


Adaptive quantum state estimation for two optical point sourcesMasataka Kimizu ^{*}*Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-0043, Japan*Fuyuhiko Tanaka [†]*Center for Education in Liberal Arts and Sciences, Osaka University, Toyonaka, Osaka 560-0043, Japan
and Center for Quantum Information and Quantum Biology, Osaka University, Toyonaka, Osaka 560-0043, Japan*Akio Fujiwara [‡]*Department of Mathematics, Osaka University, Toyonaka, Osaka 560-0043, Japan
and Center for Quantum Information and Quantum Biology, Osaka University, Toyonaka, Osaka 560-0043, Japan*

(Received 8 August 2023; accepted 21 February 2024; published 26 March 2024)

In classical optics, there is a well-known resolution limit called Rayleigh's curse in the separation of two incoherent optical sources in close proximity. Recently, Tsang *et al.* [*Phys. Rev. X* **6**, 031033 (2016)] revealed that this difficulty may be circumvented in the framework of quantum theory. Following their work, various estimation methods have been proposed to overcome Rayleigh's curse, but none of them enables us to estimate the positions of two point sources simultaneously based on single-photon measurements with high accuracy. In this study, we propose a method to *simultaneously* estimate the positions of two point sources with the highest accuracy using an adaptive quantum state estimation scheme.

DOI: [10.1103/PhysRevA.109.032434](https://doi.org/10.1103/PhysRevA.109.032434)**I. INTRODUCTION**

Discriminating two optical point sources is an important subject in optics that is expected to be applied to astronomical observations and biological imaging. However, the conventional method has a drawback called Rayleigh's curse [1] which makes it difficult to discriminate two point sources when they are close to each other. This problem can be translated as that of estimating the centroid and the separation of two point sources, and Rayleigh's curse represents the difficulty in estimating the separation when two point sources are close to each other. Recently, Tsang *et al.* [1] investigated this problem in the framework of quantum theory and showed that there is a possibility of estimating the separation of two point sources in close proximity with the same accuracy as when they are far apart. Moreover, they devised a measurement scheme called spatial mode demultiplexing (SPADE) that achieves this accuracy when the centroid of two point sources is known in advance.

The scheme SPADE allows us to accurately estimate the separation, but it requires prior knowledge of the centroid. Accordingly, a two-step procedure was proposed by Grace *et al.* [2] in which the centroid was to be estimated first. Meanwhile, Parniak *et al.* [3] and Bao *et al.* [4] investigated simultaneous estimation of the centroid and the separation, but they did not take into account the optimality of the measurement.

The optimal measurement for multiple parameters can be obtained from the simultaneous spectral decompositions of the symmetric logarithmic derivatives (SLDs) if they commute. Unfortunately, the SLDs of the centroid and the separation of two point sources do not, in general, commute [5]. In such cases, it is customary to search for a measurement that minimizes the weighted trace of the covariance matrix (or that of the inverse Fisher information matrix) [6–9]. Once the optimal measurement is obtained, the parameters can be estimated simultaneously with high accuracy using an estimation scheme called adaptive quantum state estimation (AQSE), which was proposed by Nagaoka [7] and theoretically justified by Fujiwara [8]. Since the optimal measurement generally depends on the true values of the parameters, AQSE updates the measurement sequentially.

In this study, we propose a method to *simultaneously* estimate the centroid and the separation of two point sources using AQSE. In particular, the measurement we use is the *optimal* one for estimating both the centroid and the separation, and the weighted trace of the sample covariance matrix is asymptotically the smallest in theory. Through numerical experiments, we confirm that the proposed method works effectively if the number of steps in AQSE is sufficiently large.

This paper is organized as follows. In Sec. II, we describe the mathematical formulation of our estimation problem. In Sec. III, we briefly summarize related works such as direct imaging and SPADE. In Sec. IV, we first introduce an AQSE scheme for two optical point sources using numerically obtained optimal measurements and then carry out numerical simulations of AQSE to demonstrate that the centroid and the separation can, in principle, be estimated simultaneously

^{*}kimizu@sigmath.es.osaka-u.ac.jp[†]ftanaka.celas@osaka-u.ac.jp[‡]fujiiwara@math.sci.osaka-u.ac.jp

with the best accuracy in the asymptotic limit. We also find a significant reduction in the rate of convergence of estimates as the separation of two point sources gets closer to zero; this phenomenon may correspond to Rayleigh's curse. Finally, we summarize the paper in Sec. V.

II. PROBLEM SETTING

In this section, we present the mathematical formulation of the problem we consider mainly based on Tsang *et al.* [1].

A. Mathematical formulation of our problem

The light emitted from two point sources is assumed to be quasimonochromatic and of equal brightness, and the image plane is assumed to be one-dimensional. Let $\epsilon \ll 1$ be the average number of photons observed at each temporal mode. The density operator in the image plane at each temporal mode is

$$\rho = (1 - \epsilon)\rho_0 + \epsilon\rho_1 + O(\epsilon^2), \quad (1)$$

where ρ_0 is the zero-photon state and ρ_1 is the one-photon state. Since two or more photons are almost never observed simultaneously in a single measurement when $\epsilon \ll 1$, we shall focus our attention only on the one-photon state ρ_1 .

We write $L^2(\mathbb{R})$ for the set of square integrable real-valued functions on \mathbb{R} . Let $|\psi_1\rangle, |\psi_2\rangle \in L^2(\mathbb{R})$ denote the states in the image plane of a single-photon emitted from each point source. Then, ρ_1 can be written as

$$\rho_1 = \frac{1}{2}(|\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2|). \quad (2)$$

This equation is, in fact, an approximation, but we will treat it as accurate. We assume that $|\psi_1\rangle$ and $|\psi_2\rangle$ are expressed as

$$|\psi_j\rangle = \int_{-\infty}^{\infty} dx \psi(x - x_j)|x\rangle, \quad j = 1, 2, \quad (3)$$

with $\psi(x)$ being the point-spread function and x_j being the coordinates of the j th point source satisfying $x_1 < x_2$. Here, $|x\rangle$ represents the ideal state in which the photon is localized exactly at position x . In this paper, we assume that the point-spread function $\psi(x)$ is Gaussian:

$$\psi(x) = \frac{1}{(2\pi\sigma^2)^{\frac{1}{4}}} \exp\left(-\frac{x^2}{4\sigma^2}\right), \quad (4)$$

where σ is a positive constant determined by the wavelength of the light and the properties of the lens.

Our problem is to estimate the true values of the coordinates x_1 and x_2 or, equivalently, the transformed parameters

$$\theta^1 = \frac{x_1 + x_2}{2}, \quad \theta^2 = x_2 - x_1 \quad (5)$$

simultaneously. Note that the superscripts in θ^1 and θ^2 do not represent powers. In what follows, we call θ^1 the centroid and θ^2 the separation and denote ρ_1 as ρ_θ , where $\theta = (\theta^1, \theta^2) \in \Theta = \mathbb{R} \times \mathbb{R}_{>0}$.

B. Cramér-Rao bound and quantum Fisher information matrix

In order to estimate the true values of the parameters, we apply a measurement $M = \{M(\omega) \mid \omega \in \Omega\}$ represented by a

positive operator-valued measure (POVM) to a one-photon state ρ_θ , where Ω is the set of measurement outcomes. Here, the measurement can be chosen arbitrarily, but once it is fixed, $p_\theta(\omega; M) = \text{Tr} \rho_\theta M(\omega)$ gives the probability distribution of the outcomes. This allows us to consider the Cramér-Rao inequality

$$V_\theta[M, \hat{\theta}] \geq J_\theta(M)^{-1}, \quad (6)$$

which gives a lower bound on the estimation error for any (locally) unbiased estimator $\hat{\theta}$, where $V_\theta[M, \hat{\theta}]$ is the covariance matrix and $J_\theta(M)$ is the (classical) Fisher information matrix of the parametric model $p_\theta(\cdot; M)$. In particular, given a measurement M , there is a locally unbiased estimator $\hat{\theta}$ that achieves the lower bound in (6) [6]. The asymptotic lower bound for the precision of the maximum likelihood estimator (MLE), which is not, in general, locally unbiased, is also characterized by $J_\theta(M)^{-1}$.

In quantum estimation theory, we often consider the quantum Fisher information matrix instead of the Fisher information matrix to evaluate the error bound. The quantum Fisher information matrix K_θ is defined using the Hermitian operator L_j satisfying the following equation:

$$\frac{\partial \rho_\theta}{\partial \theta^j} = \frac{1}{2}(\rho_\theta L_j + L_j \rho_\theta). \quad (7)$$

The operator L_j is called the symmetric logarithmic derivative in the direction θ^j . The (j, k) th component of the matrix K_θ is defined as

$$(K_\theta)_{jk} = \frac{1}{2} \text{Tr} \rho_\theta (L_j L_k + L_k L_j). \quad (8)$$

It is known that the Fisher information matrix $J_\theta(M)$ for a given measurement M is bounded from above by the quantum Fisher information matrix K_θ in that

$$J_\theta(M) \leq K_\theta. \quad (9)$$

If there exists a measurement M that achieves the upper bound in (9), it is the optimal measurement. In the case of single-copy measurements, a necessary condition for such a measurement to exist is that $\text{Tr}[\sqrt{\rho_\theta}[L_j, L_k]\sqrt{\rho_\theta}]$ are zero for all j, k [10,11], while a sufficient condition is that the SLDs $\{L_j\}$ commute. To the best of our knowledge, the gap between these two conditions has not yet been filled.

In our model, the quantum Fisher information matrix of ρ_θ with respect to the parameter θ^1 and θ^2 is written as

$$K_\theta = \begin{pmatrix} \frac{1}{\sigma^2} - \frac{(\theta^2)^2}{4\sigma^4} \exp\left(-\frac{(\theta^2)^2}{4\sigma^2}\right) & 0 \\ 0 & \frac{1}{4\sigma^2} \end{pmatrix}. \quad (10)$$

See [1] for a derivation. It is important to realize that, unless $\theta^2 = 2\sigma$, $\text{Tr}[\sqrt{\rho_\theta}[L_1, L_2]\sqrt{\rho_\theta}]$ is not zero, and there is no measurement that achieves the upper bound in (9) (see Fig. 1). Thus, we have to find an optimal measurement by another approach. We shall discuss this issue again in Sec. IV B.

III. BRIEF REVIEW OF PREVIOUS STUDIES

In this section, we briefly review the conventionally considered measurement and those proposed in related studies.

A. Direct imaging

Direct imaging is a simple method of measuring the position of a photon. The probability distribution of direct imaging is

$$\begin{aligned} p_\theta(x; M_{\text{direct}}) &= \text{Tr} \rho_\theta |x\rangle\langle x| = \frac{1}{2} (|\langle x|\psi_1\rangle|^2 + |\langle x|\psi_2\rangle|^2) \\ &= \frac{1}{2} [|\psi(x-x_1)|^2 + |\psi(x-x_2)|^2] \\ &= \frac{1}{2} \{ |\psi(x - (\theta^1 - \theta^2/2))|^2 \\ &\quad + |\psi(x - (\theta^1 + \theta^2/2))|^2 \}, \end{aligned} \quad (11)$$

which is a mixture of Gaussian distributions centered at θ^1 and shifted by $\pm\theta^2/2$.

In this measurement, when $\theta^2 \ll \sigma$, it is easy to estimate θ^1 because the Fisher information for θ^1 is almost equal to the quantum Fisher information, but it is difficult to estimate θ^2 because the Fisher information for θ^2 converges to zero in the limit of $\theta^2 \downarrow 0$ [1]. This fact is a variant of Rayleigh's curse in view of statistical estimation.

B. Hermite-Gaussian SPADE

Notwithstanding the above-mentioned fact, there is room for improving the precision of the estimation of θ^2 by means of a different type of measurement since the quantum Fisher information for θ^2 is a positive constant $1/(4\sigma^2)$ as seen from (10).

Hermite-Gaussian (HG) SPADE is a measurement proposed by Tsang *et al.* [1] to improve the accuracy of the estimation of θ^2 . In HG SPADE, assuming that the estimate $\hat{\theta}^1$ of the centroid is obtained *a priori*, the measurement is performed by the POVM $M_{\text{HG}} = \{|\phi_q\rangle\langle\phi_q| \mid q = 0, 1, \dots\}$, where

$$|\phi_q\rangle = \int_{-\infty}^{\infty} dx \phi_q(x - \hat{\theta}^1) |x\rangle, \quad q = 0, 1, \dots, \quad (12)$$

$$\phi_q(x) = \frac{1}{(2\pi\sigma^2)^{\frac{1}{4}}} \frac{1}{\sqrt{2^q q!}} H_q\left(\frac{x}{\sqrt{2}\sigma}\right) \exp\left(-\frac{x^2}{4\sigma^2}\right), \quad (13)$$

and $H_q(x)$ is the Hermitian polynomial. That is, we do not measure the position of the photon, but rather which spatial mode $|\phi_q\rangle$ the photon is in. The probability $p_\theta(q; M_{\text{HG}})$ of obtaining the measurement outcome q is

$$\begin{aligned} p_\theta(q; M_{\text{HG}}) &= \text{Tr} \rho_\theta |\phi_q\rangle\langle\phi_q| \\ &= \frac{1}{2} (|\langle\phi_q|\psi_1\rangle|^2 + |\langle\phi_q|\psi_2\rangle|^2) \\ &= \frac{1}{2} \left(\exp(-Q_1) \frac{Q_1^q}{q!} + \exp(-Q_2) \frac{Q_2^q}{q!} \right), \end{aligned} \quad (14)$$

where

$$Q_1 = \frac{1}{4\sigma^2} \left(\hat{\theta}^1 - \theta^1 + \frac{\theta^2}{2} \right)^2, \quad (15)$$

$$Q_2 = \frac{1}{4\sigma^2} \left(\hat{\theta}^1 - \theta^1 - \frac{\theta^2}{2} \right)^2. \quad (16)$$

Namely, $p_\theta(q; M_{\text{HG}})$ is a mixture of two Poisson distributions with parameters Q_1 and Q_2 , respectively.

Assume now that the value of the centroid is known exactly, $\hat{\theta}^1 = \theta^1$. Then

$$p_\theta(q; M_{\text{HG}}) = |\langle\phi_q|\psi_1\rangle|^2 = |\langle\phi_q|\psi_2\rangle|^2 = \exp(-Q) \frac{Q^q}{q!}, \quad (17)$$

where

$$Q = \frac{(\theta^2)^2}{16\sigma^2}, \quad (18)$$

and the Fisher information of this probability distribution for θ^2 is

$$\begin{aligned} J_\theta(M_{\text{HG}})_{22} &= \sum_{q=0}^{\infty} p_\theta(q; M_{\text{HG}}) \left\{ \frac{\partial}{\partial \theta^2} \ln p_\theta(q; M_{\text{HG}}) \right\}^2 \\ &= \frac{1}{4\sigma^2}. \end{aligned} \quad (19)$$

This is identical to the (2, 2)th entry of the quantum Fisher information matrix (10). In other words, if the true value of the centroid θ^1 is known exactly, HG SPADE is the best measurement for estimating the separation θ^2 .

However, Ref. [1] also pointed out that, if the estimate of the centroid $\hat{\theta}^1$ deviates even slightly from the true value, the Fisher information for θ^2 falls to zero in the limit of $\theta^2 \downarrow 0$.

C. Some other studies

Since HG SPADE requires accurate knowledge of the centroid θ^1 , a two-step procedure was proposed by Grace *et al.* [2] in which θ^1 was first estimated by direct imaging and then θ^2 was estimated by SPADE. Meanwhile, simultaneous estimation of θ^1 and θ^2 was studied by Parniak *et al.* [3] and Bao *et al.* [4]. Parniak *et al.* [3] used quantum correlation to measure two photons together and did not investigate simultaneous estimation with single-photon measurements without quantum correlations. Their measurement is physically feasible but is not necessarily optimal. Bao *et al.* [4], on the other hand, took a Bayesian approach to simultaneous estimation, but they also did not take into account the optimality of the measurement.

IV. ADAPTIVE PARAMETER ESTIMATION

Adaptive quantum state estimation, proposed by Nagaoka [7] and theoretically justified by Fujiwara [8], is an efficient estimation scheme for unknown parameters of a given quantum statistical model. In this section, we first briefly describe this estimation scheme and then apply it to the problem of estimating the positions of two point sources simultaneously.

A. Protocol

Given a quantum statistical model $\{\rho_\theta \mid \theta \in \Theta \subset \mathbb{R}^d\}$, let θ_* be the true value of the parameter and write $M(\cdot; \theta_*)$ for its optimal measurement, taking into account the fact that, in general, the optimal measurement depends on the unknown true value of the parameter. In order to circumvent this difficulty, one may invoke an AQSE protocol which runs as follows: choose the initial estimate $\hat{\theta}_0 \in \Theta$ arbitrarily and repeat the following for steps $n = 1, 2, \dots$

(1) Apply the measurement $M(\cdot; \hat{\theta}_{n-1})$, which is optimal at the previous estimate $\hat{\theta}_{n-1}$, to yield the n th outcome ω_n .

(2) Obtain the next estimate $\hat{\theta}_n$ from the data $(\omega_1, \dots, \omega_n)$ using the maximum likelihood method, i.e.,

$$\hat{\theta}_n = \arg \max_{\theta \in \Theta} \prod_{i=1}^n \text{Tr} \rho_\theta M(\omega_i; \hat{\theta}_{i-1}). \quad (20)$$

It was shown in [8] that, under some regularity conditions, $\hat{\theta}_n$ enjoys strong consistency,

$$\hat{\theta}_n \longrightarrow \theta_* \quad (21)$$

with probability 1, and asymptotic efficiency,

$$\sqrt{n}(\hat{\theta}_n - \theta_*) \longrightarrow N(0, J_{\theta_*}(M(\cdot; \theta_*))^{-1}) \quad (22)$$

in distribution. In actual experiments, n cannot be infinitely large and must be stopped at some point. However, if it is stopped at a sufficiently large n , the left-hand side of (22) approximately follows the distribution of the right-hand side, and a good estimation accuracy can be obtained.

B. Optimal measurement in simultaneous estimation of $\theta = (\theta^1, \theta^2)$

Prior to applying AQSE, we need to obtain a list of optimal measurements $M(\cdot; \theta)$ for all $\theta \in \Theta$. Since the asymptotic fluctuation of the estimate $\hat{\theta}_n$ obtained by AQSE is characterized by the inverse of the Fisher information matrix $J_\theta(M)$ as in (22), one may conceive that the optimal measurement would be the one that makes $J_\theta(M)^{-1}$ as small as possible. But in reality, one cannot minimize it since it is a matrix.

One approach to finding the optimal measurement is to minimize the weighted trace of the inverse of the Fisher information matrix [6–9]:

$$M(\cdot; \theta) = \arg \min_{M: \text{POVM}} \text{Tr} G J_\theta(M)^{-1}, \quad (23)$$

given a positive-definite matrix G , called the weight matrix, which may depend on the parameter θ .

In our problem, the underlying Hilbert space is $L^2(\mathbb{R})$, and thus, $M(\cdot; \theta)$ in (23) must be obtained as a POVM on $L^2(\mathbb{R})$. However, as discussed by Shao and Lu [11], ρ_θ , $\frac{\partial \rho_\theta}{\partial \theta^1}$, and $\frac{\partial \rho_\theta}{\partial \theta^2}$ that appear in $J_\theta(M)$ have support on a θ -dependent four-dimensional subspace \mathcal{V}_θ of $L^2(\mathbb{R})$, and it is enough to obtain $M(\cdot; \theta)$ as a POVM on \mathcal{V}_θ ; in fact, we need only add $I_{\mathcal{V}_\theta^\perp}$ to obtain the POVM on $L^2(\mathbb{R})$.

Unfortunately, the analytical solution for the minimization problem (23) is not known unless the underlying Hilbert space is two-dimensional [9]. We therefore invoke numerical methods to find the optimal measurement $M(\cdot; \theta)$ for each θ . Note that it is sufficient to consider 16-valued real rank-1 measurements [12], and the minimization problem is reduced to an unconstrained nonlinear programming problem [13]; see the Appendix for details. In what follows, we choose the weight matrix G to be the quantum Fisher information matrix K_θ . This choice is beneficial because the solution of the minimization problem (23) with this choice depends only on the state ρ_θ and is independent of the parametrization θ .

Figure 1 demonstrates the solution for the minimization problem (23). The horizontal axis is set to θ^2/σ because the minimum values depend only on θ^2 due to the covariant

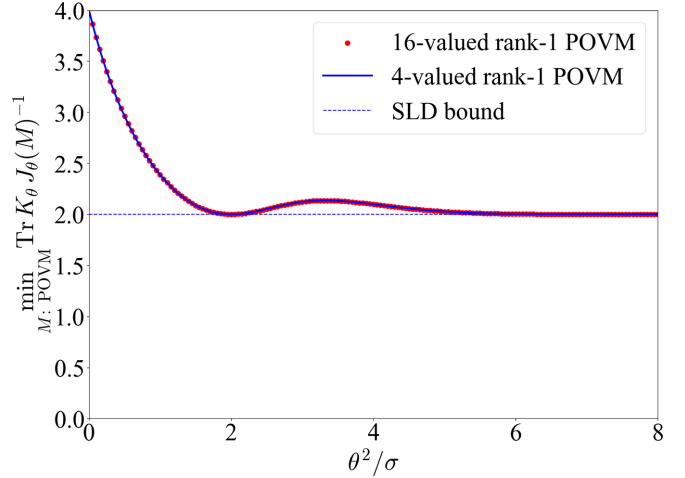


FIG. 1. Minimum values of the weighted trace of the inverse Fisher information matrix. The blue dashed line is the SLD bound $\text{Tr} G K_\theta^{-1} = 2$. The SLD bound is achieved when $\theta^2 = 2\sigma$ and is nearly achieved when $\theta^2 \gtrsim 6\sigma$.

nature of the model under parallel translation of the optical point sources. The red dots are the results of optimization with 16-valued rank-1 measurements, and the blue dashed line shows the SLD bound $\text{Tr} G K_\theta^{-1} = 2$. Figure 1 shows that the SLD bound is achieved when $\theta^2 = 2\sigma$ and is nearly achieved when $\theta^2 \gtrsim 6\sigma$. This is because the SLDs for θ^1 and θ^2 can be taken to be commutative when $\theta^2 = 2\sigma$ [5], and the two SLDs are nearly commutative when θ^2 is sufficiently large. Furthermore, the blue curve is the result of optimization with four-valued rank-1 measurements, showing that the minimum value is achieved with four-valued measurement. In particular, this four-valued rank-1 measurement is a projective measurement since \mathcal{V}_θ is four-dimensional.

For reference, Fig. 2 shows the numerically obtained wave functions $q_\omega(x) = \langle x | q_\omega \rangle$ of the basis of the optimal measurements $M(\cdot; (0, \theta^2))$, i.e.,

$$M(\omega; (0, \theta^2)) = |q_\omega\rangle\langle q_\omega| \quad (24)$$

for each θ^2 . The plot for $\theta^2 = 2\sigma$ corresponds to Fig. 3 in [5], but the slight difference is due to the nonuniqueness of the optimal measurement.

Regarding the physical realization, a measurement device that can represent arbitrary wave functions is necessary in order to sequentially update the measurement according to the estimate (see Fig. 2), but Sajjad *et al.* [14] stated that “any projective measurement on a quantum state of one photon in many (spatial) modes, which is the case for the quantum description of the state of a single temporal mode of collected light in our problem, is always realizable by a passive linear optical transformation followed by photon detection.”

Note that $\theta^2 = 2\sigma$ is the threshold at which the modality of the probability distributions for the direct imaging changes. In fact, as shown in Fig. 3, the probability distribution for direct imaging on $\rho_{(0, \theta^2)}$ is unimodal when $0 < \theta^2 < 2\sigma$, while it is bimodal when $\theta^2 > 2\sigma$.

Next, we confirm that the obtained optimal measurement is superior to the ones used in the previous studies. Figure 4 shows the components of the Fisher information

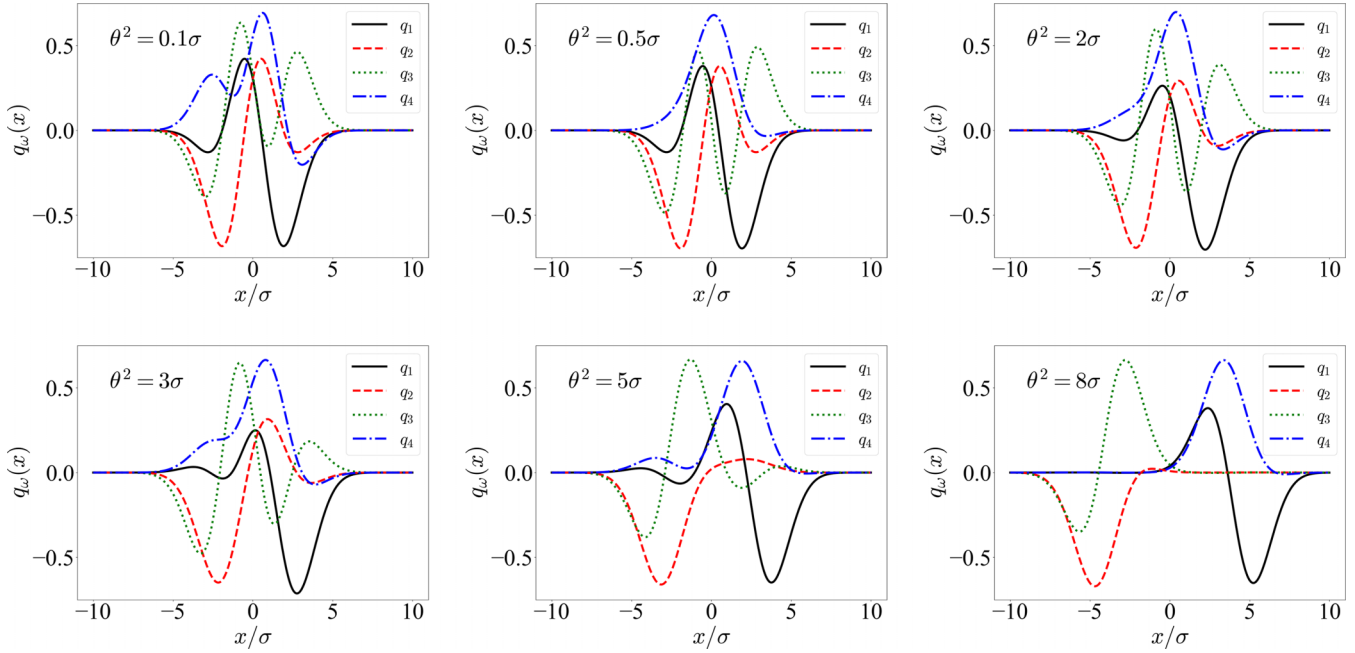


FIG. 2. The wave functions of the optimal measurement basis.

matrix $J_\theta(M(\cdot; \theta))$ of the optimal measurement. Since the off-diagonal components are zero, only the diagonal components are shown. The blue curves are the components of the quantum Fisher information matrix K_θ , which gives an upper bound of $J_\theta(M(\cdot; \theta))$ as in (9). It is noteworthy that under the framework of simultaneous estimation of the centroid θ^1 and the separation θ^2 , the Fisher information for θ^2 (dash-dotted red curve) converges to a positive value in the limit of $\theta^2 \downarrow 0$. In other words, it is possible to estimate θ^1 and θ^2 simultaneously with reasonable accuracy no matter how close to zero the separation θ^2 is.

This fact is in remarkable contrast to [2] in the following sense: They proposed a scheme to perform binary SPADE after estimating the centroid θ^1 by direct imaging, but the actual binary SPADE is misaligned since the value of θ^1 cannot be estimated exactly. Therefore, in their scheme, the

Fisher information for θ^2 falls to zero in the limit of $\theta^2 \downarrow 0$ [1]. Indeed, Fig. 4 in [2] shows that the mean squared error (MSE) increases dramatically as θ^2 approaches zero. On the other hand, Fig. 4 in our study demonstrates that our scheme has a smaller MSE than theirs when θ^2 is close to zero, even though we estimate θ^1 and θ^2 simultaneously.

The optimality of our measurements can also be verified through information regret. The (normalized-square-root) information regret with respect to θ^j is defined as [10,11]

$$\Delta_j = \sqrt{\frac{(K_\theta)_{jj} - J_\theta(M)_{jj}}{(K_\theta)_{jj}}} \quad (25)$$

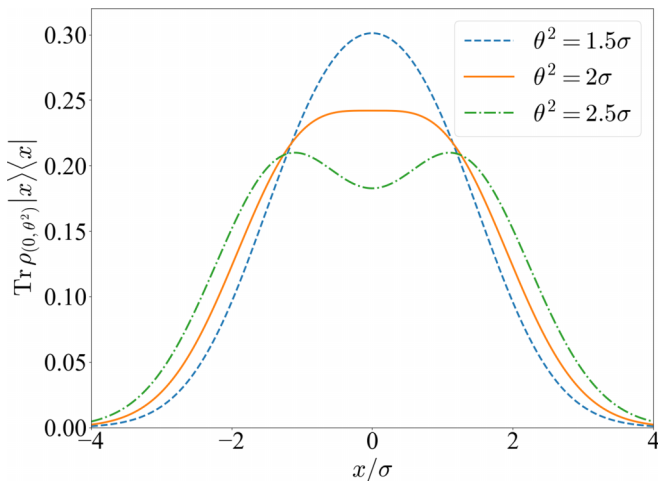


FIG. 3. Probability distributions for direct imaging. The value $\theta^2 = 2\sigma$ is the boundary between unimodal and bimodal.

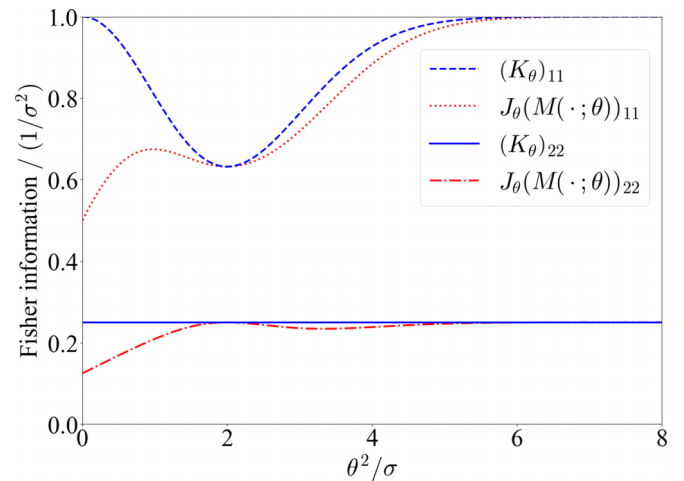


FIG. 4. Comparison of the classical Fisher information under optimal measurements (red curves) with the quantum Fisher information (blue curves). Note that the classical Fisher information for θ^1 and θ^2 converges to positive values in the limit $\theta^2 \downarrow 0$.

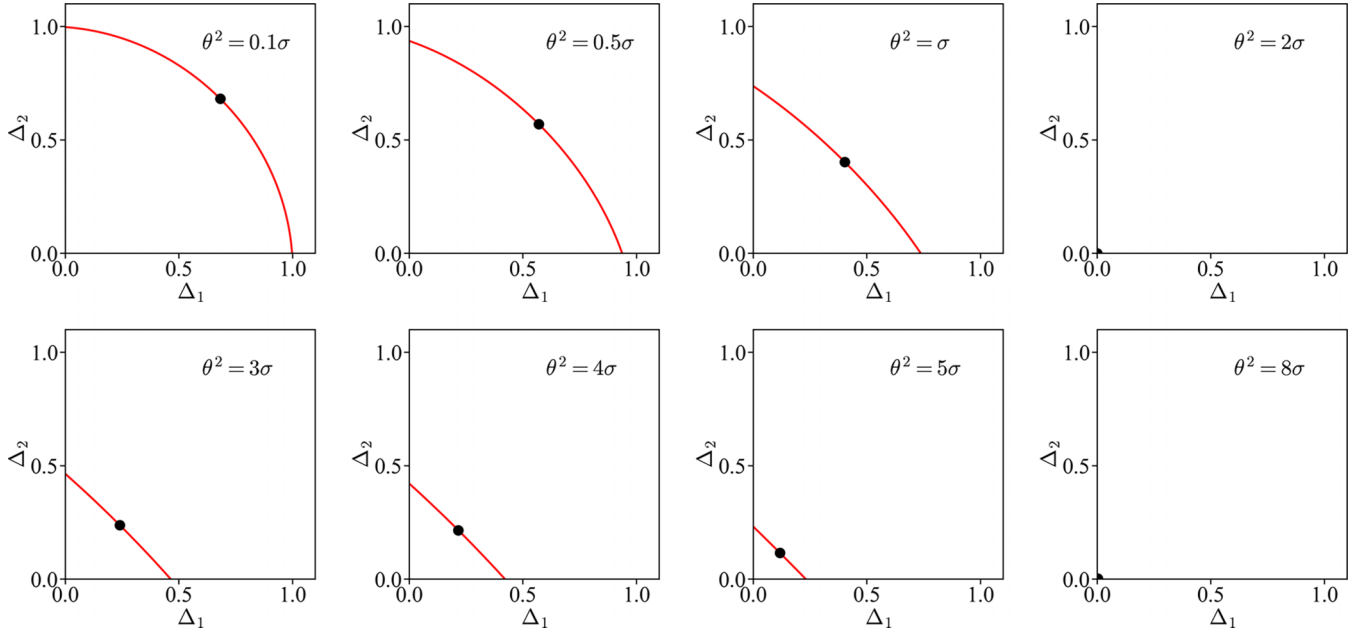


FIG. 5. The information-regret trade-off relation between θ^1 and θ^2 (red curve) and the information regrets of the numerically obtained optimal measurements (black dot).

and takes values between 0 and 1 for each j . When the equality in (9) is achieved, Δ_j are all zero, but in general, they cannot be zero for all j , and the trade-off relation

$$\Delta_j^2 + \Delta_k^2 + 2\sqrt{1 - \tilde{c}_{jk}^2} \Delta_j \Delta_k \geq \tilde{c}_{jk}^2 \quad (26)$$

holds. Here, \tilde{c}_{jk} is defined as

$$\tilde{c}_{jk} = \frac{\text{Tr}[\sqrt{\rho_\theta}[L_j, L_k]\sqrt{\rho_\theta}]}{2\sqrt{(K_\theta)_{jj}(K_\theta)_{kk}}}. \quad (27)$$

If a measurement achieves the equality in (26) for each j, k satisfying $j \neq k$, it is the best measurement among single-copy measurements. In each plot in Fig. 5, the horizontal axis represents the information regret of θ^1 , the vertical axis represents the information regret of θ^2 , and the red line shows the information regret trade-off relation (26) for θ^1 and θ^2 in the model of two point sources. The information regrets of our measurements are represented by the black dots, and all of them are on the red line.

C. Simulating AQSE

Now we proceed to numerical simulations of AQSE for two point sources using the optimal measurements obtained in the previous section. In the rest of this paper, we set $\sigma = 1$ without loss of generality.

The settings for AQSE are as follows. The true value of the parameter is $\theta_* = (\theta_*^1, \theta_*^2) = (0, 0.3)$, the initial estimate is $\hat{\theta}_0 = (1, 1)$, and the estimate is computed up to $n = 8000$ steps. Since the measurement is updated step by step, it is computationally demanding to obtain a rigorous maximum likelihood estimate. Therefore, the maximum likelihood estimate is approximately obtained by computing the log-likelihood at predefined grid points. Since we want to

check the asymptotic behavior of the estimates, the grid points are set more finely around the true value of the parameter.

We calculated a sequence of estimates $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_{8000} \in \Theta$ in each run of AQSE and repeated such runs 1000 times to obtain 1000 samples of the sequence of estimates.

First, we check the consistency (21). Figure 6 plots the estimates $\hat{\theta}_n$ for $n = 1000, 2000, 4000, \text{ and } 8000$. In each plot, the horizontal axis is θ^1 , and the vertical axis is θ^2 , with the blue dots representing estimates and the cross representing the true value. The estimates are initially widely scattered around the true value, but as the number of steps increases, the estimates get closer to the true value.

Next, we check the asymptotic normality (22). We performed goodness-of-fit tests on 1000 samples of $\hat{\theta}_{8000}$ under the null hypothesis that they follow a multivariate normal distribution. The Anderson-Darling test in the MVNTEST package of R yielded a p value of 0.9349, and the Cramér-von Mises test in the same package yielded a p value of 0.9434. The null hypothesis was accepted with a very high p value for both tests.

Finally, we check how the sample covariance matrix $V[\hat{\theta}_n]$ evolves with the number of steps. Figure 7(a) shows the weighted trace of the sample covariance matrix $\text{Tr}K_{\hat{\theta}}V[\hat{\theta}_n]$, where $K_{\hat{\theta}}$ is the quantum Fisher information matrix at the sample mean $\hat{\theta}$ of the estimates at each step. For the sake of comparison, Fig. 7(b) shows the result for the case where the true value is $\theta_* = (\theta_*^1, \theta_*^2) = (0, 0.1)$. Note that the values of the weighted trace are multiplied by n since the sample covariance matrix decreases by $1/n$. The dashed lines indicate the ultimate limits of estimation precision displayed in Fig. 1. In each case, the solid curve approaches the dashed line as the number of steps increases. This means that if the number of steps is large enough, we can estimate the centroid θ^1 and the separation θ^2 simultaneously with the best accuracy theoretically possible.

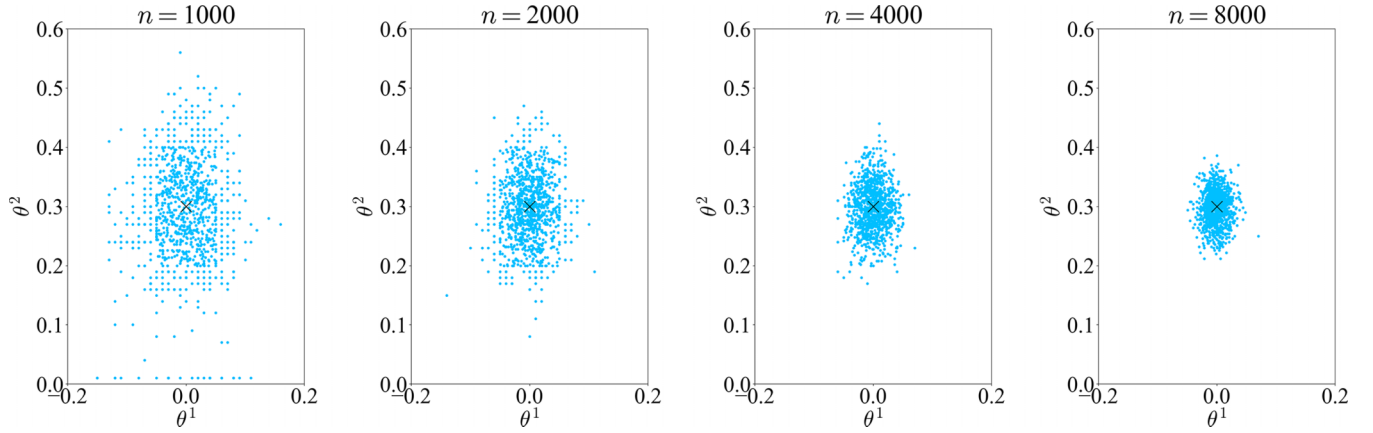


FIG. 6. Scatterplots of estimated values for different numbers of steps n when $\theta_* = (\theta_*^1, \theta_*^2) = (0, 0.3)$. Each blue dot represents an estimate, and the cross represents the true value of the parameter.

D. Trapping phenomena near $\theta^2 = 0$

It is noteworthy that, as can be seen from Fig. 7, the convergence becomes much slower as θ_*^2 gets closer to zero. We also find from the heat maps in Fig. 8 that, when θ_*^2 is small, a significant number of MLEs are trapped near the

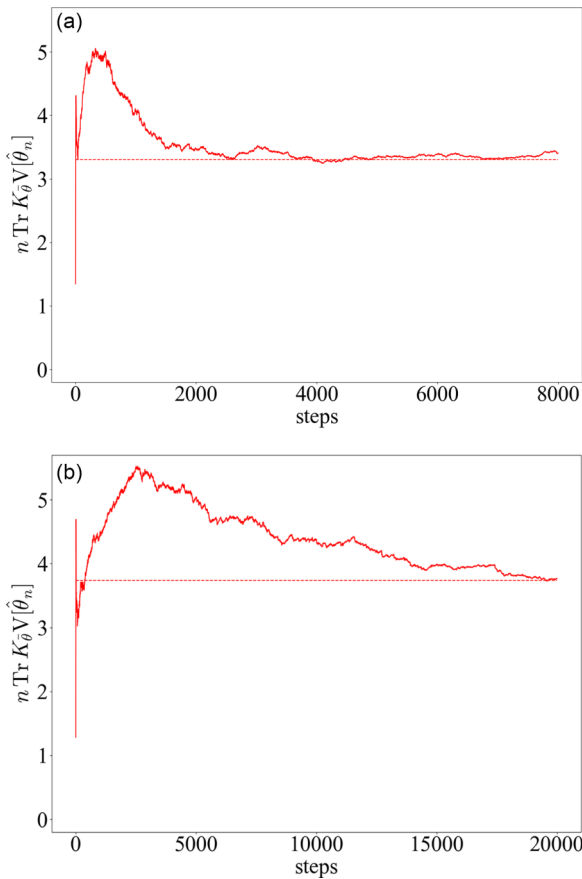


FIG. 7. Weighted trace of the sample covariance matrix for (a) $\theta_* = (\theta_*^1, \theta_*^2) = (0, 0.3)$ and (b) $\theta_* = (\theta_*^1, \theta_*^2) = (0, 0.1)$. The dashed lines indicate theoretical limits given in Fig. 1. Comparing (a) and (b), we find a notable reduction in the rate of convergence of sample covariance as θ_*^2 approaches zero.

boundary $\theta^2 = 0$ for a long time. These observations prompt us to envisage the following scenario: When θ_*^2 is small, a good number of estimates are located in the boundary region $\theta^2 \approx 0$ at an early stage of AQSE because of the large sample dispersion and are kept trapped in that region for a long time, yielding a notable slowdown of the convergence of the sample covariance matrix.

Let us examine the validity of this “boundary-effect” scenario by means of the following tentative evaluation: Because of the nature of convergence in distribution, each contour of the probability density function would converge to θ_* at the rate $\sim 1/\sqrt{n}$, so that the time τ for a certain contour to pass through the “trapping wall,” i.e., the grid line closest to the axis $\theta^2 = 0$, at a distance d from the true parameter θ_* may be evaluated as

$$d \sim \frac{C}{\sqrt{\tau}} \iff \tau \sim \frac{C^2}{d^2}, \quad (28)$$

where C is a certain constant corresponding to the contour that characterizes the trapping effect. Assume further that, after the contour moves away from the influence of the trapping wall, the time t_0 required for the estimates to converge in distribution is independent of θ_* . Then the total time $T = \tau + t_0$ of convergence in distribution would roughly be evaluated as

$$T \sim \frac{C^2}{(\theta_*^2)^2} + t_0. \quad (29)$$

Let us verify the validity of this scaling law. The following is a list of convergence times T obtained by numerical simulations for several values of θ_*^2 :

$$(\theta_*^2, T) = (0.1, 14372), \quad (0.2, 3816), \quad (0.3, 1744), \\ (0.5, 990), \quad (0.7, 653), \quad (1.0, 376);$$

the first and third data points correspond to Figs. 7(b) and 7(a), respectively. Here, we take T as the first time at which the weighted trace of the sample covariance matrix decreases to within 5% of the theoretical limit. A nonlinear parameter fitting using the FINDFIT function of *Mathematica* yields

$$T = \frac{c}{(\theta_*^2)^e} + t_0, \quad (30)$$

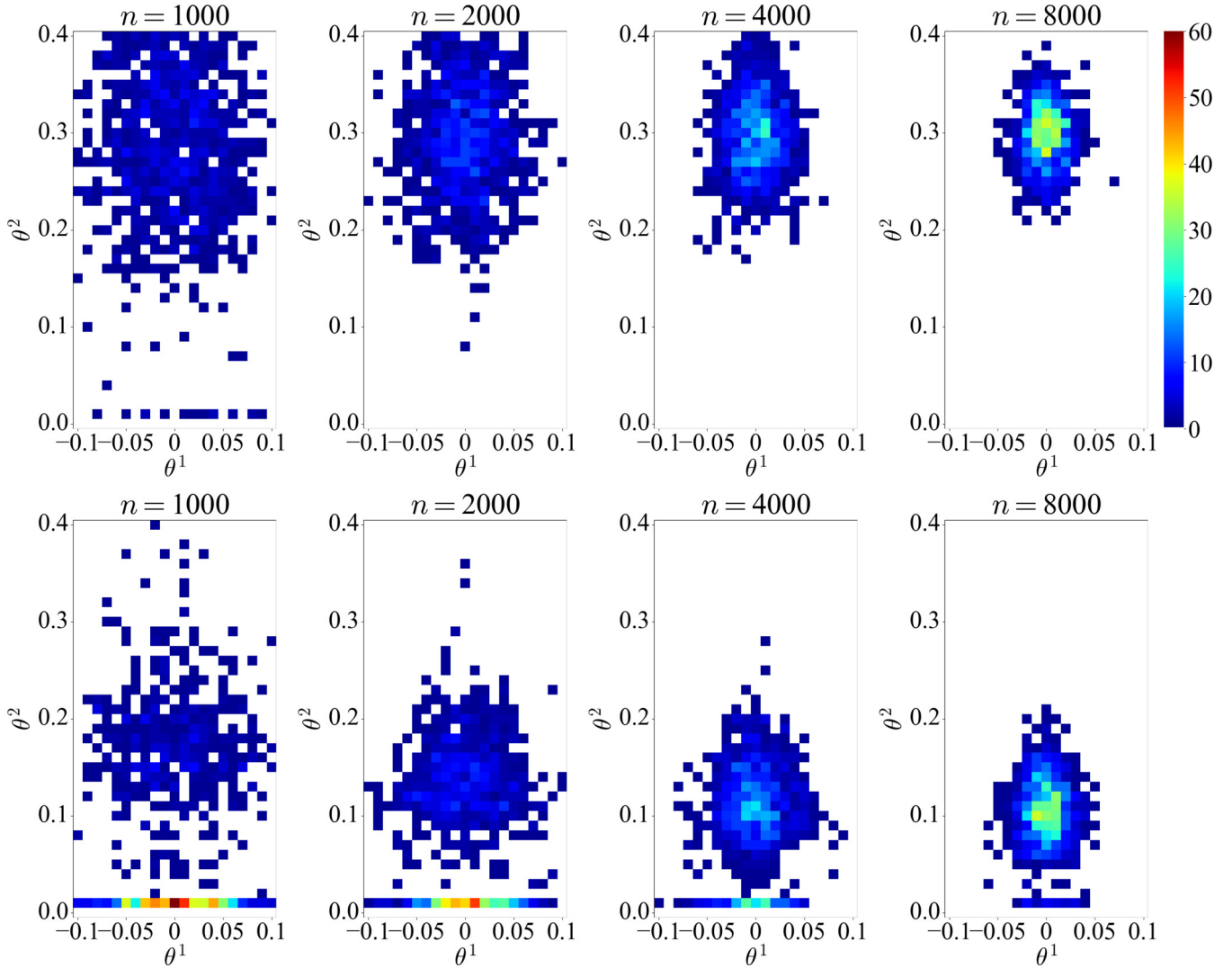


FIG. 8. Heat maps of the frequency of estimates for different numbers of steps n for $\theta_* = (\theta_*^1, \theta_*^2) = (0, 0.3)$ (top row) and $\theta_* = (\theta_*^1, \theta_*^2) = (0, 0.1)$ (bottom row). All heat maps share the same color bar, which is displayed in the upper right. Comparing the top and bottom panels, we see that when θ_*^2 is small, a significant number of estimates are trapped near the boundary $\theta^2 = 0$ for a long time.

with $e = 2.02$, $c = 133$, and $t_0 = 333$, as seen in Fig. 9. This result is reasonably consistent with the scaling law (29), supporting the validity of the trapping scenario.

In summary, although the centroid θ_*^1 and the separation θ_*^2 can, in principle, be estimated simultaneously with the best accuracy in the asymptotic limit, a notable reduction in the rate of convergence of estimates arises as the separation θ_*^2 gets closer to zero. The slowdown of the convergence of estimates as $\theta_*^2 \downarrow 0$ may be regarded as a manifestation of Rayleigh's curse in the quantum domain.

V. CONCLUSION

In this paper, we proposed a method to estimate the centroid θ^1 and the separation θ^2 of two point sources simultaneously by AQSE. Numerical experiments confirmed that the method works properly if the number of steps is large enough. It was also found that the closer to zero the θ^2 component of the true value is, the slower the convergence of estimates becomes. This phenomenon may suggest that

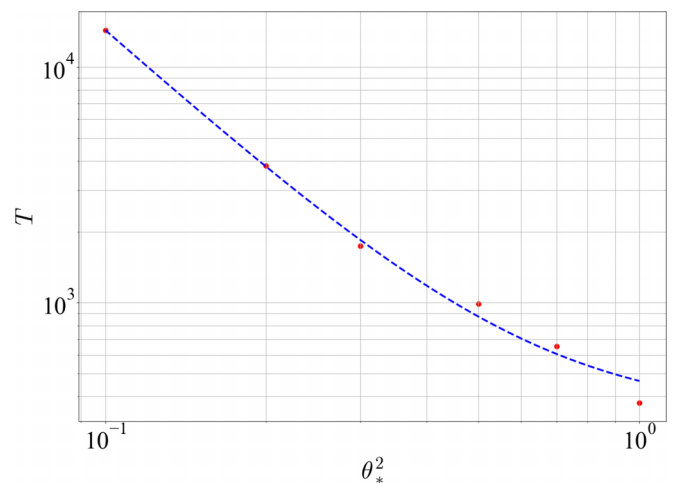


FIG. 9. Time T required for the estimates to converge in distribution for several values of θ_*^2 . The data are fitted by the curve $T = c/(\theta_*^2)^e + t_0$, with $e = 2.02$, $c = 133$, and $t_0 = 333$.

Rayleigh's curse may still survive in the framework of quantum theory, transforming itself into a *plateau* phenomenon, a notable reduction in the rate of convergence of estimates in AQSE. Nevertheless, the mechanism behind the plateau phenomenon requires further investigation.

ACKNOWLEDGMENTS

M.K. is grateful to Prof. K. Yamagata for providing him with Ref. [13] and kindly allowing him to include portions of it in this paper. This work was supported by JSPS KAKENHI Grants No. JP17H02861, No. JP20H02168, and No. JP22H00510.

APPENDIX: NUMERICAL OPTIMIZATION OF MEASUREMENT

In this Appendix, we explain how to reduce the problem of finding the optimal measurement $M(\cdot; \theta)$ given by (23) to an unconstrained nonlinear programming problem [13].

As described in Sec. IV B, ρ_θ , $\frac{\partial \rho_\theta}{\partial \theta^1}$, and $\frac{\partial \rho_\theta}{\partial \theta^2}$ in the objective function have support in the θ -dependent four-dimensional subspace \mathcal{V}_θ , so $M(\cdot; \theta)$ can be obtained as a POVM on \mathcal{V}_θ . Since the model is a real model, only the real part of the POVM needs to be considered. In addition, since the Fisher information matrix does not become smaller by decomposing the POVM into a rank-1 measurement, it is sufficient to consider only rank-1 measurements. Furthermore, according to Fujiwara [8], the optimal measurement can be achieved with at most a 16-valued measurement. Originally, he stated that $(\dim \mathcal{V}_\theta)^2 + d(d+1)$ is sufficient for the number of measurement outcomes, where d is the dimension of the parameter θ , but $(\dim \mathcal{V}_\theta)^2$ can be replaced by $\frac{1}{2} \dim \mathcal{V}_\theta (\dim \mathcal{V}_\theta + 1)$ since our model is a real model.

We now consider the parametrization of an n -valued real rank-1 measurement on a q -dimensional Hilbert space \mathbb{C}^q . The n -valued real rank-1 measurement is given by real vectors $a_1, a_2, \dots, a_n \in \mathbb{C}^q$ satisfying

$$\sum_{i=1}^n |a_i\rangle\langle a_i| = I_q. \quad (\text{A1})$$

Although a_1, a_2, \dots, a_n must satisfy the above constraint, real rank-1 measurements can be parametrized without any constraint as follows [13].

Equation (A1) can be rewritten as

$$(a_1 \quad a_2 \quad \cdots \quad a_n) \begin{pmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_n^T \end{pmatrix} = I_q. \quad (\text{A2})$$

This means that $V = (a_1 \quad a_2 \quad \cdots \quad a_n)^T$ is an isometry. Then, since the column vectors of V are orthonormal, we obtain

$$U_1 U_2 \cdots U_m V = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{n \times q} \quad (\text{A3})$$

using $m = nq - \frac{1}{2}q(q+1)$ appropriate two-level orthogonal matrices $U_1, U_2, \dots, U_m \in \mathbb{R}^{n \times n}$ (see Sec. 4.5.1 of [15]). From this it follows that

$$V = U_m^T U_{m-1}^T \cdots U_1^T \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{n \times q}. \quad (\text{A4})$$

Since each two-level orthogonal matrix can be specified with a single real parameter, the real rank-1 measurement can be specified with $m = nq - \frac{1}{2}q(q+1)$ unconstrained real parameters. We can then use the algorithm for solving the unconstrained nonlinear programming problem to obtain the optimal measurement $M(\cdot; \theta)$ for the given θ and G by using (23). For two point sources, we obtained the optimal measurements using the BASINHOPPING and POWELL algorithms of the SCIPY package.

Note that the parametrization for a rank-1 measurement with a nonzero imaginary part can also be done in the same way [13].

- [1] M. Tsang, R. Nair, and X.-M. Lu, *Phys. Rev. X* **6**, 031033 (2016).
 [2] M. R. Grace, Z. Dutton, A. Ashok, and S. Guha, *J. Opt. Soc. Am. A* **37**, 1288 (2020).
 [3] M. Parniak, S. Borówka, K. Boroszko, W. Wasilewski, K. Banaszek, and R. Demkowicz-Dobrzański, *Phys. Rev. Lett.* **121**, 250503 (2018).
 [4] F. Bao, H. Choi, V. Aggarwal, and Z. Jacob, *Opt. Lett.* **46**, 3045 (2021).

- [5] Y. Shi and X.-M. Lu, *Commun. Theor. Phys.* **75**, 045102 (2023).
 [6] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (Edizioni della Normale, Pisa, Italy, 2011).
 [7] H. Nagaoka, in *Proceedings of 12th Symposium on Information Theory and Its Applications (SITA)*, Inuyama, Aichi, Japan (1989), pp. 577–582; reprinted in *Asymptotic Theory of Quantum Statistical Inference*, edited by M. Hayashi (World Scientific, Singapore, 2005), pp 125–132.
 [8] A. Fujiwara, *J. Phys. A* **39**, 12489 (2006).

- [9] K. Yamagata, *Int. J. Quantum Inf.* **9**, 1167 (2011).
- [10] X.-M. Lu and X. Wang, *Phys. Rev. Lett.* **126**, 120503 (2021).
- [11] J. Shao and X.-M. Lu, *Phys. Rev. A* **105**, 062416 (2022).
- [12] It is theoretically known that according to Fujiwara [8], 16 values are sufficient, but in fact, according to Yamagata [13], 12 values are sufficient.
- [13] K. Yamagata, Structure of measurement space in quantum estimation theory, master's thesis, Osaka University, 2010.
- [14] A. Sajjad, M. R. Grace, Q. Zhuang, and S. Guha, *Phys. Rev. A* **104**, 022410 (2021).
- [15] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2010).