Leuchs, L. L. Sánchez-Soto, E. Karimi, and R. Boyd, "Experimental realization of quantum tomography of photonic qudits via symmetric informationally complete positive operator-valued measures," *Physical Review X*, vol. 5, no. 4, p. 041006, 2015.

[105] J. Shang, Y.-L. Seah, H. K. Ng, D. J. Nott, and B.-G. Englert, "Monte Carlo sampling from the quantum state space. I," *New Journal of Physics*, vol. 17, no. 4, p. 043017, 2015.

[106] Y.-L. Seah, J. Shang, H. K. Ng, D. J. Nott, and B.-G. Englert, "Monte Carlo sampling from the quantum state space. II," *New Journal of Physics*, vol. 17, no. 4, p. 043018, 2015.

[107] H. Zhu, "Quantum state estimation with informationally overcomplete measurements," *Physical Review A*, vol. 90, no. 1, p. 012115, 2014.

[108] Y. S. Teo, *Introduction to quantum-state estimation*. World Scientific, 2015.

[109] J. C. Butcher, *Numerical methods for ordinary differential equations*. John Wiley & Sons, 2016.

[110] P. Del Moral, A. Doucet, and A. Jasra, "Sequential Monte Carlo samplers," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 68, no. 3, pp. 411–436, 2006.

[111] C. J. Bélisle, H. E. Romeijn, and R. L. Smith, "Hit-and-run algorithms for generating multivariate distributions," *Mathematics of Operations Research*, vol. 18, no. 2, pp. 255–266, 1993.

[112] R. L. Smith, "The hit-and-run sampler: a globally reaching markov chain sampler for generating arbitrary multivariate distributions," in *Proceedings of the 28th conference on Winter simulation*, pp. 260–264, IEEE Computer Society, 1996.

[113] L. Lovász and S. Vempala, "Hit-and-run from a corner," *SIAM Journal on Computing*, vol. 35, no. 4, pp. 985–1005, 2006.

[114] S. Kiatsupaibul, R. L. Smith, and Z. B. Zabinsky, "An analysis of a variation of hit-and-run for uniform sampling from general regions," *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, vol. 21, no. 3, p. 16, 2011.

[115] B. Cousins and S. Vempala, "A practical volume algorithm," *Mathematical Programming Computation*, vol. 8, no. 2, pp. 133–160, 2016.

[116] L. Lovász, "Hit-and-run mixes fast," *Mathematical Programming*, vol. 86, no. 3, pp. 443–461, 1999.

[117] N. J. Higham, "Cholesky factorization," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 1, no. 2, pp. 251–254, 2009.

[118] S. Li, "Concise formulas for the area and volume of a hyperspherical cap," 2011.

[119] H. Zhu and B.-G. Englert, "Quantum state tomography with fully symmetric measurements and product measurements," *Physical Review A*, vol. 84, no. 2, p. 022327, 2011.

[120] E. P. Wigner, "Characteristic vectors of bordered matrices with infinite dimensions," *Annals of Mathematics*, pp. 548–564, 1955.

[121] E. P. Wigner, "Characteristics vectors of bordered matrices with infinite dimensions ii," *Annals of Mathematics*, pp. 203–207, 1957.

[122] A. Pizzo, D. Renfrew, and A. Soshnikov, "On finite rank deformations of wigner matrices," in *Annales de l'IHP Probabilités et statistiques*, vol. 49, pp. 64–94, 2013.

[123] M. L. Mehta, *Random matrices*, vol. 142. Elsevier, 2004.

[124] D. A. Bini, N. J. Higham, and B. Meini, "Algorithms for the matrix pth root," *Numerical Algorithms*, vol. 39, no. 4, pp. 349–378, 2005.

국문초록

물리적인 양의 정확한 측정은 과학 기술에서 핵심적이다. 본 학위 논문의 주제는 가우시안 상태를 이용한 양자 계측과 양자 계측에 있어서 효율적인 베이지안 오류 검정 방식을 제안하는 것이다. 먼저 우리는 광학적 간섭계에서 위상 차이를 계측하는 것을 분석한다. 광학적 간섭계는 매우 작은 변위를 측정하고자 할 때 과학 및 기술에 있어서 광범위하게 사용되는 도구이다. 최근에는 레이저 간섭계 중력파 관측소라고 불리는 곳에 있는 큰 규모의 광학적 간섭계를 이용하여 매우 작은 신호의 중력파를 관측해내는데 성공하였다. 한편, 비고전적인 양자 상태를 사용할 경우, 광학적 간섭계 등에서 높은 정확도를 같게 된다는 것이 알려졌다. 우리는 먼저 비고전적인 가우시안 상태와 실험적으로 구현 가능한 측정을 사용하였을 때, 마흐-젠더 간섭계에서 얻을 수 있는 정확도에 대해서 분석한다. 그리고 단일 모드 위상 추정에서 호모다인 측정만을 이용하여 구현 가능한 가우시안 측정을 사용하였을 때 최적의 정확도를 도달할 수 있는지 알아보고, 결론적으로 비가우시안 측정이 반드시 필요하다는 것을 밝혀내었다. 마지막으로 일반적인 가우시안 양자 계측에 있어서 최적의 측정을 찾고, 단일 모드의 경우에는 상황에 따라 세 가지 서로 다른 최적의 측정 장치가 존재한다는 것을 규명하였다.

두 번째로 우리는 양자 계측에서 핵심적인 역할을 하는 베이지안 오류 영역에 대해서 알아본다. 어떤 물리적인 값을 추정함에 있어서 우리는 추정값 뿐만 아니라 그에 대응하는 오류 영역을 반드시 제공하여야 한다. 하지만 양자 계측에 있어서 다루고자 하는 계의 차원이 커짐에 따라 오류 영역의 크기와 신용도를 계산하는 것이 기하급수적으로 오래걸린다는 것이 밝혀졌다. 우리는 이러한 문제를 해결하기 위해 점근적인 영역에서 오류 영역의 크기와 신용도의 근사적 표현을 유도하였다. 또한, 비점근적인 영역에서도 적용할 수 있는 효율적인 수치적 방법을

제시하였다.

주요어 : 양자 계측학, 가우시안 양자 정보, 베이지안 신용 영역

학번 : 2014-22371

감사의 글

박사학위를 받기까지의 학위과정에서 정말 많은 분들의 도움을 받았습니다. 가장 먼저 대학원에 진학하고 연구실에 들어와서 훌륭한 지도교수님인 정현석 교수님 밑에서 지도를 받게 된 것은 큰 행운이었습니다. 긴 박사학위 과정동안 교수님의 헌신적인 지도 덕분에 연구자로서 많은 성장을 이룰 수 있었습니다. 또한, 학부연구과정을 통해 박사과정동안 수행한 연구의 토대를 마련해주셨던 손원민 교수님께도 감사드리며, 박사학위 심사를 맡아주신 안경원 교수님, 신용일 교수님, 김도헌 교수님, 그리고 이진형 교수님께도 감사의 말씀 전합니다. 그리고 박사과정 동안에 진행하였던 연구 뿐만 아니라 인생의 선배로서 많은 도움을 주셨던 정갑균 박사님, 이수용 박사님, 그리고 이창협 박사님 정말 감사드립니다. 박사님들의 많은 조언과 도움 덕분에 성공적으로 박사 학위 과정을 마칠 수 있었습니다.

대학원 생활을 함께했던 Bobby, Yong Siah, Omkar, 영롱이형, 호용이형, 양승호 박사님, 박채연 박사님, 권혁준 박사님, 안대건 박사님, 인우, 성전, 두 석형, 성욱, 혁건이 뿐만 아니라 기민이형, 진우형, 민수형 모두 감사드립니다. 항상 연구실에서 선배님들 그리고 많은 후배들에게도 매일 새로운 것을 배울 수 있었습니다. 또한, 대학원 생활을 시작부터 함께한 우철, 다미, 용석, 윤석, 유정이 덕분에 즐겁게 박사학위 과정 지낼 수 있었고, 마찬가지로 비슷한 시기에 박사 과정 중에 있었던 많은 고민들과 고충들을 함께 토로하며 위로해준 경현이형, 수환이형, 동진이형, 현준, 동표, 그리고 비슷한 연구를 하며 많은 이야기를 나누었던 광일, 동근, 용기 모두 감사합니다. 그리고 학위 과정동안 자주 만나서 이야기하고 같이 인생에 대해서 고민해준 한석, 민성, 인봉, 승연, 민용, 은상, 그리고 성우, 진원, 정인이 모두 제가 열심히 살게 해준 원동력이 되어주었고, 정말 고맙다는 말 전하고 싶습니다. 자주 만나지는 못하지만 항상 힘이 되어주는 가족과 같은 주훈, 오평, 용우 모두 고맙다는 말 전합니다.

마지막으로, 박사 학위 과정동안 뿐만 아니라 항상 저를 사랑해주고 보살펴준 가족분들에게 무한한 감사를 표합니다. 멀리서도 항상 응원과 격려를 아낌없이 주시는 장모님, 장인어른께 감사드리며, 기나긴 박사 학위 과정을 함께하며 저에게 많은 힘이 되어준 저의 아내 유연이에게도 감사하고 사랑한다는 말 전합니다. 끝으로, 언제나 저를 믿어주고 응원해준 하나뿐인 형에게 정말 고맙다는 말 전하고, 어렸을 때부터 홀로 저를 키워주신 아버지의 노고와 헌신에 무한한 감사를 드리며, 일찍 세상을 떠나셨지만 현재의 제가 있게 해주신 어머니께도 감사드립니다.