

Efficient estimation of error bounds for quantum multiparametric imaging with constraints

Alexander Mikhalychev ^{1,*}, Saif Almazrouei,² Svetlana Mikhalycheva ¹, Abdellatif Bouchalkha ²,
Dmitri Mogilevtsev ¹ and Bobomurat Ahmedov ^{3,4}

¹*B.I. Stepanov Institute of Physics, NAS of Belarus, Nezavisimosti ave. 68-2, Minsk 220072, Belarus*

²*Directed Energy Research Centre, Technology Innovation Institute, 9639, Yas Island, Abu Dhabi, United Arab Emirates*

³*Institute of Theoretical Physics, National University of Uzbekistan, Tashkent 100174, Uzbekistan*

⁴*Institute for Advanced Studies, New Uzbekistan University, Movarounnahr str. 1, Tashkent 100000, Uzbekistan*



(Received 19 November 2024; revised 28 December 2024; accepted 4 March 2025; published 21 March 2025)

Advanced superresolution imaging techniques require specific approaches for accurate and consistent estimation of the achievable spatial resolution. Fisher information supplied to the Cramér-Rao bound (CRB) has proved to be a powerful and efficient tool for resolution analysis and optical setups optimization. However, the standard CRB is not applicable to constrained problems violating the unbiasedness condition, while such models are frequently encountered in quantum imaging of complex objects. Complementary to the existing approaches based on modifying CRB, we propose a practical algorithm for approximate construction of a modified Fisher information matrix, which takes the constraints into account and enables accurate estimation of errors for constrained problems by the standard CRB. We demonstrate the efficiency of the proposed technique by applying it to one-, two-, and multiparameter model problems in quantum imaging. The approach provides quantitative explanation of previous results with successful experimental reconstruction of objects with the spatial scale smaller than the theoretical limit predicted by the standard CRB.

DOI: [10.1103/PhysRevA.111.032618](https://doi.org/10.1103/PhysRevA.111.032618)

I. INTRODUCTION

The problem of diffraction limit of resolution in imaging was perceived and quantified more than a century ago [1,2]. The first approaches to resolution quantification (such as Rayleigh criterion) were based on formalizing visual distinguishability of small features. While being intuitive and very fruitful in physics, the approach does not account, for example, for the influence of the shot noise on the image quality. Further, more advanced approaches based on Fourier components transfer were developed [3]. Recent progress in superresolution imaging [4–7] forced the researchers to perceive the collected images (raw datasets) as quite abstract sources of information about the investigated sample rather than just visual replicas of it. Instead of measuring the intensity of the light emitted or scattered by the sample, superresolution techniques are based on the detection of photon coincidences [4,8–11], analysis of time-dependent signals for fluctuating emitters [12–16], or scanning with localized excitation of fluorophores [17–19]. Since the resulting image is typically formed by elaborate processing of the raw information, the approaches to resolution quantification are to be tailored in the way suitable for taking that into account.

A powerful mathematical tool for quantification of the information content of various images (including both raw data and the processing results) is the Fisher information (FI) [20]. It describes sensitivity of the measured signal (or processed image) to variation of the sample parameters of interest. The

Cramér-Rao bound (CRB) [21,22] connects FI with the errors of sample parameters' estimation for optimal analysis of the acquired signal. The approach has been successfully used for resolution analysis in quantum imaging of simple [23–27] and complex [24,28] model objects, optimization of detection protocols [29–34], quantum light sources [28,34], and cumulant order in superresolution optical fluctuation imaging (SOFI) [35,36]. Roughly speaking, resolution of a specific quantum or quantum-inspired imaging technique equals the smallest size of the sample details, such that the input signal (image) provides enough information for their reconstruction [34,35].

Despite strong success in applying classical and quantum FI to theoretical investigation of quantum imaging, direct exploitation of the CRB has important limitations when constraints are imposed on the considered problem. In particular, the resolution experimentally achieved in Ref. [28] for multiparametric quantum imaging was several times better than the theoretical limit derived from CRB without accounting for constraints. A similar effect (reduction of the estimation variance) was previously reported for experiments [27] and numerical simulations [29,37] for the separation of two incoherent point-like light sources, as well as for the resolution of a spectrometer [24].

The reason for those discrepancies is known from statistics [38–41] and consists in the unbiasedness assumption, which is used during derivation of CRB and, generally, becomes invalid when physical constraints are taken into account. For example, for a passive nonemitting object, the absolute values of the transmission amplitude must be in the range from 0 to 1. From physical point of view, constraints introduce additional *a priori* knowledge about the sample, complementary to the

*Contact author: mikhalychev@gmail.com

information provided by the detected signal, and enable reconstruction of smaller features. Mathematically, constraints typically lead to estimation bias [42], while biased estimators can ensure smaller mean square errors than unbiased ones [24,39,41,43,44].

Efficiency of CRB for resolution analysis on the one hand and typicality of constrained problems in practical imaging on the other hand motivate one to search for a way to restore applicability of CRB to such problems. One of the available approaches is based on constructing the modified CRB [38,40,41,43], which explicitly takes the bias into account by incorporating its derivatives over the parameters of interest. Such approach has been successfully used for simple one-parameter models in quantum imaging and metrology: estimation of the separation between two incoherent point sources [27,29,37] and inference of the wavelength by a spectrometer [24]. Minimization of the predicted estimation error over the bias allows us to construct uniform (universal) version of CRB [45–47]. The resulting trade-off between the bias and the variance of the estimate can serve for resolution quantification in biomicroscopy [48–50]. While the approach is fruitful for derivation of general theoretical bounds on estimation errors, its practical application to quantifying information content of particular quantum imaging methods and setups can be cumbersome. Analytical calculation of neither bias gradient for a selected estimator, nor the optimal bias seems feasible for a real multiparametric problem of object reconstruction in quantum imaging [28], thus leading to the necessity of time-consuming numerical Monte Carlo simulations.

In the present contribution we propose an alternative approach, which, instead of rigorous bounds, provides only an approximate assessment of the errors and resolution but can be efficiently applied to multiparametric problems in practice. The essence of the method is to make use of the correspondence between FI and Bayesian probabilities [51,52], to modify the probability distribution according to constraints as it was done previously in our paper on quantum tomography [53], and to map the new probability distribution back to FI. The modification of the probability distribution can be interpreted as taking the prior information encoded in the constraints into account. The resulting updated Fisher information matrix (FIM) already effectively includes the information about the constraints and induced estimation bias and can be put directly into the initial unmodified CRB. We apply the developed approach to multiparametric estimation problems in quantum imaging and show that it is capable of explaining and quantifying the resolution enhancement due to constraints experimentally observed in Ref. [28]. It is worth noting that the influence of the estimation bias on the predicted errors can be much more pronounced in quantum imaging based on measurement and analysis of photon coincidences (higher-order field correlation functions) relatively to traditional intensity-based microscopy, because FI tends to become singular if the investigated object has dark (nontransmitting) areas. We pay special attention to that issue and show that the case is still covered by the proposed approach.

The rest of the paper is organized as follows: First, we briefly introduce the main concepts from estimation and probability theory useful for further derivation and analysis of

the approach. Section III is devoted to constrained estimation problems and the core procedures of our paper: regularization and correction of FIM. In Sec. IV, we assess the efficiency and accuracy of the proposed approach by applying it to several model problems, ranging from a very simple one to practical multiparametric imaging.

II. FISHER INFORMATION, PROBABILITIES, AND RESOLUTION

A. Fisher information

The general purpose of measurements, including those used in imaging, is to estimate physical characteristics of the investigated system from the collected data [54]. For that purpose, one typically needs to construct the theoretical model of the signal $\mathbf{S}(\boldsymbol{\theta})$, where the vector $\boldsymbol{\theta} = \{\theta_\mu\}$ describes the parameters to be inferred from the experiment. The actually collected signal

$$\mathbf{Y} = \mathbf{S}(\boldsymbol{\theta}) + \boldsymbol{\varepsilon} \quad (1)$$

inevitably contains noise contribution $\boldsymbol{\varepsilon}$ and, therefore, represents a particular realization of random variable(s). Statistical properties of the noise can be formalized by introduction of the likelihood $L(\mathbf{Y}|\mathbf{S})$.

The ability to reconstruct the parameters of interest from the measured data and the resulting inference error are determined by the *sensitivity* of the signal to the investigated parameters. The sensitivity can be quantified by the *score* function [20]:

$$s_\mu(\boldsymbol{\theta}) = \frac{\partial l(\mathbf{Y}|\mathbf{S}(\boldsymbol{\theta}))}{\partial \theta_\mu}, \quad (2)$$

where $l(\mathbf{Y}|\mathbf{S}(\boldsymbol{\theta})) = \ln L(\mathbf{Y}|\mathbf{S}(\boldsymbol{\theta}))$ is the log-likelihood.

The covariance matrix of the score represents the Fisher information matrix (FIM) [20]:

$$F_{\mu\nu} = \text{Cov}(s_\mu(\boldsymbol{\theta}), s_\nu(\boldsymbol{\theta})) = \mathbb{E} \left(\frac{\partial l(\mathbf{Y}|\mathbf{S}(\boldsymbol{\theta}))}{\partial \theta_\mu} \frac{\partial l(\mathbf{Y}|\mathbf{S}(\boldsymbol{\theta}))}{\partial \theta_\nu} \right), \quad (3)$$

where the symbol $\mathbb{E}$ denotes mathematical expectation over the distribution of the random signal $\mathbf{Y}$. Roughly speaking, larger values of the FIM elements correspond to more informative measurements. If the signal $\mathbf{Y}$ takes discrete values, Eq. (3) can be rewritten in the following way convenient for direct calculation:

$$F_{\mu\nu} = \sum_{\mathbf{Y}} \frac{1}{L(\mathbf{Y}|\mathbf{S}(\boldsymbol{\theta}))} \frac{\partial L(\mathbf{Y}|\mathbf{S}(\boldsymbol{\theta}))}{\partial \theta_\mu} \frac{\partial L(\mathbf{Y}|\mathbf{S}(\boldsymbol{\theta}))}{\partial \theta_\nu}, \quad (4)$$

where the summation is performed over all possible realizations of the signal $\mathbf{Y}$.

The significance of FIM is enhanced by its use in the Cramér-Rao bound (CRB) [21,22]. If the parameters inference algorithm yields an *unbiased* estimate $\hat{\boldsymbol{\theta}}$, i.e., its mathematical expectation equals the true value of the parameters: $\mathbb{E}(\hat{\boldsymbol{\theta}}) = \boldsymbol{\theta}$, the covariance matrix of the estimated parameters is bounded from below by the inverse of the FIM:

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

The matrix inequality $A \geq B$ means that the matrix $A - B$ is positive semidefinite.

Equation (3) can be also rewritten as the expectation value of the log-likelihood curvature (see, e.g., Ref. [54]):

$$F_{\mu\nu} = -E \left(\frac{\partial^2 l(\mathbf{Y}|\mathbf{S}(\boldsymbol{\theta}))}{\partial \theta_\mu \partial \theta_\nu} \right). \quad (6)$$

In the limit of high signal-to-noise ratio (SNR) typical for strong signals, that representation implies that the FIM represents the Hessian of the maximum-likelihood estimate [MLE, formulated as finding $\hat{\boldsymbol{\theta}}$ maximizing the log-likelihood $l(\mathbf{Y}, \mathbf{S}(\hat{\boldsymbol{\theta}}))$ for the specific collected dataset $\mathbf{Y}$] and characterizes scattering of the estimation results for multiple repetitions of the experiment [52].

Alternatively, one can adhere to the Bayesian approach [55] and consider the posterior probability $p(\boldsymbol{\theta}|\mathbf{Y})$ of the parameter vector $\boldsymbol{\theta}$ conditioned by the measured dataset $\mathbf{Y}$. For a uniform prior probability distribution (absence of useful information about the investigated parameters before the measurement) and high SNR, the FIM F describes the shape of the posterior probability distribution near the point $\hat{\boldsymbol{\theta}}_{ML}$ of the likelihood maximum [51,52]:

$$p(\boldsymbol{\theta}|\mathbf{Y}) \propto L(\mathbf{Y}|\mathbf{S}(\boldsymbol{\theta})) \propto \exp \left[-\frac{1}{2}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{ML})^T F (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{ML}) \right]. \quad (7)$$

The application of FI in terms of CRB is directly related to accuracy and resolution in imaging, while its Bayesian interpretation is helpful for our development of the constraints treatment approach. Some insights into relations between likelihood, distribution of point estimation results, and FI can also be gained from the geometrical considerations presented in Appendix A.

It is also worth mentioning *quantum* Fisher information (QFI) [56–58], which generalizes the concept of classical FI and corresponds to the maximum of classical FI over all physically valid measurements. QFI represents a powerful tool for theoretical analysis of relatively simple systems [25,26,29,30,59–62] and can be applied even to practical multiparametric problems [63–65], but may require cumbersome calculations in that case. Due to the same structure and similar interpretation of classical FI and QFI, our approach remains applicable to QFI as well. However, in the current work we focus on practical applications considering quantum imaging with a fixed (not optimized) type of measurement and limit our analysis to classical FI.

B. Fisher information in quantum imaging

In quantum imaging, the measured signal $\mathbf{S}(\boldsymbol{\theta})$ is typically represented by field correlation functions [4]:

$$\begin{aligned} G^{(n)}(\mathbf{r}_1, t_1; \dots; \mathbf{r}_n, t_n) \\ = \langle E^{(-)}(\mathbf{r}_1, t_1) \dots E^{(-)}(\mathbf{r}_n, t_n) E^{(+)}(\mathbf{r}_n, t_n) \dots E^{(+)}(\mathbf{r}_1, t_1) \rangle, \end{aligned} \quad (8)$$

where $E^{(\pm)}(\mathbf{r}, t)$ are positive-frequency (negative-frequency) field operators. Namely, the signal $\mathbf{S}(\boldsymbol{\theta}) = (S_1(\boldsymbol{\theta}), S_2(\boldsymbol{\theta}), \dots)$ is composed of the components

$$S_i(\boldsymbol{\theta}) \propto G^{(n)}(\mathbf{r}_1^{(i)}, t_1^{(i)}; \dots; \mathbf{r}_n^{(i)}, t_n^{(i)}) \quad (9)$$

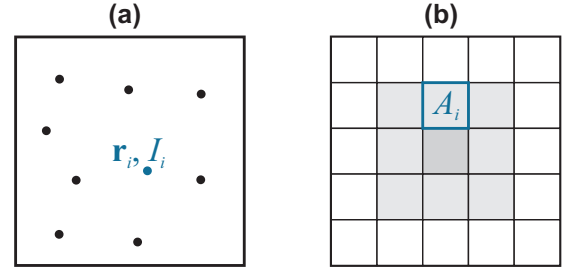


FIG. 1. Problem parametrization in quantum imaging: (a) by positions $\mathbf{r}_i$ and brightnesses I_i of point emitters; (b) by transmission amplitudes A_i of pixels.

defined by specific arguments of the correlation function.

From a practical point of view, the n th-order correlation function $G^{(n)}$ describes the rate of n -photon coincidence events for specified detection positions and time delays. The signal component $S_i(\boldsymbol{\theta})$ describes the mean number of the corresponding coincidence events, while the component Y_i of a random signal realization $\mathbf{Y} = (Y_1, Y_2, \dots)$ equals the realized (integer) number of such events.

Commonly, the detection events can be treated as independent ones and the likelihood function corresponds to independent random variables Y_i obeying Poisson distribution with the mean values $S_i(\boldsymbol{\theta})$:

$$L(\mathbf{Y}|\mathbf{S}(\boldsymbol{\theta})) = \prod_i L(Y_i|S_i(\boldsymbol{\theta})), \quad (10)$$

where

$$L(Y|S(\boldsymbol{\theta})) = \frac{[S(\boldsymbol{\theta})]^Y}{Y!} e^{-S(\boldsymbol{\theta})} \quad (11)$$

is the likelihood function for a single-variate Poisson distribution.

In that specific but commonly encountered case the expression (4) can be simplified and takes the following form (see Refs. [30,33,34,66] and Appendix B):

$$F_{\mu\nu} = \sum_i \frac{1}{S_i(\boldsymbol{\theta})} \frac{\partial S_i(\boldsymbol{\theta})}{\partial \theta_\mu} \frac{\partial S_i(\boldsymbol{\theta})}{\partial \theta_\nu}. \quad (12)$$

It is worth noting that the structure of Eqs. (4) and (12) for the FIM elements $F_{\mu\nu}$ is very similar, while their meaning is quite different. The general expression (4) is constructed from probabilities of all possible (mutually exclusive) measurement outcomes $\mathbf{Y}$, and the normalization $\sum_{\mathbf{Y}} L(\mathbf{Y}|\mathbf{S}(\boldsymbol{\theta})) = 1$ holds. In Eq. (12), the summation index i enumerates different signal components measured simultaneously in one experiment. The quantities $S_i(\boldsymbol{\theta})$ represent the expectation values of the events numbers and are not subjected to any specific kind of normalization.

The set of inferred parameters $\boldsymbol{\theta}$ is determined by the model of the investigated object. In quantum imaging, the model is typically defined either as a set of discrete point emitters, parametrized by their positions $\mathbf{r}_i$ and brightness I_i [Fig. 1(a)] [23–27,29–33,35,64], or as a pixelized mask characterized by the transmission amplitudes of the pixels A_i [Fig. 1(b)] [8–10,28,34,67,68].

The FIM itself provides valuable information about the problem structure—parameters correlations and inaccuracies

(see, e.g., Refs. [26,28,65,66]), since the variance and covariances of the parameters are related to the inverse of the FIM by the CRB. For various optimization purposes, as well as for quantification of optical resolution, it is also useful to construct a scalar real-valued characteristic providing an aggregate description of the accuracy in a multiparametric problem. There exist several common choices of such a scalar quantity [69–72]: the volume of the joint confidence region defined by CRB (D optimality—minimization of the determinant of the inverse of FIM: $\min \det F^{-1}$); the major axis of the joint confidence region [E optimality—minimization of the maximal eigenvalue of the inverse of FIM: $\min \lambda_{\max}(F^{-1})$]; the perimeter of the enclosing box of the joint confidence region (A optimality—minimization of the trace of the inverse of FIM: $\min \text{Tr } F^{-1}$). The latter choice has already been applied to quantum imaging [28,34,35] by using the lower bound of total variance

$$\Delta^2 \equiv \text{Tr } F^{-1} \leq \sum_{\mu} \text{Var } \theta_{\mu} \quad (13)$$

as a measure of diffraction-caused inaccuracy.

FI represents a powerful and useful tool for resolution quantification in imaging. If inference of the positions (separation) of point emitters is considered, CRB directly provides a resolution measure: the predicted (minimal) inaccuracy of the position reconstruction quantifies the resolving power of the microscope [23–25,27,29,31,32]. For analysis of pixelized images, one can start from the notion that a feature is resolved if the acquired image (dataset) contains enough information for inference of the feature parameters with a certain requested accuracy [23,24,34,35]. In particular, one can define the spatial resolution as the minimal spatial scale of the problem (or feature size), for which the quantity defined by Eq. (13) is lower than a certain threshold value determined by further exploitation purposes of the imaging results [34,35]: $\Delta^2 \leq \Delta_{\text{threshold}}^2$.

The close connection of FI to achievable resolution allows us to use it for optimization of the imaging setup. The increase of FI, including the removal of “Rayleigh’s curse” in simple cases, indicated apparent advantages of detection protocols based on spatial mode demultiplexing [29,30,33,73,74], modification of the point-spread function (PSF) [31,32], or probabilistic transformation of an entangled light state [34]. Analysis of the FI proved to be useful in optimization of quantum light sources [28,34] and selection of the optimal cumulant order in superresolution optical fluctuation imaging (SOFI) [35,36].

In both object parametrizations discussed above, some basic constraints can limit the range of meaningful values for the parameters. For example, brightness of any emitter or the absolute value of the separation between two emitters should be non-negative ($I_i \geq 0$, $\Delta r \equiv |\mathbf{r}_1 - \mathbf{r}_2| \geq 0$). For a passive transmitting object, the absolute value of a pixel transmission amplitude cannot exceed 1: $|A_i| \leq 1$. In a commonly encountered case of an object with a real-valued transmission amplitude (see, e.g., Refs. [8–11,75–77]), the constraint takes the form

$$0 \leq A_i \leq 1. \quad (14)$$

A remarkable illustration of the constraints influence on the practically achievable spatial resolution was provided by the experiment reported in Ref. [28]. A number of pixelized objects (one-dimensional sets of slits and two-dimensional masks) were illuminated by biphotons and pseudothermal light and the second- and third-order correlation functions were recorded. Quite unexpectedly, actually achieved resolution was several times higher than the theoretical limit, predicted without taking the constraints into account. Namely, the resolution limit (minimal pixel size $d_{\min}$ defined from FI according to the predicted inference accuracy for the pixels’ transmission amplitudes [34]) was approximately a half of the Rayleigh limit d_R for the optical system used: $d_{\min} \approx (0.4 \div 0.5)d_R$. The experimental datasets allowed successful object features reconstruction for the parametrization with a noticeably smaller pixel size: $d \approx (0.25 \div 0.28)d_R < d_{\min}$.

Those counterintuitive results were explained by a specific feature (still, quite common for model experiments) of the imaged objects: the actual transmission amplitude of their pixels took binary values $A_i = 0$ or 1. Therefore, the constraints (14) were active (i.e., influenced the reconstruction process) and provided additional information about the sample. It is worth noting that the information about such binary type of the objects was not implied *a priori* and was not used to assist the reconstruction algorithm. The connection between the resolution enhancement and the constraints was verified by Monte Carlo simulations and numerical estimation of the inference accuracy. The achievable resolution (minimal pixel size for reliable reconstruction of the transmission amplitudes) was approximately twice as good for black-and-white model objects (with active constraints) than for the corresponding gray objects (when the constraints are inactive because the values are far from 0 and 1). However, a way to quantitative theoretical FI-based prediction of such resolution improvement was not considered in Ref. [28]. In the current contribution, we aim to fill that gap.

III. CONSTRAINED ESTIMATION AND FISHER INFORMATION

A. Influence of constraints on parameter estimation

When the estimation formalism is applied to real-world tasks, the true values of the parameter may be constrained by a specific physical domain: $\theta \in \Omega$. In that case, it is reasonable to impose the constraints on the estimated values as well: $\hat{\theta} \in \Omega$. If the constraints are active (i.e., the true values of the parameters are within the estimation inaccuracy from the boundaries of the region Ω), the relation between FI and the estimation errors becomes more complicated than prescribed by CRB [Eq. (5)]. The constrained estimation typically violates unbiasedness condition [42] required for validity of CRB.

In Bayesian treatment of the problem, the constraints induce nontrivial prior probability: $p_0(\theta) = 1/|\Omega|$ for $\theta \in \Omega$ and 0 otherwise. Since prior probabilities are not taken into account by FI, it does not describe the posterior probability by Eq. (7) correctly for constraint problems.

The influence of the constraints on the distribution of estimated parameter values in the one-parameter case is shown in

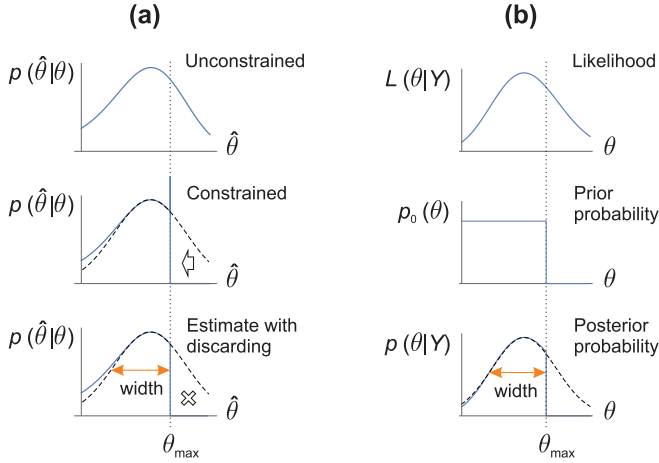


FIG. 2. Influence of a constraint $\theta \leq \theta_{\max}$ on the estimate distribution: (a) distribution of the estimated parameter value $\hat{\theta}$ for its true value θ (from top to bottom—unconstrained estimate, constrained estimate, and formally introduced estimate with discarded nonphysical values); (b) Bayesian posterior probability for the parameter θ for a given signal value Y (from top to bottom—unconstrained likelihood, prior probability distribution describing the constraint, and the resulting posterior probability). Dashed lines show Gaussian fit of the unconstrained distributions. The resulting width after the constraint application is smaller than the one predicted by FI.

Fig. 2 (provided just as an illustration without connection to any specific physical model). It is worth noting that Bayesian “cutting” of the posterior probability distribution by setting zero prior probability for nonphysical values can be formally reproduced in the distribution of point estimation results by performing unconstrained estimation and subsequent discarding of nonphysical values $\hat{\theta}_{\text{unconstr.}} \notin \Omega$. In the high-SNR case, the resulting distributions will have practically the same width. While that estimate with the values discarding does not follow typical practical procedures of parameter inference, it is useful for getting some insights for the development of the FIM modification approach.

In the one-parameter case, the constraint-induced decrease of the parameter reconstruction error (the width of the estimate distribution) is quite limited. Taking into account that the true value $\theta \in \Omega$ belongs to the physical domain and assuming that the distribution $p(\hat{\theta}|\theta)$ is symmetric around the true value θ , one can easily see that the width is minimal for θ having a boundary value ($\theta = \theta_{\max}$ for the case shown in Fig. 2) and reaches 1/2 of the initial width of the unconstrained distribution. The effect can be much more pronounced for a multiparameter problem [28]. Figure 3 illustrates the statement for a model two-parameter problem, where the error is reduced by the factor of approximately 6.

B. Effective Fisher information matrix for constrained estimation

For unconstrained problems, FIM can be used for efficient prediction of experimental inference inaccuracies via CRB [Eq. (5)]. In high-SNR regime, FIM also represents the quadratic-form kernel for a Gaussian approximation of the estimated parameters distribution [Eq. (7)]. Mathematically, the two meanings of FIM are ensured by the relation between the quadratic form and the covariance matrix of a multivariate normal distribution:

$$p(\hat{\theta}) \propto \exp \left[-\frac{1}{2} \Delta \hat{\theta}^T F \Delta \hat{\theta} \right] \Rightarrow \text{Cov}(\hat{\theta}, \hat{\theta}) = F^{-1}, \quad (15)$$

where $\Delta \hat{\theta} = \hat{\theta} - E(\hat{\theta})$.

When the constraints are applied to the problem, the basic CRB becomes invalid (due to bias of the estimate) while use of its modified versions [38,40,41,43] is typically computationally expensive. On the other hand, transformation of probability distributions $p(\hat{\theta}) \mapsto p_{\text{constr.}}(\hat{\theta})$ caused by application of the constraints is quite straightforward, as discussed above.

From practical point of view, one is interested in errors of the parameters estimation, namely, in the covariance matrix $\text{Cov}_{\text{constr.}}(\hat{\theta}, \hat{\theta})$ for the constrained distribution. If the distribution $p_{\text{constr.}}(\hat{\theta})$ can be approximated by a multivariate normal distribution $\tilde{p}(\hat{\theta})$, the relation similar to Eq. (15) will

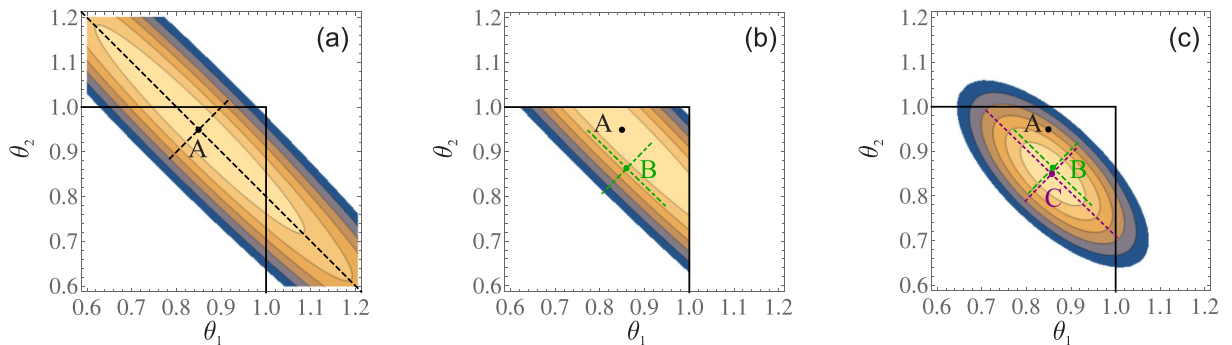


FIG. 3. Influence of constraints $\theta_1 \leq 1$ and $\theta_2 \leq 1$ on the distribution of parameters' estimates for a model two-parameter case: (a) probability distribution for an unconstrained problem, constrained distribution corresponding to posterior probability for Bayesian treatment or (b) to the formal distribution of the estimated parameters values after discarding nonphysical results, and (c) a Gaussian approximation of the constrained distribution according to Sec. III B. Capital letters A, B, and C indicate the mean values for the probability distributions shown in panels (a), (b), and (c), respectively. Dashed lines correspond to the eigenvectors of the covariance matrix and indicate the corresponding widths of the distributions. The constrained estimate is biased (the point B is shifted from the point A indicating the true values of the parameters θ_1 and θ_2). Errors of the parameters' estimation become approximately 6 times smaller when the physical constraints are taken into account.

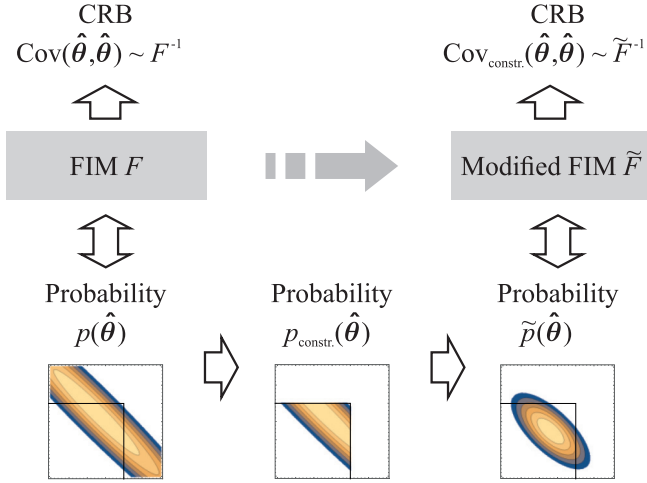


FIG. 4. Scheme of taking constraints into account by construction of modified FIM. The basic FIM F corresponds to the unconstrained problem and, in the high-SNR regime, represents the quadratic-form kernel for the probability distribution $p(\hat{\theta})$ of unconstrained parameter estimates. Then, the constraints are applied to the probability distribution and the result is approximated by a multivariate normal distribution $\tilde{p}(\hat{\theta})$. The kernel $\tilde{F}$ of that distribution's quadratic form can be interpreted as the modified FIM since it is related to the covariance of the constrained parameter estimates through the usual CRB.

hold

$$\tilde{p}(\hat{\theta}) \propto \exp \left[-\frac{1}{2} \Delta \hat{\theta}^T \tilde{F} \Delta \hat{\theta} \right] \Rightarrow \text{Cov}_{\text{constr.}}(\hat{\theta}, \hat{\theta}) \approx \tilde{F}^{-1}, \quad (16)$$

where $\Delta \hat{\theta} = \hat{\theta} - \text{E}_{\text{constr.}}(\hat{\theta})$. The quadratic form $\tilde{F}$ of the probability distribution $\tilde{p}(\hat{\theta})$ can be interpreted as the FIM for the constrained problem and quantifies the error of constrained-parameter estimation via the usual (unmodified) CRB defined by Eq. (5). The transition from the initial FIM F to the modified FIM $\tilde{F}$ is shown schematically in Fig. 4. Since the proposed procedure of FIM modification preserves the form of CRB for constrained problems, further analysis methods based on the CRB (e.g., quantification of resolution [34] or optimization of the measurement setup [28]) remain valid.

Computationally efficient procedure of transforming $p(\hat{\theta})$ into $\tilde{p}(\hat{\theta})$ has been proposed in Ref. [53] in the connection with adaptive quantum tomography with Bayesian update of the knowledge about the analyzed quantum state. The constraints are assumed to be linear and have the form $\mathbf{a}_j^T \hat{\theta} \leq b_j$. The idea of the approach consists in iterative shrinking of the probability distribution according to the current most severely violated constraint.

Let the probability distribution obtained after the i th iteration be described as

$$p^{(i)}(\hat{\theta}) \propto \exp \left[-\frac{1}{2} \Delta \hat{\theta}^T F^{(i)} \Delta \hat{\theta} \right], \quad \Delta \hat{\theta} = \hat{\theta} - \theta^{(i)}. \quad (17)$$

Then, the $(i+1)$ st iteration transforms $p^{(i)}(\hat{\theta})$ into $p^{(i+1)}(\hat{\theta})$ and includes the following steps [53] (see Appendix C for a visual illustration and additional expressions):

(1) Construct a matrix $T = T^T = (F^{(i)})^{1/2}$ and apply a linear coordinate transformation $\hat{\theta} \mapsto \theta' = T(\hat{\theta} - \theta^{(i)})$, such that the transformed distribution has zero mean $\theta^{(i)'} = 0$ and

the quadratic-form kernel represented by the unit matrix: $F^{(i)'} = T^{-1} F^{(i)} T^{-1} = 1$. The transformation is nondegenerate if $F^{(i)}$ is not singular.

(2) Recalculate all constraints to the new coordinates as $\mathbf{a}_j \mapsto \mathbf{a}_j' = \mathbf{a}_j T^{-1}$, $b_j \mapsto b_j' = b_j - \mathbf{a}_j'^T \theta^{(i)}$.

(3) Find the index of the most severely violated constraint: $j = \text{argmin}_j (b_j' / |\mathbf{a}_j'|)$.

(4) Shrink the distribution in the direction $\mathbf{d} = \mathbf{a}_j' / |\mathbf{a}_j'|$ corresponding to the found constraint: $F^{(i)'} \mapsto F^{(i)''} + \xi \mathbf{d} \mathbf{d}^T$, $\theta^{(i)'} \mapsto \theta^{(i)''} = -\delta \mathbf{d}$. Explicit expressions for δ and ξ are listed in Appendix C.

(5) Apply the inverse transform to get the resulting probability distribution after the current iteration: $F^{(i+1)} = T F^{(i)''} T = F^{(i)} + \xi T \mathbf{d} \mathbf{d}^T T$, $\theta^{(i+1)} = \theta^{(i)} - \delta T^{-1} \mathbf{d}$.

C. Ill-defined Fisher information matrix for dark objects

Both the application of the CRB (5) to the unmodified FIM and the implementation of the proposed iterative correction require the FIM to be nonsingular in order to get meaningful results. However, the multiparameter estimation problem in quantum imaging can yield an ill-defined FIM if the object contains a dark (nontransparent) region.

In particular, for a uniform object with the transmission amplitude A , the n th-order correlation function $G^{(n)}$ defined by Eq. (8) is polynomial relative to A : $G^{(n)} \propto A^{2n}$. Then the FI calculated for the problem parameter A scales as $F \propto A^{2(n-1)}$. For conventional intensity-based imaging ($n = 1$), the singularity is absent: $2(n-1) = 0$ and F remains nonzero for $A = 0$. However, for detection of correlation functions with $n \geq 2$, one has $F = 0$ for $A = 0$ (quite similarly to “Rayleigh’s curse” for inference of two point light sources separation [29]).

As another illustrative example of quantum imaging, one can consider illumination of the object by ideally correlated biphotons. The second-order correlation function can be represented as [4]

$$G^{(2)}(\mathbf{r}_1, \mathbf{r}_2) \propto |\Psi(\mathbf{r}_1, \mathbf{r}_2)|^2 \equiv \left| \int d^2 \mathbf{s} A^2(\mathbf{s}) h(\mathbf{s}, \mathbf{r}_1) h(\mathbf{s}, \mathbf{r}_2) \right|^2, \quad (18)$$

where $A(\mathbf{s})$ is the transmission amplitude of the object at the point with the transverse position $\mathbf{s}$, and $h(\mathbf{s}, \mathbf{r})$ is the PSF (Green’s function) of the optical system.

For a pixelized object, an array detector, and a general case of nonideally correlated biphotons, the second-order correlation between the i th and j th pixels of the detector can be expressed as [28]

$$G_{ij}^{(2)} \propto |\Psi_{ij}|^2 = \left| \sum_{m,l} D_{ml}^{(ij)} A_m A_l \right|^2, \quad (19)$$

with certain coefficients $D_{ml}^{(ij)}$ (real valued for simple models [28]). If biphotons are ideally correlated, the additional relation $D_{ml}^{(ij)} = 0$ holds for $m \neq l$. According to Eq. (12), diagonal elements of FIM are proportional to

$$F_{mm} \propto A_m^2 \sum_{i,j} (D_{mm}^{(ij)})^2 \propto A_m^2. \quad (20)$$

for the considered parametrization. Therefore, the adverse scaling of FIM for $A_m \rightarrow 0$ remains the same as for the toy example of a uniform object.

D. Regularization of ill-defined Fisher information matrix

An insight to the reasons of FIM singularity and inapplicability of CRB for quantum imaging of dark objects can be gained from Eqs. (6) and (7). Roughly speaking, the local properties of the log-likelihood, characterized by its second derivative over the parameters, are used to estimate the distribution shape and to quantify its width. However, such local properties can be nonrepresentative if the distribution shape is far from the assumed one.

To assist our intuition, it is worth considering a simple problem resembling the discussed behavior of FI. Let us suppose we are given a profile $y(x)$ with the maximum at $x = 0$ and need to estimate the width of that peak. If the peak has Gaussian shape [$y(x) \propto e^{-x^2/(2\sigma^2)}$], the half width σ can be estimated by calculating the second derivative of the profile logarithm:

$$\sigma = \frac{1}{\sqrt{|Y^{(2)}(0)|}}, \quad Y^{(2)}(x) \equiv \frac{d^2}{dx^2}(\ln y). \quad (21)$$

Now, let us consider the profiles $y_1 = e^{-\max(0, (|x|-x_0))^2/(2\sigma^2)}$ and $y_2 = e^{-|x|^k/(2\sigma^k)}$, $k > 2$ [Fig. 5(a)]. In both cases, we have $Y^{(2)}(0) = 0$ and Eq. (21) is inapplicable for estimation of σ , while the actual width remains finite. To get meaningful width estimates, one can probe the slope of the peak by calculating $Y^{(2)}(x')$ for shifted positions $x' \neq 0$ and construct the quantities

$$\Delta(x') = |x'| + \sigma(x') \equiv |x'| + |Y^{(2)}(x')|^{-1/2}. \quad (22)$$

Geometrically, we fit different parts of the peak slope by a Gaussian-shape peak and estimate the resulting width [Fig. 5(b)]. For the profile $y_1(x)$, the value

$$\sigma_{\min} = \min_{x'} \Delta(x') \quad (23)$$

will be exactly the half width of the peak [Fig. 5(b)]. For the function $y_2(x)$, Eq. (23) provides an approximate result, which, however, is still quite accurate [see Fig. 5(c) and Appendix D].

One should treat the discussed illustrative example with caution: the toy example deals with a continuous dependence $y(x)$, while the actually detected signal $\mathbf{Y}$ in quantum imaging typically represents a vector of discrete integer-valued random variables. Moreover, it is the discreteness of the detected signal that leads to singularity of FIM for a dark object (Appendix E). Still, the gained insights can be used for regularization of FIM.

To design the regularization procedure for FIM, we start from considering the one-parameter case. When the error Δ of estimating the parameter θ is analyzed on the base of CRB, FI $F(\theta)$ (which, in the case of a single parameter, is a scalar— 1×1 matrix) plays the role similar to the quantity $Y^{(2)}(x)$, introduced in Eq. (21):

$$\Delta \sim \frac{1}{\sqrt{F(\theta)}}. \quad (24)$$

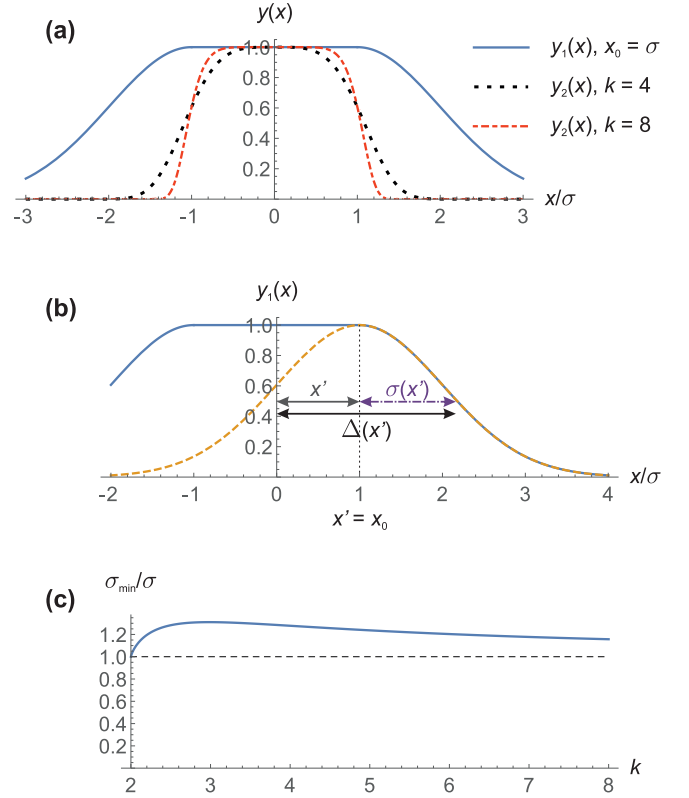


FIG. 5. Width estimation by the second derivative. (a) Model profiles defined in the text (blue solid line— $y_1(x)$ with $x_0 = \sigma$; black dotted line— $y_2(x)$ with $k = 4$; red dashed line— $y_2(x)$ with $k = 8$). (b) Width estimation by “scanning” the profile by a Gaussian-shaped peak: total half width $\Delta(x')$ includes the half width $\sigma(x')$ estimated at the shifted position x' and the displacement x' . (c) Half-width estimate $\sigma_{\min}$ for $y_2(x)$ according to the proposed approach normalized by the true half width σ of the profile (at the height $e^{-1/2}$).

That analogy is also supported by definition of FIM according to Eq. (6).

In the case of a dark object ($F \approx 0$), Eq. (24) strongly overestimates the error, since the second derivative is not representative enough for the probability distribution shape. Following the idea of Eqs. (22) and (23), one can define

$$\Delta(\theta') = |\theta' - \theta| + F^{-1/2}(\theta') \quad (25)$$

and estimate the inaccuracy as

$$\Delta_{\min} = \min_{\theta'} \Delta(\theta'). \quad (26)$$

The approach is illustrated in Appendix A.

Now, we aim at constructing a modified FI $\tilde{F}(\theta)$ such that the initial form of the relation $\Delta = \Delta_{\min} \sim [\tilde{F}(\theta)]^{-1/2}$ holds. This requirement can be fulfilled by defining

$$\tilde{F}(\theta) = \frac{1}{\Delta_{\min}^2} = \max_{\theta'} \frac{F(\theta')}{(1 + |\theta' - \theta| \sqrt{F(\theta')})^2}. \quad (27)$$

In a general multiparameter case, one can apply the discussed regularization along main axes of FIM independently.

If FIM $F(\theta)$ is decomposed as

$$F(\theta) = \sum_i \lambda_i \mathbf{v}_i \mathbf{v}_i^T, \quad (28)$$

where λ_i and $\mathbf{v}_i$ are eigenvalues and eigenvectors, respectively, and $\mathbf{v}_i^T \mathbf{v}_j = \delta_{ij}$, one can consider the shifts $\theta' = \theta + \mathbf{v}_i \delta\theta$ for each separate direction $\mathbf{v}_i$. The direct analog of the scalar FI $F(\theta')$ for the one-parameter case is the projection of FIM for the shifted parameters at the considered direction:

$$\lambda_i(\delta\theta) = \mathbf{v}_i^T F(\theta + \mathbf{v}_i \delta\theta) \mathbf{v}_i. \quad (29)$$

Therefore, the generalization of Eq. (27) to the multiparameter case is

$$\tilde{F}(\theta) = \sum_i \tilde{\lambda}_i \mathbf{v}_i \mathbf{v}_i^T, \quad (30)$$

where

$$\tilde{\lambda}_i = \max_{\delta\theta} \frac{\lambda_i(\delta\theta)}{(1 + |\delta\theta| \sqrt{\lambda_i(\delta\theta)})^2}. \quad (31)$$

It is worth noting that we neither provide accurate proof of the proposed regularization procedure nor claim that the constructed modified FIM $\tilde{F}(\theta)$ ensures a rigorous bound via CRB for the estimation error. Rather, the approach is expected to be a useful practical tool for approximate assessment of the errors and resolution for multiparametric problems in quantum imaging. The next section illustrates the approach with several examples, starting from a simple one-parameter toy example and covering resolution estimation in practical multiparametric quantum imaging.

IV. APPLICATION OF THE APPROACH TO QUANTUM IMAGING

A. One-parameter model

To illustrate the proposed procedures of the FIM correction for constrained problems in quantum imaging and assess the accuracy of the approach, we start from a simple toy example of a uniform object with the transmission amplitude A , $0 \leq A \leq 1$, representing the only parameter of interest. Let us assume that the light source emits groups of n time-correlated photons and n -photon coincidence events are detected. For $n = 2$, such process corresponds to emission of biphotons by spontaneous parametric down-conversion (SPDC) in a nonlinear crystal. Let the number of the emitted photon groups per an experiment have the mean value N and be described by Poisson distribution. Then, the detected signal Y has the expectation value

$$S(A) = N\eta^n A^{2n}, \quad (32)$$

where η is the light collection efficiency.

The likelihood for the signal Y (taking integer values) corresponds to Poisson distribution (11) with the parameters vector θ containing the only component A . Equation (12) for the FIM is applicable in the considered case and yields the result

$$F(A) = 4n^2 N \eta^n A^{2(n-1)}. \quad (33)$$

It is worth noting that the considered toy problem can be parametrized in different ways. For example, one can choose

the probability of photon transmission $T = A^2$ as the parameter of interest instead of the transmission amplitude A and get a different expression for the FI (Appendix F). However, such parametrization would be quite unnatural for a more general multiparameter case with light interference effects: for example, the transmission amplitude directly enters the expression (19) for imaging with nonideal spatial correlations of biphotons. The main goal of the current section is the illustration of the developed approach by a simple example and preparation for analysis of practical multiparametric problems in Sec. IV C. For that reason, we follow the general course here and use A as the parameter of interest, rather than search for the optimal parametrization of the specific toy problem. Further discussion of parametrization influence on FI is provided in Appendix F.

As references for evaluation of the proposed approach, we consider the following two estimates of the transmission amplitude A :

(1) Constrained MLE:

$$\begin{aligned} \hat{A}_{\text{MLE}}(Y) &= \operatorname{argmax}_{0 \leq A' \leq 1} L(Y|S(A')) = \min(1, S^{-1}(Y)) \\ &= \min(1, Y^{1/(2n)} N^{-1/(2n)} \eta^{-1/2}). \end{aligned} \quad (34)$$

(2) Mean over posterior probability distribution according to Bayes' formula:

$$\hat{A}_{\text{B}}(Y) = \frac{\int_0^1 dA' L(Y|S(A')) A'}{\int_0^1 dA' L(Y|S(A'))}. \quad (35)$$

It is worth noting that the latter also represents the optimal biased estimate (following the spirit of Refs. [43,44,47]), minimizing the mean squared error (MSE) averaged over the range of the true values $A \in [0, 1]$ (Appendix G 1).

The bias of the estimates $\hat{A}_{\text{MLE}}$ and $\hat{A}_{\text{B}}$ is shown in Fig. 6(a) for $N = 200$, $\eta = 0.7$, $n = 2$. One can see that the estimates are, indeed, biased near the boundary values $A = 0$ and $A = 1$ (in the regions I and III, respectively).

Figure 6(b) shows the dependence of the estimation error

$$\Delta = \sqrt{\text{MSE}} = \sqrt{\mathbb{E}[(\hat{A} - A)^2]} \quad (36)$$

on the true value A of the transmission amplitude for the two estimators (blue and orange lines correspond to MLE and Bayesian estimators respectively). For each estimator, the MSE is calculated directly from the probability distribution (solid colored lines) and from modified CRB taking the bias into account [43] (dashed colored lines). The details of the calculations are described in Appendix G 2.

One can notice that, in the region I, the error strongly depends on the estimator used; changing the bias, one can “redistribute” the inaccuracy over the range of the parameter A [41]. MLE is heavily biased to zero: it maps zero signal $Y = 0$ to $\hat{A}_{\text{MLE}}(0) = 0$, while such an outcome remains the most probable one for the whole range $A \in [0, A_0]$ with $A_0 = N^{-1/(2n)} \eta^{-1/2} = 0.32$. The estimation error is vanishing for $A = 0$ but becomes quite large around the value $A = A_0$. On the opposite, the Bayesian estimate takes into account that zero signal is probable for $A \in [0, A_0]$ and maps it to $\hat{A}_{\text{B}}(0) \approx A_0 \Gamma(1/n) / \Gamma(1/(2n)) = 0.49 A_0$. It is worth also mentioning the modified MLE, introduced in Appendix 5 of Ref. [30], where the MLE mapping is overridden by $\hat{A}'_{\text{MLE}} = A_0/2$ to

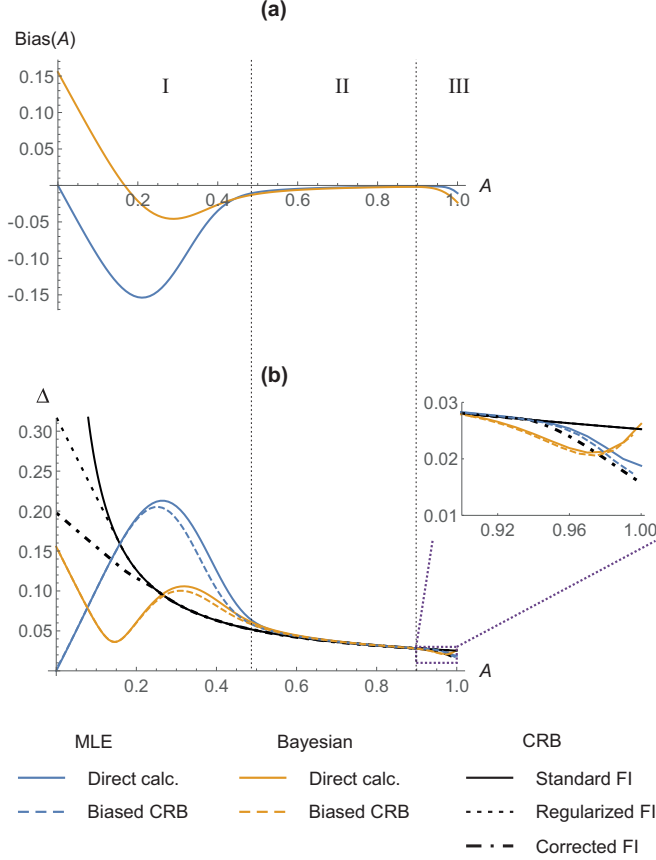


FIG. 6. Transmission amplitude estimation for a uniform object (single-parameter model): (a) estimation bias and (b) error. The notations of line styles are listed in the text. The regions I and III correspond to biased estimates for dark (A close to 0) and bright (A close to 1) objects, respectively. In region II, the value A is far from the boundaries and the estimate is almost unbiased. The inset in panel (b) shows an enlarged part of the region III, where the constraint $A \leq 1$ leads to the estimation bias.

improve the worst-case estimation error. Still, the MSE remains finite for the whole range $A \in [0, 1]$ for the considered estimators.

Black lines in Fig. 6(b) correspond to application of basic CRB [namely, Eq. (24)] to the unmodified FI $F(A)$ (solid black line), FI after regularization according to Eq. (27) (dotted black line; see Appendix G 3), and FI $\tilde{F}(A)$ regularized and then corrected for the constraints (dot-dashed black line; see Appendix G 4). The regularization procedure resolves the divergence of the basic CRB with unmodified FI at $A = 0$. The correction, described in Sec. III B, further improves the error estimate and makes it consistent with the actual inaccuracies indicated by the colored lines in Fig. 6(b) (up to the strong variation of the error for different estimators in the region I).

B. Two-parameter model

As the next step of the proposed procedure illustration, we consider the following two-parameter model. The object consists of two parts with the transmission amplitudes A_1 and

A_2 , representing the parameters of interest: $\theta \equiv \mathbf{A} = (A_1, A_2)$. The physical domain of the parameter values is naturally defined as $\Omega = \{\mathbf{A} | 0 \leq A_i \leq 1\}$. We assume that the object is illuminated by an SPDC light source, producing on average N photon pairs ($n = 2$) per an experiment, and the second-order autocorrelations are measured by a pair of detectors with the efficiency η . The signal vector $\mathbf{S}(\mathbf{A})$ contains two components, which are described according to Eq. (19) as

$$S_i(\mathbf{A}) = N\eta^2 G_{ii}^{(2)} = N\eta^2 \left| \sum_{m,l=1,2} D_{ml}^{(ii)} A_m A_l \right|^2, \quad i = 1, 2. \quad (37)$$

For simplicity, we assume that the coefficients $D_{ml}^{(ii)}$ are real, symmetric, and correspond to ideal spatial correlations of photons:

$$D_{ml}^{(ii)} \in \mathcal{R}, \quad D_{11}^{(11)} = D_{22}^{(22)} = h_0, \quad D_{22}^{(11)} = D_{11}^{(22)} = h_1, \\ D_{12}^{(ii)} = D_{21}^{(ii)} = 0. \quad (38)$$

Therefore, the signal components can be expressed as

$$S_1(\mathbf{A}) = N\eta^2 (h_0 A_1^2 + h_1 A_2^2)^2, \quad (39)$$

$$S_2(\mathbf{A}) = N\eta^2 (h_1 A_1^2 + h_0 A_2^2)^2. \quad (40)$$

The realizations of the measured signal $\mathbf{Y} = (Y_1, Y_2)$ are assumed to be characterized by Poisson distribution with Y_i having the expectation value $S_i(\mathbf{A})$. The FIM for the problem can be calculated according to Eq. (12) and equals

$$F(\mathbf{A}) = 16N\eta^2 (h_0^2 + h_1^2) \begin{pmatrix} A_1^2 & \zeta A_1 A_2 \\ \zeta A_1 A_2 & A_2^2 \end{pmatrix}, \quad (41)$$

where

$$\zeta = \frac{2h_0 h_1}{h_0^2 + h_1^2}, \quad |\zeta| < 1 \text{ for } h_0 \neq h_1. \quad (42)$$

For comparison with the error prediction by the proposed approach, we consider the following estimates of the parameters of interest $\mathbf{A}$:

(1) Constrained MLE:

$$\hat{\mathbf{A}}_{\text{MLE}}(\mathbf{Y}) = \underset{\mathbf{A}' \in \Omega}{\text{argmax}} [L(\mathbf{Y}|\mathbf{S}(\mathbf{A}'))] \quad (43)$$

where the likelihood for the two-component signal is defined according to Eq. (10).

(2) Bayesian mean *a posteriori* estimate:

$$\hat{\mathbf{A}}_{\text{B}}(\mathbf{Y}) = \frac{\int_0^1 dA'_1 \int_0^1 dA'_2 L(\mathbf{Y}|\mathbf{S}(\mathbf{A}')) \mathbf{A}'}{\int_0^1 dA'_1 \int_0^1 dA'_2 L(\mathbf{Y}|\mathbf{S}(\mathbf{A}'))}. \quad (44)$$

Figure 7 shows the results for the considered two-parameter model. The estimators listed above are applied to 1000 random samples of the signal $\mathbf{Y}$ for the true values of the transmission amplitudes localized in the lower part [Fig. 7(a)], the center [Fig. 7(b)], and the upper part [Fig. 7(c)] of the physical domain. The estimation results $\mathbf{Y} \mapsto \hat{\mathbf{A}}$ are shown as points in each part of the figure. The statistics of each generated dataset is shown as an ellipse, indicating the mean value $\langle \hat{\mathbf{A}} \rangle$ and the covariance $C = \text{Cov}(\hat{\mathbf{A}}, \hat{\mathbf{A}})$ according to the approach described in Appendix H:

$$\Delta \mathbf{A}^T C^{-1} \Delta \mathbf{A} = 2 \ln 2, \quad \Delta \mathbf{A} = \mathbf{A} - \langle \hat{\mathbf{A}} \rangle. \quad (45)$$

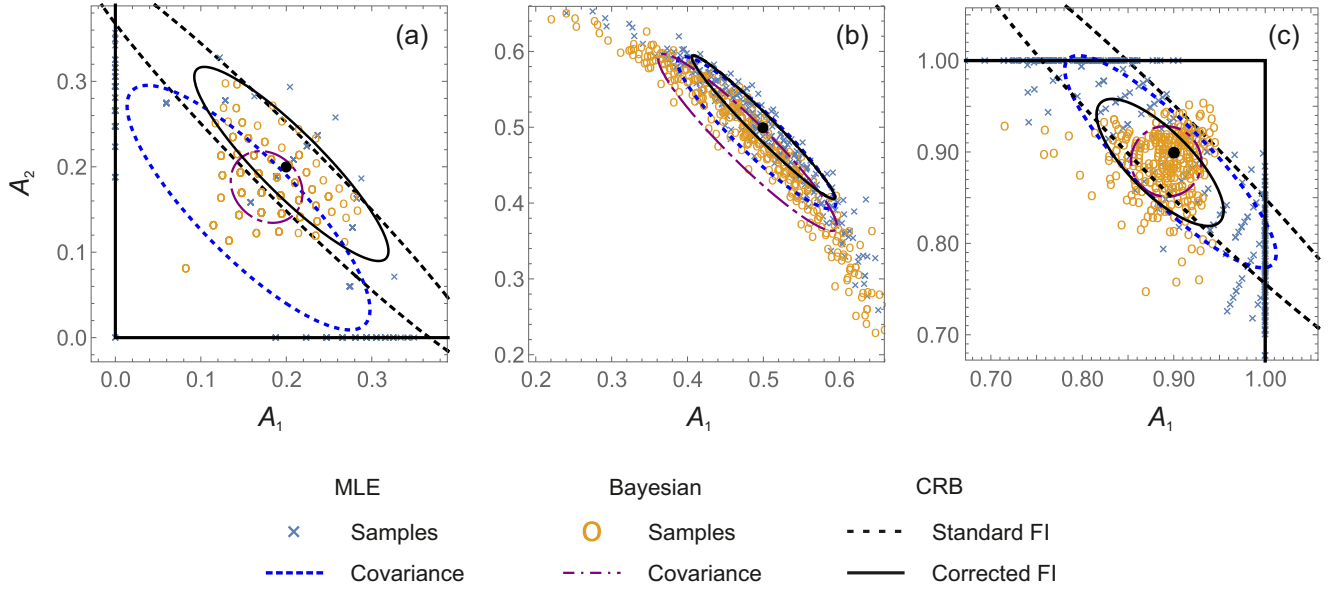


FIG. 7. Transmission amplitudes estimation for the two-parameter model. The true parameters values are (a) $A_1 = A_2 = 0.2$, (b) 0.5 , and (c) 0.9 , as indicated by black dots. Blue crosses and orange circles show the parameter inference results of a randomly sampled signal for MLE and a Bayesian estimate, respectively. One thousand samples were used for statistics estimation; 300 samples are shown in the figure; some results coincide due to discreteness of the signal. The modeling parameters are (a), (b) $\eta = 0.7$, $h_0 = 1$, $h_1 = 0.8$; $N = 1000$ and (c) 50 . Ellipses (blue dotted for MLE and purple dot-dashed for Bayesian estimate) are centered at the mean of the sampled results and have the shape defined by the corresponding covariance matrix (Appendix H). Black dotted ellipse (extending beyond the plotted regions) shows the predicted results scattering according to CRB with the standard FIM $F(\mathbf{A})$. Black solid ellipse corresponds to the regularized and corrected FIM $\tilde{F}(\mathbf{A})$. Straight black lines show the constraints imposed on the parameters A_1 and A_2 .

The results of random sampling are compared with the error predictions based on the CRB and FIM: the standard one $F(\mathbf{A})$, calculated according to Eq. (41), and the FIM $\tilde{F}(\mathbf{A})$ obtained by applying the proposed regularization and correction procedures to $F(\mathbf{A})$. The FIMs $F(\mathbf{A})$ and $\tilde{F}(\mathbf{A})$ are also visualized as ellipses (Appendix H).

First, one can notice that in the case, shown in Fig. 7(b), the true value $\mathbf{A} = (0.5, 0.5)$ is far from the boundaries. The regularization and correction procedures act trivially: $\tilde{F}(\mathbf{A}) = F(\mathbf{A})$ and the dashed and solid black ellipses coincide. The statistics of MLE and Bayesian estimates almost coincide with each other and with the prediction according to the CRB (still, some estimation bias is present).

For the cases when the constraints are active [Figs. 7(a) and 7(c)], the CRB with the standard FIM strongly overestimates the errors. The sampling statistics depends on the particular estimator used. Up to that variability, the corrected FIM provides a reasonable prediction of the parameters estimation inaccuracies.

C. Practical multiparametric models in quantum imaging

To conclude testing of the proposed approach, we apply it to multiparametric quantum imaging problems analyzed in Refs. [28,34]. The considered imaging setup is schematically shown in Fig. 8(a) and follows the ideas of the biphoton imaging experiment in Ref. [28]. An object consists of vertical slits [each having width d and a constant transmission amplitude—Fig. 8(b)] and is illuminated by entangled photon pairs. The radiation is collected by an optical system with the resolution limit d_R (classical Rayleigh limit for incoherent

imaging). The registered signal represents coincidence counts number for a pair of detectors.

First, we consider a ten-parameter model with the object shown in Fig. 8(b) [34] and parametrized by the transmission amplitudes A_m (each describing the region $(m-1)d \leq x \leq md$ of the object): $\theta \equiv \mathbf{A} = (A_1, \dots, A_{10})$. We choose $A_{\min} = 0.9$ for testing the correction-only procedure and $A_{\min} = 0$ as the case when both regularization and correction of the FIM are required.

The incident photon pairs are assumed to be ideally correlated, and the signal is described by Eq. (19) with $i = j$, $m = l$, and the sampling points corresponding to the step $d/2$ in the object plane:

$$S_j(\mathbf{A}) = NG^{(2)}(x_j, x_j) = N \left| \sum_m D_{mm}^{(jj)} A_m^2 \right|^2, \quad (46)$$

where N is the expectation value of the total number of detected coincidence events (already taking into account the detectors efficiency), $x_j = jd/2$,

$$D_{mm}^{(jj)} = 4k_{\max}^2 \int_{(m-1)d}^{md} ds \sin^2[k_{\max}(s - x_j)], \quad (47)$$

and $k_{\max} = 3.83/d_R$ is the momentum-transfer cutoff of the optical system [34].

Similarly to previous examples, Poisson statistics of the signal is assumed, and the FIM is calculated according to Eq. (12). The total value of the estimation error, predicted by the CRB, can be quantified by the quantity Δ^2 defined in Eq. (13) and applied to both the standard FIM $F(\mathbf{A})$ and its regularized and corrected counterpart $\tilde{F}(\mathbf{A})$. As a reference

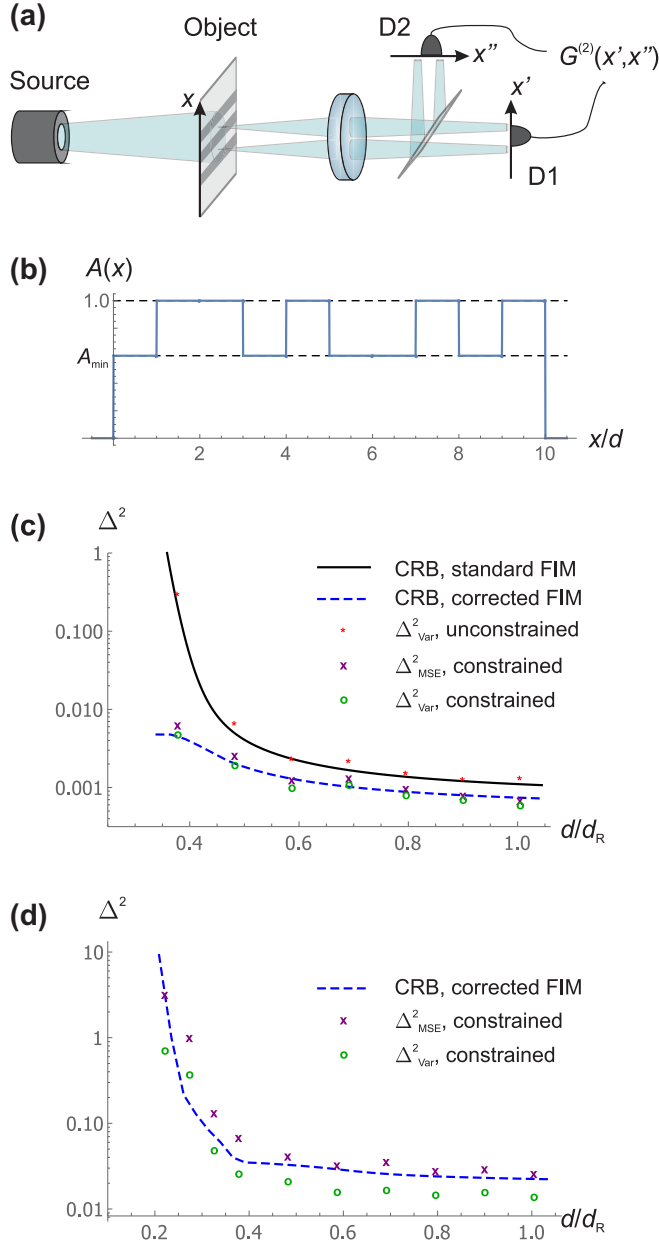


FIG. 8. Transmission amplitudes estimation for the ten-parameter model. (a) Scheme of the modeled imaging setup. Biphoton radiation, passing through the investigated object, is collected by an optical system and split between two single-photon detectors. The signal is represented by the correlation function $G^{(2)}(x', x'')$ depending on the detector positions x' and x'' . (b) Transmission amplitude of the model object as a function of the transverse coordinate x normalized by the width d of the slits (extended pixels) used for pixelization of the object. (c), (d) The dependence of the estimation error on the slit width d normalized by the Rayleigh resolution limit d_R for (c) $A_{\min} = 0.9$ and (d) $A_{\min} = 0$. The notations and line styles are described in the text.

for quality assessment of the predictions, we also perform a direct estimation of the transmission amplitudes $\hat{\mathbf{A}}$ for the signal $\mathbf{Y}$ randomly sampled according to its likelihood $L(\mathbf{Y}|\mathbf{S}(\mathbf{A}))$. The inference is performed by minimizing the distance

between the realized signal $\mathbf{Y}$ and its parametrized expectation value $\mathbf{S}(\hat{\mathbf{A}})$:

$$\hat{\mathbf{A}} = \operatorname{argmin}_{\mathbf{A}'} |\mathbf{Y} - \mathbf{S}(\mathbf{A}')|^2. \quad (48)$$

The estimation inaccuracy is characterized by the total variance and MSE for the sampling results:

$$\Delta_{\text{Var}}^2 = \sum_m \text{Var}(\hat{A}_m), \quad \Delta_{\text{MSE}}^2 = \langle |\hat{\mathbf{A}} - \mathbf{A}|^2 \rangle. \quad (49)$$

Figure 8(c) shows the dependence of the estimation error Δ^2 on the spatial scale of the problem d/d_R for the model object presented in Fig. 8(b) with $A_{\min} = 0.9$. The number of the detected coincidence events is taken to be $N = 10^4$. The error Δ^2 calculated for the standard FIM $F(\mathbf{A})$ (solid line) accurately describes the variance for unconstrained estimation (red star-marked points). If the constraint $0 \leq A_m \leq 1$ is imposed during the transmission amplitudes inference according to Eq. (48), the variance (green circles) and MSE (purple crosses) become smaller. Such error decrease is accurately reproduced by using CRB with the corrected FIM $\tilde{F}(\mathbf{A})$ (dashed blue line).

Figure 8(d) presents similar results for the case $A_{\min} = 0$, which requires regularization of FIM. Application of the CRB to the standard FIM does not yield meaningful results in that case since $F(\mathbf{A})$ is singular. For that reason, only the constrained estimation was sampled and compared with the error prediction by the regularized and corrected FIM $\tilde{F}(\mathbf{A})$. One can see that the agreement between the results of the proposed approach and direct sampling is good.

As the last example, we consider the object from a model quantum imaging experiment presented in Ref. [28]. The spatial resolution, achieved in practice, was higher than the theoretical bound predicted by CRB with the standard FIM. Such a counterintuitive result in Ref. [28] was the motivation for the current research. The measurement setup in the discussed experiment corresponded to the scheme in Fig. 8(a). The resolution of the optical system was artificially worsened by decreasing the numerical aperture for demonstration of sub-Rayleigh imaging and corresponded to the Rayleigh limit $d_R = 30.0 \mu\text{m}$. In contrast with the previous idealized ten-parameter imaging model, the spatial correlation of biphotons is not assumed to be ideal here (i.e., the photons from a pair have nonzero probability of passing through different parts of the object).

The object is shown in Fig. 9(a). Physically, it represents three nontransparent (dark) slits on a transparent background (a part of a positive USAF 1951 resolution target) with the slit width $31.25 \mu\text{m}$. For the statement of the estimation problem, the object was divided into 24 slit-like pixels and parametrized by 24 values of the transmission amplitude $\theta \equiv \mathbf{A} = (A_1, \dots, A_{24})$. The true values $\mathbf{A}$ used for our modeling [Fig. 9(a)] corresponded to the actual object used in the experiment [28].

The measured signal $\mathbf{S}(\mathbf{A})$ is represented by the full correlation matrix $G^{(2)}(x_i, x_j)$ without the constraint $i = j$ [28]:

$$S_{ij}(\mathbf{A}) = N G^{(2)}(x_i, x_j), \quad (50)$$

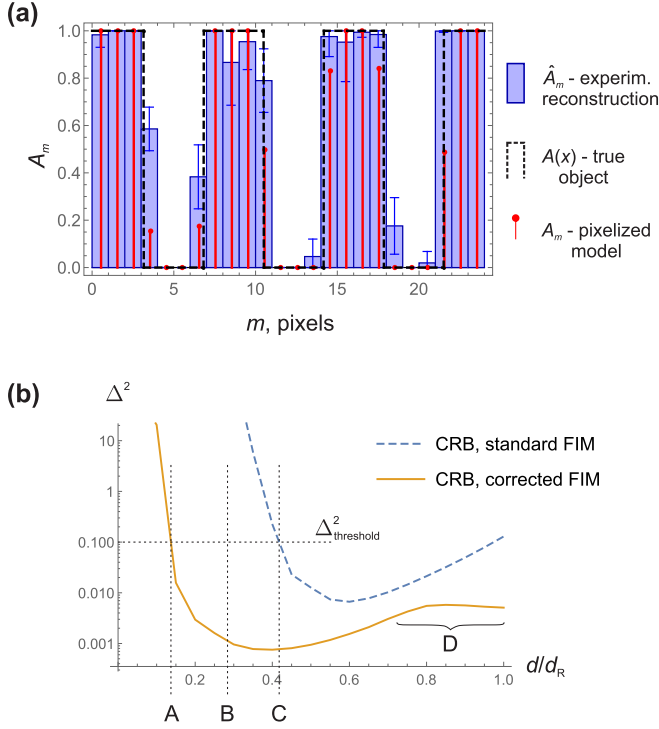


FIG. 9. Transmission amplitude estimation for the 24-parameter model from Ref. [28]. (a) Transmission amplitude of the model object: pixelized reconstruction result $\hat{\mathbf{A}} = \{\hat{A}_m\}$ from Ref. [28] (filled blue rectangles) with error bars built on the base of repeated measurement results; true transmission amplitude of the model object $A(x)$ (dashed black line); and its pixelized counterpart representing the true values $\mathbf{A} = \{A_m\}$ of the parameters of interest for the solved problem (red vertical bars). (b) The dependence of the estimation error on the slit width d normalized by the Rayleigh resolution limit d_R via the CRB with the standard FIM (dashed blue line) and the regularized and corrected FIM (solid orange line). The dotted vertical lines indicate characteristic resolution values: A—the resolution limit for the corrected FIM, defined for the threshold error value $\Delta^2_{\text{threshold}} = 0.1$; B—the spatial scale of the successful experimental inference of the object transmission amplitude [28]; C—the resolution limit for the standard FIM. In region D, the regularizing effect of the nonideal biphoton correlations decreases with the growth of the problem scale d .

where $G^{(2)}(x_i, x_j)$ is calculated according to Eq. (19), the expressions for the coefficients $D_{ml}^{(ij)}$ are provided in Ref. [28], and N is the expectation value of the total number of detected coincidence events. As previously, we assume independent Poissonian statistics for the components Y_{ij} of the measured signal $\mathbf{Y}$. Therefore, the FIM is described by Eq. (12) and takes the form

$$F_{ml}(\mathbf{A}) = 4N \sum_{i,j} \left(\sum_{l'} D_{ij}^{(ml')} A_{l'} \right) \left(\sum_{m'} D_{ij}^{(m'l)} A_{m'} \right). \quad (51)$$

Typically, it is sufficient to have at least one nondark pixel ($\exists l : A_l \neq 0$) to ensure nonzero values of the expressions in parentheses and, therefore, nonsingular FIM $F(\mathbf{A})$. The nonideal biphoton correlations produce regularizing effect by shifting the problem closer to coherent light imaging, where the FIM singularity issue does not arise.

Figure 9(b) shows the results of resolution estimation via the CRB for $N = 10^5$. The total reconstruction error Δ^2 , defined by Eq. (13), is calculated for the standard FIM $F(\mathbf{A})$ (dashed blue line) and the regularized and corrected one $\tilde{F}(\mathbf{A})$ according to the proposed approach (solid orange line). The achievable resolution can be defined as the minimal problem scale (pixel size) $d = d_{\min}$, for which the reconstruction error Δ^2 remains within certain reasonable limit: $\Delta^2 \leq \Delta^2_{\text{threshold}}$ [34]. For illustration purposes, the threshold value $\Delta^2_{\text{threshold}} = 0.1$, consistent with the actual experimental noise in Ref. [28], is used in Fig. 9(b). For such selection of the threshold, the theoretical resolution limit for the standard FIM is $d_{\min}/d_R \approx 0.42$, which is larger than the spatial scale $d/d_R = (8.51 \mu\text{m})/(30.0 \mu\text{m}) = 0.28$ of the problem, successfully solved experimentally [28]. FIM correction according to the proposed approach shifts the resolution limit to the value $d_{\min}/d_R \approx 0.14$, thus, resolving the seeming contradiction of beating the ultimate theoretical limit experimentally.

An interesting and rather counterintuitive feature of the obtained dependencies is the growth of the predicted error Δ^2 when the spatial scale d of the problem increases (region D). The effect is closely related to the regularizing influence of the finite transverse correlation length of nonideal biphotons, discussed after Eq. (51). When the pixel size increases and the correlation length remains constant, the cross-term coefficients $D_{ml}^{(ij)}$ with $m \neq l$ tend to zero and the standard FIM becomes ill-defined. However, the effect is less pronounced for the regularized and corrected version of FIM.

V. CONCLUSIONS

A priori constraints, imposed on quantum imaging problems, can improve the stability of the object parameters inference and decrease the errors beyond the CRB-based theoretical limit (strictly speaking, inapplicable due to estimation bias). The approach, proposed in the current paper, restores the applicability of the usual form of the CRB to such constrained problems by constructing a modified FIM that includes the constraints and is regularized if the initial (standard) FIM turns out to be ill-defined.

The proposed FIM correction and regularization technique represents a practical tool suitable for complex multiparametric models and accurate enough to be used for resolution analysis. A competitive approach, based on a modified CRB [43,47], requires knowledge of the bias derivatives, which may be infeasible for analytical calculation and very time-consuming for estimation by numerical Monte Carlo simulations. Even for a relatively simple two-parameter case, the analytical description of biased constrained estimates is cumbersome.

The proposed approach was successfully applied to 1-, 2-, 10-, and 24-parameter imaging problems and exhibited a good agreement with other methods of error estimation. For the simplest one-parameter model, it is possible to describe various estimators and to quantify resolution analytically. Many fundamental theoretical works have already been devoted to such simple models. For two parameters, it is still possible to estimate the errors directly, but the calculations are more complicated. When a multiparametric problem is considered,

the proposed procedure continues to work well and yields meaningful results, while other approaches require huge computational efforts. In particular, the technique developed of FIM modification helped us to explain the high resolution achieved experimentally in Ref. [28] and to predict the resolution limit for such a constrained problem quantitatively. While the illustrations provided assumed a Poisson distribution of the signal noise and the applicability of Eq. (12), the proposed method itself is not limited to such models in any way and can be applied to the FIM regardless of the specific type of expression used for its calculation. It is worth noting that the applicability of the developed approach can be extended to problems with a larger number of parameters (e.g., hundreds or thousands) by analyzing subproblems of a reasonable size after splitting the initial problem into almost independent windows [28].

The ability to construct the modified FIM for constrained problems can be useful for resolution estimation (as shown in the last considered example), problem structure analysis (similar to the one performed in Ref. [28], but more accurate for constrained problems), imaging setup optimization, and other similar tasks. The applicability of the approach extends far beyond quantum optics. For example, Monte Carlo simulations for high-resolution x-ray diffraction in Ref. [66] demonstrated decrease of the estimation variance when the natural physical constraints imposed on the parameters of interest became active. The proposed technique is capable of predicting such effect on the base of the modified FIM and CRB.

Since the QFI equals the classical FI for the optimal measurement, it is affected by the *a priori* constraints in a very similar way. Therefore, the developed approach is not limited to classical FI and can be applied to QFI as well.

ACKNOWLEDGMENTS

A.M., S.M., and D.M. acknowledge support from the National Academy of Sciences of Belarus program “Convergence.” A.M. and S.M. acknowledge partial support from BRFFR Grant No. F24KI-058 and from MSHE-RF Grant No. 075-15-2024-556. A.M. and D.M. also acknowledge partial support from BRFFR Grant No. F23UZB-064.

APPENDIX A: GEOMETRY OF LIKELIHOOD AND FISHER INFORMATION

The close connection of FI with distribution of the estimation results (via the CRB) on the one hand and with the likelihood [via the definition (3)] on the other hand gives rise to its frequentist and Bayesian interpretations (see, e.g., Refs. [51,52]). The geometry of the connection between the likelihood, distribution of the estimation results, and FI is illustrated by Fig. 10. The maps schematically (without direct relation to any specific physical model) represent the likelihood $L(Y|\theta)$ for a simple one-parameter problem with a scalar (one-dimensional) signal Y .

Figure 10(a) corresponds to the interpretation of FI in terms of the estimation results distribution for a fixed model. If the true value of the parameter equals θ , the possible realizations of the detected signal Y are distributed

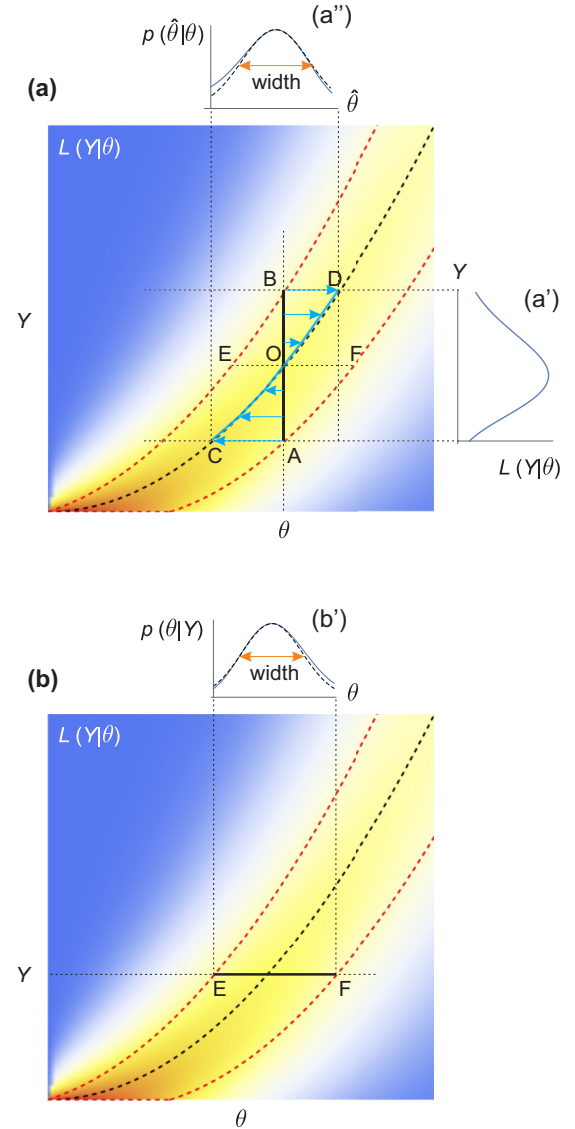


FIG. 10. Geometry of likelihood and FI: (a) distribution of the estimated parameter value $\hat{\theta}$ for a given true value θ of the parameter; (b) Bayesian posterior probability for the parameter θ for a given signal value Y . The details are provided in the text of Appendix A. The segment EF is also shown in panel (a) for the purpose of comparing the two interpretations. For simplicity, the signal Y is shown as a continuous variable, while for quantum-optical measurements it is typically a discrete variable taking integer values (the number of detected photons or coincidence events).

according to the likelihood $L(Y|\theta)$, shown in the Fig. 10(a') for specific θ and representing the vertical cross section AB of the map. Inference of the investigated parameter through its MLE $\hat{\theta}$ corresponds to processing the detected signal value Y by mapping $Y \mapsto \hat{\theta} = \arg\max_{\hat{\theta}} L(Y|\hat{\theta})$. Geometrically, each point of the segment AB is mapped onto the curve CD. The resulting distribution $p(\hat{\theta}|\theta)$, shown in the Fig. 10(a''), describes frequencies of getting the estimated values $\hat{\theta}$ in a detection-inference procedure if the true value of the investigated parameter equals θ . In the high-SNR limit, the distribution can be approximated by a Gaussian profile [dashed line in Fig. 10(a'')]. The width of the distribution

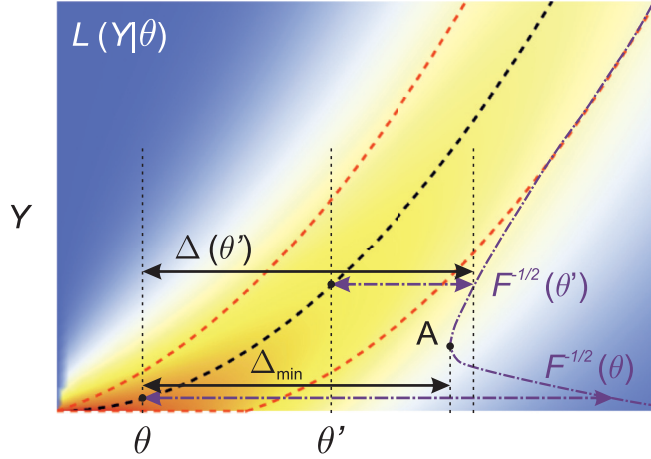


FIG. 11. FI regularization for one-parameter problem. The same likelihood diagram as in Fig. 10 is used. Purple dot-dashed line schematically shows the error estimate (upper bound for $\hat{\theta}$) based on FI [Eq. (24)], which diverges at $\theta \rightarrow 0$. The modified error estimate $\Delta(\theta')$ defined by Eq. (25) takes into account FI-based error estimate and the displacement $\theta' - \theta$ for a shifted parameter value θ' . The resulting regularized error estimate $\Delta_{\min}$ for θ is constructed by minimizing the quantity $\Delta(\theta')$ over θ' [Eq. (26)—i.e., by finding the rightmost point A along the purple dot-dashed curve].

[total length $l(\text{CA}) + l(\text{BD})$ of the segments CA and BD] characterizes the estimation error and is determined by FI.

The geometry for the Bayesian interpretation of FI is illustrated in Fig. 10(b). For a given value of the measured signal Y , the posterior distribution $p(\theta|Y)$ of the parameter θ [Fig. 10(b')] represents a horizontal cross-section EF of the likelihood map. In the high-SNR limit, the Gaussian approximation of the posterior probability is valid [dashed line in Fig. 10(b')], with its width $l(\text{EF})$ determined by the FI according to Eq. (7).

By overlaying the segment EF onto the plot in Fig. 10(a), one can see that, in the high-SNR limit, the two interpretations of the FI lead to the same predicted error (width of the probability distribution). In that limit, the curve CD is almost a straight line. Therefore, the shapes EOB and OBD, as well as COA and AOF are close to triangles, which are equal to each other pairwise. The width $l(\text{EF})$ of the posterior probability distribution in Bayesian interpretation is approximately the same as the deviation between the estimated and the true values of the parameter in the frequentist approach:

$$l(\text{EF}) = l(\text{EO}) + l(\text{OF}) \approx l(\text{BD}) + l(\text{CA}). \quad (\text{A1})$$

The introduced geometrical treatment of the likelihood and FI is also useful for illustration of the regularization procedure described by Eqs. (25) and (26) in Sec. III D—Fig. 11.

APPENDIX B: FISHER INFORMATION FOR INDEPENDENT EVENTS

For the specific case of the likelihood, defined by Eq. (10), the general expression (4) can be simplified in the following way: Due to additivity of FI [54] and independence of the signal components [factorization of the expression (10) for the

likelihood], the FIM is a sum of separate contributions $F_{\mu\nu}^{(i)}$:

$$F_{\mu\nu} = \sum_i F_{\mu\nu}^{(i)}, \quad (\text{B1})$$

where a term

$$F_{\mu\nu}^{(i)} = \sum_{Y_i=0}^{\infty} \frac{1}{L(Y_i|S_i(\theta))} \frac{\partial L(Y_i|S_i(\theta))}{\partial \theta_\mu} \frac{\partial L(Y_i|S_i(\theta))}{\partial \theta_\nu} \quad (\text{B2})$$

represents the FI for the i th component of the signal.

For the single-component likelihood, described by Eq. (11), the derivatives over the parameters can be expressed as

$$\frac{\partial L(Y_i|S_i(\theta))}{\partial \theta_\mu} = \frac{\partial S_i(\theta)}{\partial \theta_\mu} \left(\frac{Y_i}{S_i(\theta)} - 1 \right) L(Y_i|S_i(\theta)). \quad (\text{B3})$$

Substitution of Eq. (B3) into Eq. (B2) yields the resulting expression

$$\begin{aligned} F_{\mu\nu}^{(i)} &= \frac{\partial S_i(\theta)}{\partial \theta_\mu} \frac{\partial S_i(\theta)}{\partial \theta_\nu} \sum_{Y_i=0}^{\infty} \left(\frac{Y_i}{S_i(\theta)} - 1 \right)^2 L(Y_i|S_i(\theta)) \\ &= \frac{1}{S_i(\theta)} \frac{\partial S_i(\theta)}{\partial \theta_\mu} \frac{\partial S_i(\theta)}{\partial \theta_\nu}. \end{aligned} \quad (\text{B4})$$

Equation (12) is obtained by substitution of expression (B4) into Eq. (B1).

The presented approach to construction of the FIM is based on treating the measured dataset as a whole and considering the multivariate probability distribution for the signal vector $\mathbf{Y}$. An alternative approach, typical for quantum imaging [27,29–31], is based on calculation of FI $F^{(\text{event})}$ for a single-detection event with subsequent multiplication of the result by the number N of detected events. A single-detection event is characterized by the probability distribution $p_i(\theta)$, where the index i enumerates the possible outcomes. For example, when just field intensity is measured [$n = 1$ in Eq. (8)], the index i describes the position of a single-photon detection $\mathbf{r}^{(i)}$. The probabilities are normalized as $\sum_i p_i(\theta) = 1$, and the FIM can be constructed according to Eq. (4) as

$$F_{\mu\nu} = N F_{\mu\nu}^{(\text{event})} = N \sum_i \frac{1}{p_i(\theta)} \frac{\partial p_i(\theta)}{\partial \theta_\mu} \frac{\partial p_i(\theta)}{\partial \theta_\nu}. \quad (\text{B5})$$

One can easily relate the single-event probability distribution $p_i(\theta)$ with the mean number $S_i(\theta)$ of such events used in our approach as

$$S_i(\theta) = N p_i(\theta). \quad (\text{B6})$$

Therefore, Eqs. (12) and (B5) yield exactly the same result. Such equivalence of the approaches is not surprising since the Poisson distribution assumed during derivation of Eq. (12) is directly related to the independence of the events used in Eq. (B5).

APPENDIX C: ITERATIVE ALGORITHM OF FISHER INFORMATION MATRIX MODIFICATION

The iterative procedure of shrinking the multivariate normal probability distribution to fit certain linear constraints, described in Sec. III B of the main text, is illustrated by Fig. 12.

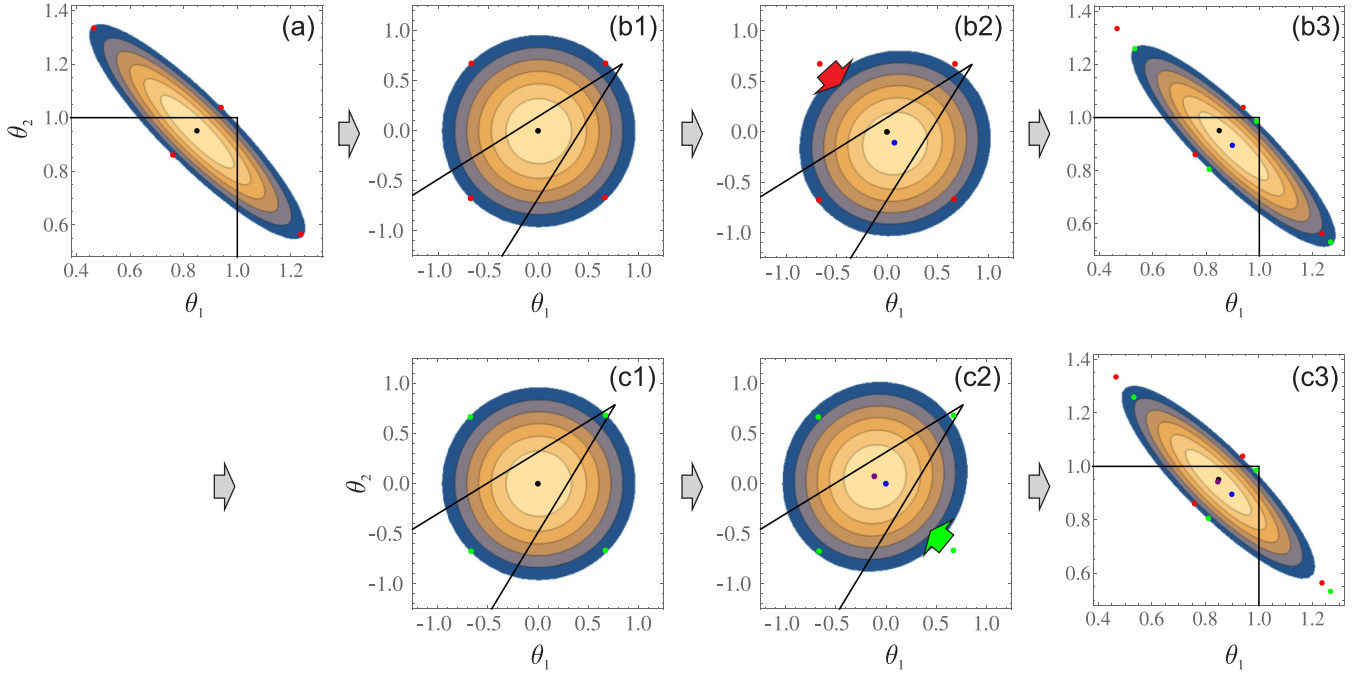


FIG. 12. Iterative shrinking of a model probability distribution in the two-parameter case: initial distribution in panel (a) and two subsequent iterations in panels (b1)–(b3) and (c1)–(c3). The constraints are $\theta_1 \leq 1$ and $\theta_2 \leq 1$. The black dot indicates the true values of the parameters $(\theta_1, \theta_2) = (0.85, 0.95)$. (b1), (c1) Each iteration starts from the coordinate transformation described in Sec. III B of the main text. (b2), (c2) Then, the resulting distribution is modified to reduce the probability of violating one of the constraints. (b3), (c3) Finally, the inverse coordinate transform is applied. Red and green points indicate the shape of the probability distribution before the first and the second iterations, respectively. Blue and purple points show the mean values for the modified distributions after the first and the second iterations, respectively.

The modification of the probability distribution follows the ideas from Ref. [53], but with certain changes discussed below.

After the coordinate transformation at the i th iteration, one ends up with the probability distribution

$$p^{(i)}(\theta') \propto \exp\left[-\frac{1}{2}(\theta')^T \theta'\right]. \quad (\text{C1})$$

For the k th constraint $(\mathbf{a}'_k)^T \theta' \leq b'_k$, one can define the probability of its violation as

$$P_k = \int_{(\mathbf{a}'_k)^T \theta' > b'_k} p^{(i)}(\theta') d\theta' = \frac{1}{2} \left[1 - \operatorname{erf} \left(\frac{b'_k}{\sqrt{2} |\mathbf{a}'_k|} \right) \right]. \quad (\text{C2})$$

The index of the most severely violated constraint is defined as

$$j = \operatorname{argmax}_k P_k = \operatorname{argmin}_k \frac{b'_k}{|\mathbf{a}'_k|}. \quad (\text{C3})$$

If the maximal probability of constraint violation P_j is smaller than certain threshold value (0.01 for the presented numerical calculations), the iterations are stopped and the probability distribution $p^{(i)}(\theta)$ is treated as the final result. Otherwise, one can introduce the variable $x = \mathbf{d}^T \theta'$ (where $\mathbf{d} = \mathbf{a}'_j / |\mathbf{a}'_j|$), use the symmetry of the distribution (C1), and consider the marginal distribution $w'(x) \propto \exp(-x^2/2)$ with the constraint $x \leq x_0 \equiv b'_j / |\mathbf{a}'_j|$. Then, the considered iteration modifies the marginal distribution as

$$w'(x) \mapsto w''(x) \propto \exp\left[-\frac{1+\xi}{2}(x+\delta)^2\right], \quad (\text{C4})$$

with the parameters ξ and δ being chosen according to the following requirements:

(1) The probability of the constraint violation should be reduced by certain quantity η (small enough to ensure convergence of iterations— $\eta = 0.1$ for the presented numerical calculations): $P_j \mapsto P'_j = \max(P_j/2, P_j - \eta)$:

$$P'_j = \frac{1}{2} \left[1 - \operatorname{erf} \left(\frac{(x_0 + \delta)\sqrt{1+\xi}}{\sqrt{2}} \right) \right]. \quad (\text{C5})$$

(2) The width of the distribution in the prescribed domain $x \leq x_0$ should be preserved:

$$\operatorname{Var}(x|w'(x), x \leq x_0) = \operatorname{Var}(x|w''(x), x \leq x_0). \quad (\text{C6})$$

The latter requirement differs from the one considered in Ref. [53], where preservation of the mean value $E(x|w'(x), x \leq x_0) = E(x|w''(x), x \leq x_0)$ was more appropriate for the quantum state reconstruction task.

The two listed requirements imply the following expressions for the quantities ξ and δ :

$$\xi = \frac{V(P'_j, x'_0)}{V(P_j, x_0)} - 1, \quad (\text{C7})$$

$$\delta = \frac{x'_0}{\sqrt{1+\xi}} - x_0, \quad (\text{C8})$$

where $x'_0 = \sqrt{2} \operatorname{erf}^{-1}(1 - 2P'_j)$ and

$$V(p, x) = 1 - \frac{xe^{-x^2/2}}{\sqrt{2\pi}(1-p)} - \frac{e^{-x^2}}{2\pi(1-p)^2}. \quad (\text{C9})$$

APPENDIX D: EXAMPLES OF WIDTH ESTIMATION BY THE SECOND DERIVATIVE WITH REGULARIZATION

For the model profile $y_1(x)$, defined in the main text, the second derivative of the logarithm takes the values

$$Y^{(2)}(x') = \begin{cases} 0 & \text{for } |x'| < x_0 \\ 1/\sigma^2 & \text{for } |x'| \geq x_0. \end{cases} \quad (\text{D1})$$

Therefore, according to the definition (22), $\Delta(x') \rightarrow \infty$ for $x' < x_0$ and $\Delta(x') = |x'| + \sigma$ otherwise. The minimum of $\Delta(x')$ is reached at $x' = \pm x_0$ and yields the accurate estimate of the half width,

$$\sigma_{\min} = x_0 + \sigma. \quad (\text{D2})$$

For the profile $y_2(x)$, the quantity $Y^{(2)}(x')$ equals

$$Y^{(2)}(x') = -\frac{k(k-1)|x'|^{k-2}}{2\sigma^k}. \quad (\text{D3})$$

The displaced half width $\Delta(x')$, defined by Eq. (22) is equal to

$$\Delta(x') = |x'| + \sqrt{\frac{2\sigma^k}{k(k-1)}}|x'|^{1-k/2} \quad (\text{D4})$$

and reaches the minimal value at

$$x' = \pm \left[\frac{(k-2)^2}{2k(k-1)} \right]^{1/k} \sigma. \quad (\text{D5})$$

The resulting half width estimate is

$$\sigma_{\min} = \left[\frac{(k-2)^2}{2k(k-1)} \right]^{1/k} \frac{k}{k-2} \sigma. \quad (\text{D6})$$

The dependence of the ratio $\sigma_{\min}/\sigma$ on k is shown in Fig. 5.

APPENDIX E: FISHER INFORMATION FOR ONE-PARAMETER PROBLEM WITH DISCRETE AND CONTINUOUS SIGNAL

Here, we show that discreteness of the detected signal strongly influences the behavior of FI for dark objects. For simplicity, the one-parameter model for a uniform object with the transmission amplitude A is considered and the mean value of the detected signal is assumed to scale as $S(A) \propto A^{2n}$.

First, let us consider the typical case, when the detected signal Y takes integer values distributed according to the Poisson distribution (11) with the mean $S(A)$. According to Eq. (12), the FI equals

$$F_P = \frac{1}{S(A)} \left(\frac{dS(A)}{dA} \right)^2 \propto A^{2(n-1)}. \quad (\text{E1})$$

Now, let us approximate the Poisson distribution by a continuous-variable Gaussian distribution with the same mean and variance:

$$L_G(Y|S(A)) = \frac{1}{\sqrt{2\pi S(A)}} \exp \left(-\frac{[Y - S(A)]^2}{2S(A)} \right). \quad (\text{E2})$$

Then, the FI can be calculated as

$$F_G = \int_{-\infty}^{\infty} dY \frac{1}{L(Y|S(A))} \left(\frac{dL(Y|S(A))}{dA} \right)^2 = \left(\frac{1}{S(A)} + \frac{1}{2S^2(A)} \right) \left(\frac{dS(A)}{dA} \right)^2 \quad (\text{E3})$$

In the limit of strong signal (bright object) $S(A) \gg 1$, the two noise distributions yield the same FI: $F_G \approx F_P$. However, for weak signal (dark object), the asymptotics is different:

$$F_G \sim \frac{1}{2S^2(A)} \left(\frac{dS(A)}{dA} \right)^2 \propto A^{-2} \rightarrow \infty, \quad (\text{E4})$$

while $F_P \rightarrow 0$ for $A \rightarrow 0$ and $n > 1$.

APPENDIX F: INFLUENCE OF PROBLEM PARAMETRIZATION ON FISHER INFORMATION

1. Parametrization-dependence of Fisher information

The toy problem, introduced in Sec. IV A, can also be parametrized by the transmission probability $T = A^2$. In that case, the FI, calculated according to Eq. (12) for such parametrization, will be equal to

$$\tilde{F}(T) = n^2 N \eta^n T^{n-2} \propto A^{2(n-2)}. \quad (\text{F1})$$

Comparing Eqs. (33) and (F1), one can notice that the FI has different asymptotics for $A \rightarrow 0$: if $n = 2$, FI for the transmission amplitude tends to zero, $F(A) \propto A^{2(n-1)} \rightarrow 0$, while $\tilde{F}(T) = n^2 N \eta^n$ remains constant. From that perspective, the following questions arise:

- (1) To what extent is FI parametrization-dependent?
- (2) Does reparametrization of a problem help to resolve the issue of the predicted error divergence in practice?

For a one-parameter problem, one can define the following change of variables:

$$\tilde{\theta} = g(\theta). \quad (\text{F2})$$

According to Eq. (3) and the chain rule for derivatives, the FI for the new parameter $\tilde{\theta}$ equals

$$\tilde{F} = \frac{1}{J^2} F, \quad J = \frac{d\tilde{\theta}}{d\theta} = \frac{dg(\theta)}{d\theta}. \quad (\text{F3})$$

CRB for the new parameter $\tilde{\theta}$ takes the form

$$\Delta\tilde{\theta}^2 \geq 1/\tilde{F} = J^2/F. \quad (\text{F4})$$

If FI tends to zero as $F = \Theta(\theta^{2m})$ for $\theta \rightarrow 0$ (in Knuth's notations [78]), one can choose the new parameter by setting $g(\theta) = \theta^{m+1}$ and ensure that the FI $\tilde{F}$ would scale as $\tilde{F} = \Theta(1)$ and the error $\Delta\tilde{\theta}$ would remain finite for $\theta \rightarrow 0$. For FI, defined by Eq. (33), the parametrization $\tilde{\theta} = A^n$ leads to $\tilde{F} = 4N\eta^n = \text{const}$ and resolves the FI singularity at $A = 0$.

Another important example is the so-called ‘‘Rayleigh’s curse’’—scaling of FI as $F = \Theta(d^2)$ for estimation of a small separation d of a pair of point-like incoherent light sources by direct imaging [29,31,33,79]. For the parametrization $\tilde{\theta} = d^2$, one has $\tilde{F} = \Theta(1)$ and the error divergence does not appear [the inaccuracy $\Delta(d^2)$ remains finite for $d \rightarrow 0$].

The presented considerations can also be extended to multiparameter problems [62,80]. The change of variables can be

introduced as

$$\tilde{\theta} = \mathbf{g}(\theta). \quad (\text{F5})$$

The FIM for the new parametrization can be calculated according to Eq. (3) and the chain rule for derivatives:

$$\tilde{F} = (J^{-1})^T F J^{-1}, \quad (\text{F6})$$

where J is the Jacobian matrix, $J_{\mu\nu} = \partial\tilde{\theta}_\mu/\partial\theta_\nu$. The CRB for the parameters $\tilde{\theta}$ takes the form

$$\text{Cov}(\tilde{\theta}, \tilde{\theta}) \geq \tilde{F}^{-1} = J F^{-1} J^T. \quad (\text{F7})$$

To conclude, the problem parametrization influences the behavior of FIM significantly and can change its asymptotics. Therefore, a methodological “chicken-or-egg” question arises: What comes first—the measurement or the problem parametrization? Do we seek for *any* kind of information about the investigated object *after* the measurement has been performed? Or do we specify the measurement *purpose* (the set of parameters θ to be inferred) at the stage of just *planning* the experiment? For practical applications, the latter approach is generally favored. Some remarks on the optimal choice of parameters are presented in Appendix F3.

2. Fisher information and error analysis for a predefined set of parameters

In the main text, we do not consider the task of finding the optimal task-specific parametrization and assume that the problem statement already prescribes certain fixed “natural” set of parameters θ and enforces us to report the results in terms of those specific predefined variables. After estimation of the errors for the introduced parameters $\tilde{\theta}$ according to Eqs. (F4) and (F7), one needs to return to the initial set of parameters θ . According to the error-propagation rules, the errors for the initial variables can be estimated as

$$\Delta\theta^2 = \Delta\tilde{\theta}^2/J^2 \quad (\text{F8})$$

for a one-parameter model and

$$\text{Cov}(\theta, \theta) = J^{-1} \text{Cov}(\tilde{\theta}, \tilde{\theta}) (J^{-1})^T \quad (\text{F9})$$

in the general case.

The transition from $F = \Theta(\theta^{2m}) \rightarrow 0$ for $\theta \rightarrow 0$ to $\tilde{F} = \Theta(1)$, discussed above, requires $J = \Theta(\theta^m) \rightarrow 0$. Therefore, the problem of the error divergence re-appears as soon as we return to the initial variable(s). For example, in the case of Rayleigh’s curse, a small error $\Delta(d^2)$ of the squared separation d^2 does not necessarily imply accurate estimation of the separation d itself for small d , since $\Delta d \approx \Delta(d^2)/(2d)$. Similarly, a small error of $\tilde{\theta} = A^n$ in the toy example from Sec. IV A does not guarantee accurate inference of the transmission amplitude A .

As a result, the reparametrization of the problem just shifts the necessity of regularization from the FIM F to the errors recalculation according to Eqs. (F8) and (F9). For a one-parameter case, one can go beyond the linear approximation lying behind the error propagation rules and estimate $\Delta\theta$ in a more accurate way (for example, by recalculating Bayesian posterior probabilities). In a general multiparametric case, the problem of regularizing Eq. (F9) looks at least as complicated

as our approach operating with initial FIM F itself without resorting to the reparametrization.

3. Optimal parametrization for Fisher information analysis

In this section, we briefly discuss how the choice of the problem parametrization can improve calculation of FI and application of CRB, if no strict constraints are imposed on the object parameters to be reported.

If the problem is described by a single parameter $\bar{\theta}$, defined by Eq. (F2) and known *a priori* to have the value close to $\bar{\theta}_0$, one can construct the efficient (saturating CRB) locally unbiased estimator [56,62]:

$$\begin{aligned} \hat{\bar{\theta}}(Y, \bar{\theta}_0) &= \bar{\theta}_0 + \bar{F}^{-1} \frac{d}{d\bar{\theta}} \ln L(Y|S(\bar{\theta})) \Big|_{\bar{\theta}=\bar{\theta}_0} \\ &= \left(\bar{\theta}_0 - \frac{S(\bar{\theta}_0)}{S'(\bar{\theta}_0)} \right) + \frac{Y}{S'(\bar{\theta}_0)}, \end{aligned} \quad (\text{F10})$$

where we took into account Eq. (11) and introduced the notation $S'(\bar{\theta}_0) = dS(\bar{\theta})/d\bar{\theta}|_{\bar{\theta}=\bar{\theta}_0}$. Here, we do not take into account the additional bias imposed by the upper bound ($A \leq 1$ in the toy example in Sec. IV A).

If, for certain parametrization (specific choice of $\bar{\theta}$), both terms on the right-hand side of Eq. (F10) do not depend on $\bar{\theta}_0$, $\hat{\bar{\theta}}(Y)$ represents the globally unbiased efficient estimator:

$$\begin{aligned} S'(\bar{\theta}_0) &= \alpha, \\ \frac{S(\bar{\theta}_0)}{S'(\bar{\theta}_0)} - \bar{\theta}_0 &= \beta, \end{aligned} \quad (\text{F11})$$

where α and β are constants (independent of $\bar{\theta}_0$). The system of equations (F11) implies that the relation between the signal S and the parameter $\bar{\theta}$ must be linear:

$$S(\bar{\theta}) = \alpha\bar{\theta} + \beta. \quad (\text{F12})$$

In particular, for the one-parameter toy example, $\bar{\theta}$ can be chosen as $\bar{\theta} = A^{2n}$. It is worth noting that the constructed unbiased estimator $\hat{\bar{\theta}}(Y)$ coincides with MLE for that parametrization.

In contrast with the optimal parametrization, the initial parametrization of the toy problem by the transmission amplitude A does not allow construction of a globally unbiased estimator. Since $S(0) = 0$, the mean bias of an estimator $\bar{A}(Y)$ at the point $A = 0$ equals

$$\langle \bar{A} \rangle \Big|_{A=0} = \sum_{Y=0}^{\infty} \bar{A}(Y) L(Y|S(0)) = \bar{A}(0). \quad (\text{F13})$$

Therefore, only estimators with $\bar{A}(0) = 0$ are unbiased at $A = 0$. The derivative of the mean bias at $A = 0$ for the toy example equals

$$\frac{d\langle \bar{A} \rangle}{dA} \Big|_{A=0} = \bar{A}(1) \frac{dS(A)}{dA} \Big|_{A=0} = 2n\bar{A}(1)N\eta^n A^{2n-1} \Big|_{A=0} = 0 \quad (\text{F14})$$

for $n \geq 1$, while for an unbiased estimator the equality

$$\frac{d\langle \bar{A} \rangle}{dA} = 1 \quad (\text{F15})$$

must hold.

Treatment of a general multiparametric case is similar. An efficient locally unbiased estimator is constructed as [62]

$$\begin{aligned}\hat{\boldsymbol{\theta}}(\mathbf{Y}, \bar{\boldsymbol{\theta}}_0) &= \bar{\boldsymbol{\theta}}_0 + \bar{\mathbf{F}}^{-1} \nabla \ln L(\mathbf{Y}|\mathbf{S}(\bar{\boldsymbol{\theta}}))|_{\bar{\boldsymbol{\theta}}=\bar{\boldsymbol{\theta}}_0} \\ &= \bar{\boldsymbol{\theta}}_0 + \sum_i F^{-1}(\bar{\boldsymbol{\theta}}_0) \left(\frac{Y_i}{S_i(\bar{\boldsymbol{\theta}}_0)} - 1 \right) \nabla S_i(\bar{\boldsymbol{\theta}})|_{\bar{\boldsymbol{\theta}}=\bar{\boldsymbol{\theta}}_0} \\ &\equiv \mathbf{C}(\bar{\boldsymbol{\theta}}_0) \mathbf{Y} + \mathbf{D}(\bar{\boldsymbol{\theta}}_0),\end{aligned}\quad (\text{F16})$$

where

$$\mathbf{D}(\bar{\boldsymbol{\theta}}_0) = \bar{\boldsymbol{\theta}}_0 - \mathbf{C}(\bar{\boldsymbol{\theta}}_0) \mathbf{S}(\bar{\boldsymbol{\theta}}_0). \quad (\text{F17})$$

The estimator is globally unbiased and efficient if the matrix $\mathbf{C}(\bar{\boldsymbol{\theta}}_0) = \mathbf{C}$ and the vector $\mathbf{D}(\bar{\boldsymbol{\theta}}_0) = \mathbf{D}$ do not depend on the parameter values $\bar{\boldsymbol{\theta}}_0$. In that case Eq. (F17) implies that such estimator can be constructed only for a parametrization with the following linear relation between the parameter vector $\boldsymbol{\theta}$ and the signal $\mathbf{S}(\boldsymbol{\theta})$:

$$\boldsymbol{\theta} = \mathbf{C} \mathbf{S}(\boldsymbol{\theta}) + \mathbf{D}. \quad (\text{F18})$$

For the two-parameter toy example, discussed in Sec. IV B, the parameters choice

$$\bar{\theta}_1 = S_1(\mathbf{A}), \quad \bar{\theta}_2 = S_2(\mathbf{A}) \quad (\text{F19})$$

indeed ensures that the MLE is globally unbiased and efficient (except for vicinity of the upper bound $A_i \leq 1$). However, such parametrization is far from being practical. The ultimate goal of imaging (microscopy) is to retrieve *local* information about the object features from the results of available measurements. Even allowing nonlinear reparametrization, one would prefer to retain locality by defining the i th new parameter as a function of the i th initial parameter only:

$$\bar{\theta}_i = g_i(\theta_i). \quad (\text{F20})$$

For the considered two-parameter toy example, Eqs. (F18) and (F20) can be satisfied simultaneously if the equality

$$2h_0h_1(h_1^2 - h_0^2) = 0 \quad (\text{F21})$$

holds, which is possible only in trivial special cases: $h_0 = 0$, or $h_1 = 0$ (no blurring), or $h_0 = h_1$ [degenerate problem with $S_1(\mathbf{A}) \equiv S_2(\mathbf{A})$].

For the multiparametric problems, discussed in Sec. IV C, an attempt to satisfy the conditions of linearity [Eq. (F17)] and locality [Eq. (F20)] simultaneously leads to the same issue: the number of additional constraints imposed on the imaging system parameters $D_{ml}^{(ij)}$ grows quadratically relatively to the number of the object parameters A_i (and $\bar{\theta}_i$). Therefore, the existence of the optimal parametrization allowing construction of the globally unbiased efficient estimator and satisfying the locality restriction (F20) is limited to certain very special cases.

The necessity to deal with parametrizations not allowing construction of globally unbiased efficient estimators is common for quantum multiparameter estimation problems [62,80–82]. In the limit of a large number of detected events (measurement repetitions), the MLE becomes asymptotically unbiased (except for the parameters belonging to the border of the physical region Ω) and the CRB provides an accurate estimate of error [62,80,81]. The effect manifests itself as a

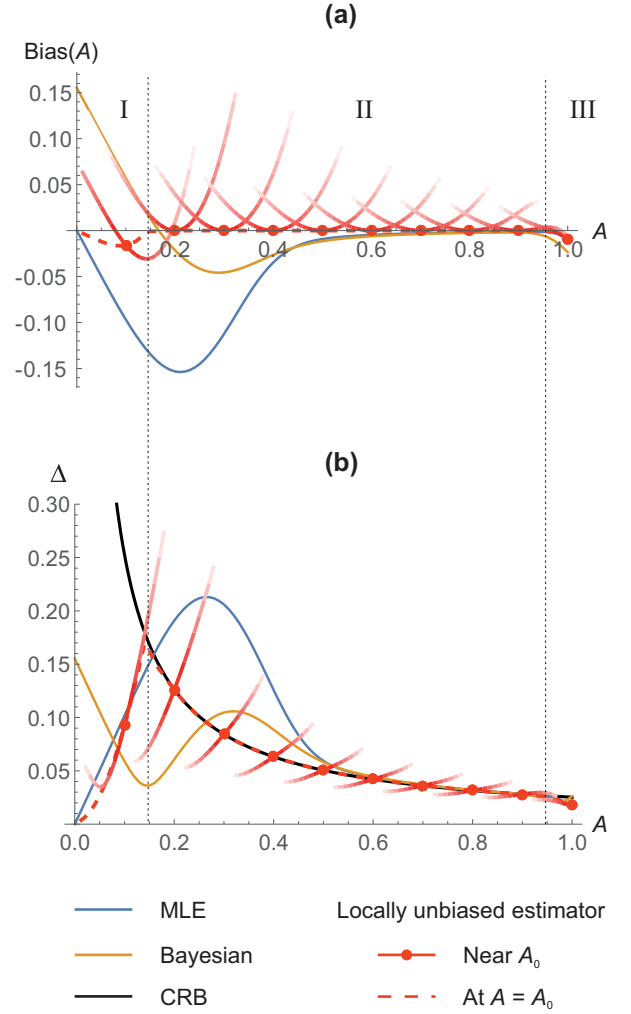


FIG. 13. Transmission-amplitude estimation for a uniform object (single-parameter model): (a) estimation bias and (b) MSE. Solid red lines show the behavior of the locally unbiased estimator $\hat{A}(Y, A_0)$ defined by Eq. (F22) for ten values A_0 indicated by red dots. Dashed red lines aggregate the bias and MSE for the estimators $\hat{A}(Y, A_0)$ with the value A_0 equal to the true value A . Other lines are taken from Fig. 6. The regions I–III are defined according to the bias of the estimators $\hat{A}(Y, A_0)$ with $A_0 = A$ (biased in I and III, locally unbiased in II).

small bias in the region II in Fig. 6 even for the relatively small mean number of detection events $N\eta^n = 98$.

If the number of detection events is not large, one can use the locally unbiased estimator [56,62], defined by Eqs. (F10) and (F16), with the rough estimate of $\bar{\boldsymbol{\theta}}_0$ being found by the procedure outlined in Ref. [82]. For the one-parameter toy example with the initial parametrization by the transmission amplitude A , the resulting estimator takes the form

$$\hat{A}(Y, A_0) = \min \left\{ 1, A_0 \left[1 + \frac{1}{2n} \left(\frac{Y}{S(A_0)} - 1 \right) \right] \right\}, \quad (\text{F22})$$

where the upper bound $A \leq 1$ is taken into account explicitly. The bias and MSE of that estimator for the parameters, used in Fig. 6, are shown in Fig. 13. As expected, the constructed estimator is locally unbiased and saturates the CRB in the region

II, not including the vicinity of the boundary values $A = 0$ and $A = 1$. The constraint, introduced by the function $\min\{1, \dots\}$ in Eq. (F22), leads to the estimation bias and decrease of MSE relatively to the CRB in regions I and III. To account for such effect, the FI modification and regularization approach proposed in the current paper can be applied. It is interesting to note that the bias in region I is still caused by the *upper* bound $A \leq 1$: the locally unbiased estimator $\hat{A}(Y, A_0)$, defined by Eq. (F10), diverges for $A_0 \rightarrow 0$ [$\lim_{A_0 \rightarrow 0} \hat{A}(1, A_0) = \infty$], while the constraint introduced in Eq. (F22) regularizes it at the cost of additional bias.

APPENDIX G: CALCULATIONS FOR THE ONE-PARAMETER PROBLEM OF TRANSMISSION AMPLITUDE ESTIMATION

1. Optimal biased estimate

Optimizing the bias of a parameter estimate, one can reduce the MSE of the parameter relatively to unbiased estimates [41,43,44,47]. For the model, described in Sec. IV A, one can specify the cost function for the bias optimization as the MSE averaged over the whole physical range of the true parameter values $A \in [0, 1]$:

$$\langle \Delta^2 \rangle_A \equiv \int_0^1 dA E[(\hat{A}(Y) - A)^2] \quad (\text{G1})$$

where $\hat{A}(Y)$ is the estimator to be optimized, and the expectation value over the signal realization is defined as

$$E[f(Y)] = \sum_{Y=0}^{\infty} L(Y|S(A))f(Y). \quad (\text{G2})$$

Substituting Eq. (G2) into Eq. (G1), one can obtain the following expression for the cost function:

$$\begin{aligned} \langle \Delta^2 \rangle_A &= \sum_{Y=0}^{\infty} [(L(Y|S(A)))_A \hat{A}^2(Y) - 2(L(Y|S(A)))_A \hat{A}(Y) \\ &\quad + (L(Y|S(A)A^2))_A], \end{aligned} \quad (\text{G3})$$

where the brackets $\langle \dots \rangle$ denote averaging over A as in Eq. (G1). The optimization is performed over the discrete set of variables $\hat{A}(Y)$, parametrized by an integer index Y :

$$\frac{\partial}{\partial \hat{A}(Y)} \langle \Delta^2 \rangle_A = 0 \text{ for all } Y = 0, 1, \dots \quad (\text{G4})$$

The minimum is reached when the following equation is satisfied:

$$\hat{A}(Y) = \frac{(L(Y|S(A)))_A}{(L(Y|S(A)))_A} = \frac{\int_0^1 dA' L(Y|S(A'))A'}{\int_0^1 dA' L(Y|S(A'))}. \quad (\text{G5})$$

The result coincides with the mean *a posteriori* estimate defined by Eq. (35).

2. Mean squared error estimation

For numerical analysis of the estimation error in a one-parameter case, one can start from constructing the mapping dictionary $Y \mapsto \hat{A}(Y)$ for the considered estimate. Equations (34) and (35) are to be used for MLE and Bayesian estimates defined in the main text. Then, the MSE and the

bias are calculated as functions of the true parameter value A as

$$\text{MSE}(A) = E[(\hat{A}(Y) - A)^2], \quad \text{Bias}(A) = E[\hat{A}(Y) - A], \quad (\text{G6})$$

where the expectation value is defined by Eq. (G2).

FI-based error evaluation includes application of standard CRB [$\text{Var}(A) \sim 1/F$], modified the CRB for a biased estimator [41,43]

$$\begin{aligned} \text{MSE}(A) &= \text{Var}(A) + [\text{Bias}(A)]^2 \sim \left(1 + \frac{\partial \text{Bias}(A)}{\partial A}\right)^2 \frac{1}{F} \\ &\quad + [\text{Bias}(A)]^2, \end{aligned} \quad (\text{G7})$$

and the proposed approach with the corrected FI and the standard CRB ($\text{MSE} \sim 1/\tilde{F}$).

3. Fisher information regularization for dark objects

Regularization of FI for small A is performed according to Eq. (27) after substitution of Eq. (33):

$$\tilde{F}(A) = \max_{A'} \frac{KA'^{2(n-1)}}{(1 + |A' - A|\sqrt{KA'^{n-1}})^2}, \quad (\text{G8})$$

where $K = 4n^2N\eta^n$. For $A' < A$ the derivative of the optimized expression is positive, i.e., the optimal A' is not less than A . For $A' > A$ the maximum of the optimized expression is reached at

$$A'_{\text{opt}} = \left(\frac{n-1}{\sqrt{K}}\right)^{1/n} \quad (\text{G9})$$

if $A \leq A'_{\text{opt}}$. Otherwise, the optimal value is $A' = A$ and no regularizing correction is applied.

4. Fisher information correction for the constraints

For the one-parameter model, considered in Sec. IV A, the constraints to be taken into account by FI modification are $0 \leq A \leq 1$. The input FI is the result of regularization described in Appendix G 3. To avoid confusion with the final result of correction, we denote the input FI as F [not $\tilde{F}$ as in Eq. (G8)]. We follow the procedure from Sec. III B.

Both the FIM F and the transformation matrix T have dimensions 1×1 in the considered case and represent just scalars: $T = \sqrt{F}$. The parameter transformation (for the first iteration) is $A' = \sqrt{F}(A - A^{(0)})$, where $A^{(0)}$ is the transmission amplitude value, for which FI is analyzed. Parametrization of the constraints in the form $a_j A \leq b_j$ is

$$A \leq 1 \Rightarrow a_1 = 1, \quad b_1 = 1, \quad (\text{G10})$$

$$A \geq 0 \Rightarrow -A \leq 0 \Rightarrow a_0 = -1, \quad b_0 = 0. \quad (\text{G11})$$

The transformed parameters of the constraints are

$$a'_1 = \frac{1}{\sqrt{F}}, \quad b'_1 = 1 - A^{(0)}, \quad a'_0 = -\frac{1}{\sqrt{F}}, \quad b'_0 = A^{(0)}. \quad (\text{G12})$$

The local variable x , introduced in Appendix C, coincides with the transformed parameter up to the sign: $x = \pm A'$,

where the “+” and “−” signs correspond to the constraint indices 1 and 0, respectively. The limiting values $x_0^{(j)} = b_j/|a_j|$ for the variable x are equal to

$$x_0^{(1)} = \sqrt{F}(1 - A^{(0)}), \quad x_0^{(0)} = \sqrt{F}A^{(0)}. \quad (\text{G13})$$

The probabilities P_1 and P_0 of the constraints violation are defined by Eq. (C2). For the model, discussed in Sec. IV A, the following relations hold for those probabilities:

$$P_1 \leq P_{\text{threshold}} \equiv 0.01 \text{ for } A \leq 0.937, \quad (\text{G14})$$

and

$$P_0 \leq P_{\text{threshold}} \text{ for } A \geq 0.242. \quad (\text{G15})$$

Therefore, for each A at most one of the constraints is active and requires application of the correction procedure. Due to the additivity of corrections applied to the same direction [53], one can avoid multiple iterations and reduce the constraint violation probability from its initial value to the target value P_{target} in one step.

The shrinking parameter ξ for the discussed correction step is prescribed by Eq. (C7) with $P'_j = P_{\text{target}}$. According to the last step of the correction procedure described in Sec. III B, the resulting corrected FI equals $\tilde{F} = (1 + \xi)F$, where F is the input FI (namely, the regularization result).

APPENDIX H: VISUALIZATION OF FISHER INFORMATION AND COVARIANCE MATRIX FOR A TWO-PARAMETER MODEL

Since the FIM is a positive semidefinite matrix, it can serve as a quadratic form defining an ellipse in a two-parameter case $\theta = (\theta_1, \theta_2)$. The meaning of the ellipse is clearly seen from Eqs. (15) and (16) (see also Fig. 1 in Ref. [66] for illustration).

If the estimation results are scattered according to multivariate normal distribution,

$$p(\theta) = \frac{\sqrt{\det F}}{2\pi} \exp \left[-\frac{1}{2} \Delta\theta^T F \Delta\theta \right], \quad \Delta\theta = \theta - \theta_0, \quad (\text{H1})$$

centered at some point θ_0 , the equation

$$\Delta\theta^T F \Delta\theta = 2 \ln 2 \quad (\text{H2})$$

defines the boundary between the regions with $p(\theta) > 1/2$ and $p(\theta) < 1/2$. Moreover, each region will, on average, contain half of the sampled results:

$$\int_{\Delta\theta^T F \Delta\theta \leq 2 \ln 2} d^2\theta p(\theta) = \frac{1}{2\pi} \int_{|\mathbf{r}| \leq 1} d^2\mathbf{r} e^{-\mathbf{r}^2/2} = \frac{1}{2}. \quad (\text{H3})$$

Since the CRB with the standard FIM is valid for unbiased estimates, it is reasonable to take θ_0 equal to the true value of the parameters θ for visualization of the standard FIM. The correction procedure (Sec. III B) shifts the probability distribution. Therefore, θ_0 should be taken as the center $\theta^{(i_{\text{max}})}$ of the distribution after the last iteration (with the index i_{max}) for visualization of the corrected FIM $\tilde{F}$.

To visualize statistics of a set of randomly sampled points (in two-dimensional space), one can assume a multivariate normal distribution based on the sample mean $\langle\theta\rangle$ and covariance $C = \text{Cov}(\theta, \theta)$:

$$p(\theta) = \frac{1}{2\pi\sqrt{\det C}} \exp \left[-\frac{1}{2} \Delta\theta^T C^{-1} \Delta\theta \right], \quad \Delta\theta = \theta - \langle\theta\rangle. \quad (\text{H4})$$

The quadratic form of the distribution corresponds to the ellipse defined by the equation

$$\Delta\theta^T C^{-1} \Delta\theta = 2 \ln 2 \quad (\text{H5})$$

in the sense discussed above.

-
- [1] E. Abbe, Beiträge zur theorie des mikroskops und der mikroskopischen wahrnehmung, *Arch. Mikr. Anat.* **9**, 413 (1873).
 - [2] L. Rayleigh, XXXI. Investigations in optics, with special reference to the spectrocope, *Philos. Mag.* **8**, 261 (1879).
 - [3] K. Rossmann, Point spread-function, line spread-function, and modulation transfer function: tools for the study of imaging systems, *Radiology* **93**, 257 (1969).
 - [4] Y. Shih, *An Introduction to Quantum Optics: Photon and Biphoton Physics* (CRC Press, New York, 2018).
 - [5] O. S. Magaña-Loaiza and R. W. Boyd, Quantum imaging and information, *Rep. Prog. Phys.* **82**, 124401 (2019).
 - [6] L. Schermelleh, A. Ferrand, T. Huser, C. Eggeling, M. Sauer, O. Biehlmaier, and G. P. C. Drummen, Super-resolution microscopy demystified, *Nat. Cell Biol.* **21**, 72 (2019).
 - [7] K. Prakash, B. Diederich, R. Heintzmann, and L. Schermelleh, Super-resolution microscopy: A brief history and new avenues, *Philos. Trans. R. Soc. A* **380**, 20210110 (2022).
 - [8] V. Giovannetti, S. Lloyd, L. Maccone, and J. H. Shapiro, Sub-Rayleigh-diffraction-bound quantum imaging, *Phys. Rev. A* **79**, 013827 (2009).
 - [9] D.-Q. Xu, X.-B. Song, H.-G. Li, D.-J. Zhang, H.-B. Wang, J. Xiong, and K. Wang, Experimental observation of sub-Rayleigh quantum imaging with a two-photon entangled source, *Appl. Phys. Lett.* **106**, 171104 (2015).
 - [10] M. Unternährer, B. Bessire, L. Gasparini, M. Perenzoni, and A. Stefanov, Super-resolution quantum imaging at the Heisenberg limit, *Optica* **5**, 1150 (2018).
 - [11] E. Toninelli, P.-A. Moreau, T. Gregory, A. Mihalyi, M. Edgar, N. Radwell, and M. Padgett, Resolution-enhanced quantum imaging by centroid estimation of biphotons, *Optica* **6**, 347 (2019).
 - [12] T. Dertinger, R. Colyer, G. Iyer, S. Weiss, and J. Enderlein, Fast, background-free, 3D super-resolution optical fluctuation imaging (SOFI), *Proc. Natl. Acad. Sci. USA* **106**, 22287 (2009).
 - [13] S. T. Hess, T. P. K. Girirajan, and M. D. Mason, Ultra-high resolution imaging by fluorescence photoactivation localization microscopy, *Biophys. J.* **91**, 4258 (2006).
 - [14] M. J. Rust, M. Bates, and X. Zhuang, Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM), *Nat. Methods* **3**, 793 (2006).

- [15] J. Schnitzbauer, M. T. Strauss, T. Schlichthaerle, F. Schueder, and R. Jungmann, Super-resolution microscopy with DNA-PAINT, *Nat. Protoc.* **12**, 1198 (2017).
- [16] M. Gu, Y. Cao, S. Castelletto, B. Kouskousis, and X. Li, Super-resolving single nitrogen vacancy centers within single nanodiamonds using a localization microscope, *Opt. Express* **21**, 17639 (2013).
- [17] S. W. Hell and J. Wichmann, Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy, *Opt. Lett.* **19**, 780 (1994).
- [18] S. W. Hell and M. Kroug, Ground-state-depletion fluorescence microscopy: A concept for breaking the diffraction resolution limit, *Appl. Phys. B: Lasers Opt.* **60**, 495 (1995).
- [19] P. C. Maurer, J. R. Maze, P. L. Stanwix, L. Jiang, A. V. Gorshkov, A. A. Zibrov, B. Harke, J. S. Hodges, A. S. Zibrov, A. Yacoby *et al.*, Far-field optical imaging and manipulation of individual spins with nanoscale resolution, *Nat. Phys.* **6**, 912 (2010).
- [20] R. A. Fisher, Theory of statistical estimation, in *Mathematical Proceedings of the Cambridge Philosophical Society* (Cambridge University Press, Cambridge, 1925), Vol. 22, pp. 700–725.
- [21] H. Cramer, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, 1946).
- [22] C. R. Rao, Information and accuracy attainable in the estimation of statistical parameters, *Bull. Calcutta Math. Soc.* **37**, 81 (1945).
- [23] S. Ram, E. Sally Ward, and R. J. Ober, Beyond Rayleigh's criterion: A resolution measure with application to single-molecule microscopy, *Proc. Natl. Acad. Sci. USA* **103**, 4457 (2006).
- [24] L. Motka, B. Stoklasa, M. D'Angelo, P. Facchi, A. Garuccio, Z. Hradil, S. Pascasio, F. V. Pepe, Y. S. Teo, J. Řeháček *et al.*, Optical resolution from Fisher information, *Eur. Phys. J. Plus* **131**, 130 (2016).
- [25] M. Paúr, B. Stoklasa, Z. Hradil, L. L. Sánchez-Soto, and J. Rehacek, Achieving the ultimate optical resolution, *Optica* **3**, 1144 (2016).
- [26] J. Řeháček, Z. Hradil, B. Stoklasa, M. Paúr, J. Grover, A. Krzic, and L. L. Sánchez-Soto, Multiparameter quantum metrology of incoherent point sources: Towards realistic superresolution, *Phys. Rev. A* **96**, 062107 (2017).
- [27] Y. Zhou, J. Yang, J. D. Hassett, S. MohammadH. Rafsanjani, M. Mirhosseini, A. Nick Vamivakas, A. N. Jordan, Z. Shi, and R. W. Boyd, Quantum-limited estimation of the axial separation of two incoherent point sources, *Optica* **6**, 534 (2019).
- [28] A. B. Mikhalychev, B. Bessire, I. L. Karuseichyk, A. A. Sakovich, M. Unternährer, D. A. Lyakhov, D. L. Michels, A. Stefanov, and D. Mogilevtsev, Efficiently reconstructing compound objects by quantum imaging with higher-order correlation functions, *Commun. Phys.* **2**, 134 (2019).
- [29] M. Tsang, R. Nair, and X.-M. Lu, Quantum theory of superresolution for two incoherent optical point sources, *Phys. Rev. X* **6**, 031033 (2016).
- [30] M. Tsang, Subdiffraction incoherent optical imaging via spatial-mode demultiplexing, *New J. Phys.* **19**, 023054 (2017).
- [31] M. Paúr, B. Stoklasa, J. Grover, A. Krzic, L. L. Sánchez-Soto, Z. Hradil, and J. Řeháček, Tempering Rayleigh's curse with PSF shaping, *Optica* **5**, 1177 (2018).
- [32] M. Paúr, B. Stoklasa, D. Koutný, J. Řeháček, Z. Hradil, J. Grover, A. Krzic, and L. L. Sánchez-Soto, Reading out Fisher information from the zeros of the point spread function, *Opt. Lett.* **44**, 3114 (2019).
- [33] Y. L. Len, C. Datta, M. Parniak, and K. Banaszek, Resolution limits of spatial mode demultiplexing with noisy detection, *Int. J. Quantum Inform.* **18**, 1941015 (2020).
- [34] A. B. Mikhalychev, P. I. Novik, I. L. Karuseichyk, D. A. Lyakhov, D. L. Michels, and D. S. Mogilevtsev, Lost photon enhances superresolution, *npj Quantum Inf.* **7**, 125 (2021).
- [35] S. Vlasenko, A. B. Mikhalychev, I. L. Karuseichyk, D. A. Lyakhov, D. L. Michels, and D. Mogilevtsev, Optimal correlation order in superresolution optical fluctuation microscopy, *Phys. Rev. A* **102**, 063507 (2020).
- [36] S. Kurdzialek and R. Demkowicz-Dobrzański, Super-resolution optical fluctuation imaging—fundamental estimation theory perspective, *J. Opt.* **23**, 075701 (2021).
- [37] M. Tsang, Conservative classical and quantum resolution limits for incoherent imaging, *J. Mod. Opt.* **65**, 1385 (2018).
- [38] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Filtering Theory* (John Wiley and Sons, Inc., New York, 1968).
- [39] P. Stoica and R. L. Moses, On biased estimators and the unbiased Cramér-Rao lower bound, *Signal Process.* **21**, 349 (1990).
- [40] J. D. Gorman and A. O. Hero, Lower bounds for parametric estimation with constraints, *IEEE Trans. Inf. Theory* **36**, 1285 (1990).
- [41] Y. C. Eldar, Rethinking biased estimation: Improving maximum likelihood and the Cramér–Rao bound, *Fnt Signal Process.* **1**, 305 (2008).
- [42] J. O. Berger, On the inadmissibility of unbiased estimators, *Stat. Probab. Lett.* **9**, 381 (1990).
- [43] Y. C. Eldar, Minimum variance in biased estimation: Bounds and asymptotically optimal estimators, *IEEE Trans. Signal Process.* **52**, 1915 (2004).
- [44] Z. Ben-Haim and Y. C. Eldar, A lower bound on the Bayesian MSE based on the optimal bias function, *IEEE Trans. Inf. Theory* **55**, 5179 (2009).
- [45] A. O. Hero, J. A. Fessler, and M. Usman, Exploring estimator bias-variance tradeoffs using the uniform CR bound, *IEEE Trans. Signal Process.* **44**, 2026 (1996).
- [46] C. H. Hua, N. H. Clinthorne, S. J. Wilderman, J. W. LeBlanc, and W. L. Rogers, Quantitative evaluation of information loss for Compton cameras, *IEEE Trans. Nucl. Sci.* **46**, 587 (1999).
- [47] Y. C. Eldar, Uniformly improving the Cramér-Rao bound and maximum-likelihood estimation, *IEEE Trans. Signal Process.* **54**, 2943 (2006).
- [48] L. J. Meng and N. H. Clinthorne, A modified uniform Cramer-Rao bound for multiple pinhole aperture design, *IEEE Trans. Med. Imaging* **23**, 896 (2004).
- [49] L. J. Meng, N. H. Clinthorne, S. Skinner, R. V. Hay, and M. Gross, Design and feasibility study of a single photon emission microscope system for small animal I-125 imaging, *IEEE Trans. Nucl. Sci.* **53**, 1168 (2006).
- [50] L.-J. Meng and N. Li, A vector uniform Cramer-Rao bound for SPECT system design, in *2008 IEEE Nuclear Science Symposium Conference Record* (IEEE, Dresden, 2008), pp. 4054–4064.
- [51] E. T. Jaynes, *Probability Theory: The Logic of Science* (Washington University, St. Louis, 1996).

- [52] M. Vallisneri, Use and abuse of the Fisher information matrix in the assessment of gravitational-wave parameter-estimation prospects, *Phys. Rev. D* **77**, 042001 (2008).
- [53] A. Mikhalychev, D. Mogilevtsev, Y. S. Teo, J. Řeháček, and Z. Hradil, Bayesian recursive data-pattern tomography, *Phys. Rev. A* **92**, 052106 (2015).
- [54] E. L. Lehmann and G. Casella, *Theory of Point Estimation* (Springer Science & Business Media, New York, 1998).
- [55] K.-R. Koch, *Introduction to Bayesian Statistics* (Springer Science & Business Media, Berlin, 2007).
- [56] C. W. Helstrom, Quantum detection and estimation theory, *J. Stat. Phys.* **1**, 231 (1969).
- [57] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland Publishing Company, Amsterdam, 1982), Vol. 1.
- [58] M. G. A. Paris, Quantum estimation for quantum technology, *Int. J. Quantum Inform.* **07**, 125 (2009).
- [59] M. Tsang, Quantum limits to optical point-source localization, *Optica* **2**, 646 (2015).
- [60] R. Nair and M. Tsang, Far-field superresolution of thermal electromagnetic sources at the quantum limit, *Phys. Rev. Lett.* **117**, 190801 (2016).
- [61] Z. Dutton, R. Kerviche, A. Ashok, and S. Guha, Attaining the quantum limit of superresolution in imaging an object's length via predetection spatial-mode sorting, *Phys. Rev. A* **99**, 033847 (2019).
- [62] R. Demkowicz-Dobrzański, W. Górecki, and M. Guţă, Multi-parameter estimation beyond quantum fisher information, *J. Phys. A: Math. Theor.* **53**, 363001 (2020).
- [63] C. Lupo, Z. Huang, and P. Kok, Quantum limits to incoherent imaging are achieved by linear interferometry, *Phys. Rev. Lett.* **124**, 080503 (2020).
- [64] S. Zhou and L. Jiang, Modern description of Rayleigh's criterion, *Phys. Rev. A* **99**, 013808 (2019).
- [65] E. Bisketzi, D. Branford, and A. Datta, Quantum limits of localisation microscopy, *New J. Phys.* **21**, 123032 (2019).
- [66] A. Mikhalychev, K. Zhevno, S. Vlasenko, A. Benediktovitch, T. Ulyanenkova, and A. Ulyanekov, Fisher information for optimal planning of x-ray diffraction experiments, *J. Appl. Crystallogr.* **54**, 1676 (2021).
- [67] G. B. Lemos, V. Borish, G. D. Cole, S. Ramelow, R. Lapkiewicz, and A. Zeilinger, Quantum imaging with undetected photons, *Nature (London)* **512**, 409 (2014).
- [68] A. A. Pushkina, G. Maltese, J. I. Costa-Filho, P. Patel, and A. I. Lvovsky, Superresolution linear optical imaging in the far field, *Phys. Rev. Lett.* **127**, 253602 (2021).
- [69] J. Kiefer, Optimum experimental designs, *J. Roy. Stat. Soc. B Met.* **21**, 272 (1959).
- [70] S. P. Asprey and S. Macchietto, Designing robust optimal dynamic experiments, *J. Process Control* **12**, 545 (2002).
- [71] J. R. Banga, K. J. Versyck, and J. F. Van Impe, Computation of optimal identification experiments for nonlinear dynamic process models: a stochastic global optimization approach, *Ind. Eng. Chem. Res.* **41**, 2425 (2002).
- [72] D. Faller, U. Klingmüller, and J. Timmer, Simulation methods for optimal experimental design in systems biology, *Simulation* **79**, 717 (2003).
- [73] I. Karuseichyk, G. Sorelli, M. Walschaers, N. Treps, and M. Gessner, Resolving mutually-coherent point sources of light with arbitrary statistics, *Phys. Rev. Res.* **4**, 043010 (2022).
- [74] C. Rouvière, D. Barral, A. Grateau, I. Karuseichyk, G. Sorelli, M. Walschaers, and N. Treps, Ultra-sensitive separation estimation of optical sources, *Optica* **11**, 166 (2024).
- [75] I. N. Agafonov, M. V. Chekhova, T. Sh. Iskhakov, and L.-A. Wu, High-visibility intensity interference and ghost imaging with pseudo-thermal light, *J. Mod. Opt.* **56**, 422 (2009).
- [76] X.-H. Chen, I. N. Agafonov, K.-H. Luo, Q. Liu, R. Xian, M. V. Chekhova, and L.-A. Wu, Arbitrary-order lensless ghost imaging with thermal light, *Opt. Lett.* **35**, 1166 (2010).
- [77] P.-A. Moreau, E. Toninelli, P. A. Morris, R. S. Aspden, T. Gregory, G. Spalding, R. W. Boyd, and M. J. Padgett, Resolution limits of quantum ghost imaging, *Opt. Express* **26**, 7528 (2018).
- [78] D. E. Knuth, Big omicron and big omega and big theta, *SIGACT News* **8**, 18 (1976).
- [79] E. Bettens, D. Van Dyck, A. J. Den Dekker, J. Sijbers, and A. Van den Bos, Model-based two-object resolution from observations having counting statistics, *Ultramicroscopy* **77**, 37 (1999).
- [80] F. Albarelli, M. Barbieri, M. G. Genoni, and I. Gianani, A perspective on multiparameter quantum metrology: From theoretical tools to applications in quantum imaging, *Phys. Lett. A* **384**, 126311 (2020).
- [81] Y. Yang, G. Chiribella, and M. Hayashi, Attaining the ultimate precision limit in quantum state estimation, *Commun. Math. Phys.* **368**, 223 (2019).
- [82] O. E. Barndorff-Nielsen and R. D. Gill, Fisher information in quantum statistics, *J. Phys. A: Math. Gen.* **33**, 4481 (2000).