Quantum Multi-Parameter Adaptive Bayesian Estimation and Application to Super-Resolution Imaging

Kwan Kit Lee
Optical Sciences
The University of Arizona
Tucson, AZ 85721, USA
kklee@email.arizona.edu

Christos Hagatsos
Optical Sciences
The University of Arizona
Tucson, AZ 85721, USA
cgagatsos@optics.arizona.edu

Saikat Guha
Optical Sciences
The University of Arizona
Tucson, AZ 85721, USA
saikat@optics.arizona.edu

Amit Ashok
Optical Sciences
The University of Arizona
Tucson, AZ 85721, USA
ashok@optics.arizona.edu

Abstract—It is well known in Bayesian estimation theory that the estimator \( \hat{\theta} = E[\theta|l] \) attains the minimum mean squared error (MMSE) for estimating a scalar parameter of interest \( \theta \) from the observation of \( l \) through a noisy channel \( P_{l|\theta} \), given a prior \( P_{\theta} \) on \( \theta \). In quantum, e.g., optical and atomic, sensing tasks the user gets \( \rho_0 \), the quantum state that encodes \( \theta \). They choose a measurement, a positive-operator valued measure (POVM) \( \Pi_l \), which induces the channel \( P_{\theta,l} = \text{Tr}(\rho_0 \Pi_l) \) to the measurement outcome \( l \), on which the aforesaid classical MMSE estimator is employed. Personick found the optimum POVM (MMSE) for estimating a scalar parameter of interest from the observation of \( \theta \). However, there has been little work on adaptive measurement scheme for multi-parameter estimation in the quantum setting. In this paper, we build upon Personick’s result to construct a Bayesian adaptive measurement scheme for multi-parameter estimation when \( N \) copies of \( \rho_0 \) are available. We illustrate an application to localizing a cluster of point emitters in a highly sub-Rayleigh angular field-of-view, an important problem in fluorescence microscopy and astronomy. Our algorithm translates to a multi-spatial-mode transformation prior to a photon-detection array, with electro-optic feedback to adapt the mode sorter. We show that this receiver performs far superior to quantum-noise-limited focal-plane direct imaging.

I. INTRODUCTION

A sensing and imaging measurement design problem can be abstracted to having access to an \( N \)-copy quantum state \( \rho(\theta)^\otimes N \) encoding \( M \) parameters of interest, denoted by the vector \( \theta = [\theta_1, \theta_2, ..., \theta_M]^T \). The receiver realizes a quantum measurement, described by a positive-operator-valued measure (POVM) \( \{\Pi_l\} \), operating on a single copy of \( \rho(\theta) \), resulting in a vector-valued measurement outcome \( l = [l_1, l_2, ..., l_N]^T \). The choice of a POVM induces the classical measurement channel defined by the probability density \( p(l|\theta) = \text{Tr}(\rho(\theta)\Pi_l) \). When more than one copy of the quantum state is available (\( N \geq 2 \)), the receiver can: (1) in the most general setting, choose a joint-measurement POVM \( \{\Pi_{l(N)}\} \) acting collectively on \( \rho(\theta)^\otimes N \), producing the outcome \( l_{(N)} \); (2) employ the so-called local operations and classical communications (LOCC) scheme, where each copy of the state is measured by an independent measurement, where the POVM \( \{\Pi_l\} \) acting on the \( \tau^{th} \) copy of \( \rho(\theta) \) is chosen based on the information available from the previous set of measurement outcomes \( \{l^{(1)}, l^{(2)}, ..., l^{(\tau−1)}\} \), \( 1 \leq \tau \leq N \); or (3) use independent identical measurements on each copy of the state, described by the POVM \( \{\Pi_l\} \). No matter which strategy the receiver may use, after measuring all \( N \) copies, the receiver generates an estimate of \( \theta \), i.e., \( \hat{\theta}(l_{\text{set}}) \) where \( l_{\text{set}} = l_{(N)} \) for case (1) above, and \( l_{\text{set}} = [l^{(1)}, l^{(2)}, ..., l^{(N)}] \) for cases (2) and (3) above. The receiver chooses the estimator to optimize a desired objective/loss function. A natural choice of the objective function associated with sensing and imaging tasks is mean (expected) squared-error (MSE), \( E[||\hat{\theta} − \theta||^2] \).

For any given measurement POVM \( \{\Pi_l\} \), assuming strategy (3) above, i.e., the same measurement acts on each copy of \( \rho(\theta) \), the problem reduces to the standard classical estimation theory problem of estimating \( \theta \) from \( N \) i.i.d. samples of \( l \), each produced by \( p(l|\theta) \). The covariance \( \text{Cov}(\hat{\theta}(l_{\text{set}}), \theta) \) for any unbiased estimator \( \hat{\theta} \) of \( \theta \) is lower bounded by \( \Sigma_C \). This means \( \text{Cov}(\hat{\theta}(l_{\text{set}}), \theta) − \Sigma_C \) is a semi-positive definite matrix, denoted compactly as \( \Sigma_C = \text{Cov}(\hat{\theta}(l_{\text{set}}), \theta) \geq \Sigma_C \). The receiver’s task is to pick the optimal estimator \( \hat{\theta}^\text{opt}(l_{\text{set}}) \) on the measurement outcomes \( l_{\text{set}} \), such that \( \text{Cov}(\hat{\theta}^\text{opt}(l_{\text{set}}), \theta) \) saturates the bound \( \Sigma_C \) when possible.

Tools of quantum estimation theory let us find a tighter lower bound to \( \text{Cov}(\hat{\theta}(l), \theta) \), which automatically optimizes over all physically-permissible choices of a POVM \( \{\Pi_l\} \) (again, assuming that the same measurement is used to detect each copy of \( \rho(\theta) \)). The \( \text{Cov}(\hat{\theta}(l), \theta) \) is lower bounded by \( \Sigma_Q \) (a quantum bound), which itself is an infimum of all bounds \( \Sigma_C \) associated with all possible choices of \( \{\Pi_l\} \). For certain cases (for example when \( \theta \) is a single scalar parameter),
quantum estimation theory also tells us the optimal receiver POVM \( \{\Pi_{l}^{opt}\} \). Once the optimal receiver is chosen, it uses the optimal estimator \( \hat{\bf{\theta}}^{opt}(l_{set}) \) using standard classical estimation tools, such that \( \text{Cov}(\hat{\bf{\theta}}^{opt}(l_{set}), \bf{\theta}) \) saturates \( \Sigma_{Q} \) when permissible. In general \( \text{Cov}(\hat{\bf{\theta}}(l_{set}), \bf{\theta}) \geq \Sigma_{C} \geq \Sigma_{Q} \), where \( \Sigma_{C} \) could correspond to any POVM choice.

The aforementioned lower bounds on the covariance of multi-parameter estimators can be defined within the statistical inference frameworks the frequentist, i.e., Fisherian (with no prior), or the Bayesian (with prior \( p(\bf{\theta}) \)) inference settings. We review below some known bounds in both the settings.

In the Fisherian (frequentist) framework, when no prior \( p(\bf{\theta}) \) is available, the Cramer-Rao lower bound (CRLB) \( \Sigma_{C} \) on the covariance \( \text{Cov}(\hat{\bf{\theta}}(l), \bf{\theta}) \) of an unbiased estimator is given by the inverse of the Fisher information (FI) matrix \( I(\bf{\theta}) \):

\[
I_{ij} = \int \left[ \frac{\partial}{\partial \theta_i} \ln p(l|\bf{\theta}) \right] \left[ \frac{\partial}{\partial \theta_j} \ln p(l|\bf{\theta}) \right] p(l|\bf{\theta}) dl, \tag{1}
\]

with \( 1 \leq i, j \leq M \), and the likelihood \( p(l|\bf{\theta}) = \text{Tr}(\rho(\bf{\theta})\Pi_{l}) \). The quantum version of this lower bound \( \Sigma_{Q} \), which only depends on \( \rho(\bf{\theta}) \) (since the measurement \( \Pi_{l} \) is automatically optimized over) is given by the inverse of the quantum Fisher information (QFI) matrix \( Q(\bf{\theta}) \), the elements of which are:

\[
Q_{ij} = \frac{1}{2} \text{Tr} \left[ \rho(\bf{\theta}) L_{i}L_{j} + L_{j}L_{i} \right], \tag{2}
\]

where \( L_{i} \) is the symmetric logarithmic derivative (SLD) operator which can be evaluated from the implicit relationship:

\[
2\frac{\partial}{\partial \theta_i} \rho(\bf{\theta}) = \rho(\bf{\theta}) L_{i} + L_{i} \rho(\bf{\theta}), \tag{3}
\]

with \( 1 \leq i \leq M \). Thus, we have \( \text{Cov}(\hat{\bf{\theta}}(l_{set}), \bf{\theta}) \geq I^{-1} \geq Q^{-1} \) in the Fisher framework. For \( \text{N}-\text{copy}\ i.d\ . \text{measurement of } \rho(\bf{\theta})^{\otimes N} \), both the classical and quantum bounds above are lowered by a factor of \( 1/N \). The classical one is asymptotically attained by the maximum likelihood estimator (MLE). The quantum CRLB \( (Q^{-1}) \) is not saturable in general for \( M > 1 \).

The corresponding Bayesian lower bounds on the covariance \( \text{Cov}(\hat{\bf{\theta}}(l), \bf{\theta}) \) of any estimator \( \hat{\bf{\theta}}(l) \) are found in [4]. Given a prior \( p(\bf{\theta}) \) on the parameter vector \( \bf{\theta} \), the Bayesian Cramer-Rao lower bound (BCLRBB) \( \Sigma_{C} \) is given by:

\[
\Sigma_{C} = \int p(\bf{\theta}) \bf{\theta}^{T} d\bf{\theta} - J, \tag{4}
\]

where the \( M \)-by-\( M \) matrix \( J \) is defined as:

\[
J_{ij} = \int \left[ \frac{\partial}{\partial \theta_i} \ln p(l|\bf{\theta}) \right] \left[ \frac{\partial}{\partial \theta_j} \ln p(l|\bf{\theta}) \right] p(l) dl, \tag{5}
\]

and \( p(l, \bf{\theta}) = p(l|\bf{\theta})p(\bf{\theta}) \) is the joint distribution of \( l \) and \( \bf{\theta} \). For the quantum version of this lower bound, we first define the following operators, for \( 1 \leq i \leq M \) and \( k = 0, 1, 2 \) [11]::

\[
\Gamma_{i,k} = \int d\bf{\theta} p(\bf{\theta}) \rho(\bf{\theta}) \theta_{i}^{k}, \tag{6}
\]

and operators \( B_{i}, 1 \leq i \leq M \), that satisfy:

\[
2\Gamma_{i,1} = \Gamma_{0} B_{i} + B_{i} \Gamma_{0}. \tag{7}
\]

For \( k = 0 \), \( \Gamma_{i,0} = \Gamma_{0} \delta_{i,0} \), \( \forall (i,j) \), thus we can drop the first index and denote it as \( \Gamma_{0} = \int d\bf{\theta} p(\bf{\theta}) \rho(\bf{\theta}) \), the average received state. The quantum BCLRBB \( \Sigma_{Q} \) can be written as:

\[
\Sigma_{Q} = \int p(\bf{\theta}) \bf{\theta}^{T} d\bf{\theta} - G, \tag{8}
\]

where

\[
G_{ij} = \frac{1}{2} \text{Tr} \left[ \Gamma_{0} B_{i} B_{j} + B_{j} B_{i} \right]. \tag{9}
\]

Thus in a Bayesian inference framework, we have \( \text{Cov}(\hat{\bf{\theta}}(l), \bf{\theta}) \geq \Sigma_{C} \geq \Sigma_{Q} \). Table I summarizes these bounds.

| Table I | Summary of Classical and Quantum CRBs. |
| --- | --- |
| Classical | Information Matrix: \( I \) | Covariance Matrix: \( \Sigma_{C} \) |
| Quantum | Information Matrix: \( Q \) | Covariance Matrix: \( \Sigma_{Q} \) |
| Bounds | \( \text{Cov}(\bf{\theta}, \bf{\theta}) \geq I^{-1} \geq Q^{-1} \) | \( \text{Cov}(\bf{\theta}, \bf{\theta}) \geq \Sigma_{C} \geq \Sigma_{Q} \) |

To achieve the quantum bound, an optimal measurement is required (i.e. an optimal choice POVM, that acts on each copy of \( \rho(\bf{\theta}) \)). For a single parameter problem \( (M = 1) \), the projective measurement onto the eigenvectors of the SLD operator \( L \) in Eq. (3) saturates the Fisher bound, i.e., the \( I \) for the SLD measurement equals \( Q \). Likewise, the Bayesian quantum bound on the covariance is saturated (i.e. \( \Sigma_{C} = \Sigma_{Q} \)), for the case of a single parameter \( (M = 1) \) by a projective measurement onto the eigenvectors of the operator \( B \) in Eq. (7) [11].

For multi-parameter estimation, if the operators associated with parameter \( \theta_{i} \), \( L_{i} \), and \( B_{i} \), \( 1 \leq i \leq M \) commute with one another, for the Fisher and Bayesian frameworks respectively, the corresponding covariance bound can be saturated by the above-said measurements, calculated by evaluating the eigenvectors of \( L_{i} \) or \( B_{i} \), respectively (which \( i \) does not matter as they are simultaneously diagonal). However, if the operators do not commute, which is the case in general, a measurement that is jointly optimal for all parameters may not exist and likely to be challenging to derive.

In the quantum case, the Holevo Cramer-Rao bound (HCRB) [5] is the most fundamental scalar lower bound on the weighted mean square error \( \text{Tr}[W \text{Cov}(\hat{\bf{\theta}}(l), \bf{\theta})] \), for a positive definite \( W \). The HCRB represents the best precision attainable with a collective measurement (discussed as case (1) above) on an asymptotically large number of identical copies of \( \rho(\bf{\theta}) \).

In this paper, we propose a sequential adaptive measurement scheme (an LOCC) within a full Bayesian inference framework. We leverage tools from Bayesian quantum estimation theory. The details of our scheme are discussed in Sec. III. In Sec. III, we employ our measurement scheme to the problem of localizing an unknown number of point-emitters placed in a sub-Rayleigh (below diffraction-limit)
field of view in an optical imaging context. This application is motivated by the fact that traditional direct focal-plane imaging, which employs intensity measurements followed by electronic-domain processing, is known to be highly sub-optimal [6] in the sub-Rayleigh regime. We compare our quantum-inspired adaptive sequential measurement design with the direct imaging technique to quantify the significant performance improvement in optical resolution achieved by our scheme.

II. ADAPTIVE SEQUENTIAL MEASUREMENT SCHEME

Consider a system in the state described by the density operator:

\[
\rho'(\theta) = \sum_{i=1}^{P} b_i(\theta) |\psi_i(\theta)\rangle \langle \psi_i(\theta)|,
\]

where \( \theta = [\theta_1, \theta_2, ..., \theta_M]^T \) are the parameters of interest, \(|\psi_i(\theta)\rangle\) and \(b_i(\theta)\) are the parameter-dependent pure states and the corresponding weights respectively. \( P \) itself, in general, is an unknown parameter (positive integer) such that: \( P_{\text{min}} \leq P \leq P_{\text{max}} \). Here we assume that \( P \) is upper bounded by \( P_{\text{max}} \), i.e., a prior on \( P \). If the lower bound \( P_{\text{min}} \) is not known (available), we can set it to 1. When \( P_{\text{min}} \neq P_{\text{max}} \), both \( P \) and \( \theta \) need to be estimated. On the contrary, if \( P_{\text{min}} = P = P_{\text{max}} \), i.e., \( P \) is known a priori exactly, then we only need to estimate the parameters \( \theta \).

A single copy of a quantum state \( \rho'(\theta) \), defined in Eq. (10), can be measured with a POVM \( \{\Pi_i\} \) such that the probability of getting a scalar outcome \( l \) is \( p(l|\theta) = Tr[\rho'(\theta)\Pi_l] \). Now, let \( K \) denote the block-length (of copies of \( \rho'(\theta) \)) over which the measurement \( \Pi_l \) stays the same. The density operator \( \rho'(\theta) = \rho'(\theta) \otimes K \) and the probability of observing the measurement outcome \( l = [l_1, l_2, ..., l_K]^T \) is \( p(l|\theta) = Tr[\rho'(\theta)\Pi_l] = \prod_{i=1}^{K} Tr[\rho'(\theta)\Pi_{l_i}] \), where \( \Pi_l \equiv \Pi_{l_1} \otimes \cdots \otimes \Pi_{l_K} \).

A. Measurement Scheme

To illustrate the proposed scheme, we begin with the \( P \) known exactly case. In the next section, we discuss an extension of this scheme where we relax this prior on \( P \). Let us take \( N = K \times S \) to be the total number of copies of \( \rho'(\theta) \) available to us. We adapt the measurement choice \( S \) times, denoted by \( \tau \) as the iteration index, \( 1 \leq \tau \leq S \). Thus, at the \( \tau \)th iteration of the sequential measurement, a single copy of \( \rho'(\theta) \) is measured with the POVM \( \{\Pi_{l_k}^{(\tau)}\} \) yielding the outcome vector \( l^{(\tau)} = [l_1^{(\tau)}, l_2^{(\tau)}, ..., l_K^{(\tau)}]^T \). At the end of the sequential measurement scheme, a \( S \)-copy state \( \rho'(\theta) \otimes S \) has been measured. The parameter estimate \( \hat{\theta}^{(\tau)} \), after the \( \tau \)th sequential measurement is denoted by \( \hat{\theta}^{(\tau)} = [\hat{\theta}_1^{(\tau)}, \hat{\theta}_2^{(\tau)}, ..., \hat{\theta}_M^{(\tau)}]^T \). In a Bayesian inference setting, the parameter estimate \( \hat{\theta}^{(\tau)} \) is given by posterior mean: \( \hat{\theta}^{(\tau)} = E_{p(\theta|l^{(\tau)})}[\theta] \). The posterior \( p(\theta|l^{(\tau)}) = p(l^{(\tau)}|\theta) \cdot p(\theta)/p(l^{(\tau)}) \), where \( p(l^{(\tau)}|\theta) \) is the prior at the \( \tau \)th iteration. Note that the prior \( p(l^{(\tau)}|\theta) \) in turn equals the posterior \( p(\theta|l^{(\tau-1)}) \) at the previous \( (\tau-1) \)th iteration. The density operator at the \( \tau \)th iteration is represented as \( \rho^{(\tau)}(\theta) \). Now what remains to be determined is how we choose the POVM \( \{\Pi_{l_k}^{(\tau)}\} \) at the \( \tau \)th iteration.

It is known that for a single parameter estimation problem, the eigen-projection measurement of \( B_1 \) in Eq. (7) saturates the quantum bound \( \Sigma_Q \), which reduces to a lower bound of the variance of the scalar parameter. The minimum mean square error (MMSE) is given by \( \Sigma_Q = Tr[\Gamma_{1,2} - B_1 \Gamma_{1,1}] \), where \( \Gamma_{i,j} \) are defined in Eq. (6). We refer to this measurement as the Personick projection in this work. For the multi-parameter problem, the counter-part of \( Tr[B_1 \Gamma_{1,1}] \) is a matrix \( G \) in Eq. (9). If all \( B_i \) operators commute, the quantum optimal measurement is given by the eigen-projections onto any of the \( B_i \) operators [4], however, there is no such guarantee that the optimal measurement for all parameters exists or can be found. At the \( \tau \)th iteration of sequential measurement we propose to use a single parameter \( \gamma^{(\tau)} \), defined as a linear combination of the \( M \) parameters given by the eigenvector of the matrix \( G \) with the smallest eigenvalue. Note that this matrix \( G \) is defined per Eq. (9) for the density operator \( \rho^{(\tau-1)}(\theta) \). The scalar parameter \( \gamma^{(\tau)} \) is used to construct the operator \( B_{\gamma^{(\tau)}} \). The corresponding Personick projection constructed using \( B_{\gamma^{(\tau)}} \) is chosen as the POVM \( \{\Pi_{l_k}^{(\tau)}\} \) at the \( \tau \)th iteration.

The sequential measurements are terminated when we have used all the \( N \) available copies of \( \rho'(\theta) \).

B. Extension: \( P \) not known a priori

If the scalar \( P \) in Eq. (10) is unknown, we can employ and initialize multiple models of density operators \( \rho(\theta_P) \) with the corresponding prior \( p(\theta_P) \), where \( \theta_P = [\theta_1, \theta_2, ..., \theta_{MP}]^T \) for \( P_{\text{min}} \leq P \leq P_{\text{max}} \). In such scenario, the number of parameters, denoted by \( MP \), for each model corresponding to a \( P \) can be different in general. In \( \tau \)th iteration of the sequential measurement, one model is selected and used to construct the Personick measurement. The model can be selected randomly at \( \tau = 0 \), and the one that maximizes \( p(l^{(\tau-1)}) \) can be used for the \( \tau \)th measurement iteration. After model selection, the measurement scheme defined in the previous section can be applied. Note that at \( \tau \)th iteration, not only selected model but all the models are updated, in a Bayesian inference setting, using the measurement outcome \( l^{(\tau)} \). When the sequential measurements eventually terminate, we simply pick a model using the same model selection criteria described above and compute the final multi-parameter estimate as the posterior mean.

III. APPLICATION: MULTI POINT-EMITTER ESTIMATION

A. Formulation

We apply the adaptive sequential measurement scheme to estimate the location and relative brightness of incoherent point emitters in a cluster/constellation per the formulation in [6]. The quantum state of photons incident on the image point emitters in a cluster/constellation per the formulation in [6]. The quantum state of photons incident on the image point emitters in a cluster/constellation per the formulation in [6].
where $|0\rangle$ is the vacuum state, $\rho'$ is the single photon state density operator, which has the form of Eq. (10), and $\epsilon$ is the total number of photons arriving on the image plane within the coherence time of the source. Assuming that $\epsilon \ll 1$ (valid for weak thermal source), the photon states with order $O(\epsilon^2)$ are negligible. As the vacuum state $|0\rangle$ provides no information, we can focus on $\rho'$. Thus, the components of Eq. (10) have the following meaning: $P$ is the number of emitters, $\{b_i\}_{i=1}^P$ are the hyper-parameters of the Dirichlet distribution. Thus, $p(x,y)$ are the relative brightness of each source (sum normalized to 1) and the states $|\psi_i\rangle$ are given by:

$$|\psi_i\rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi(x-x_i, y-y_i)|x,y\rangle dxdy,$$

(12)

such that $(x_i, y_i)$ are the scaled coordinates of the $i^{th}$ point source on the image plane. The point spread function (PSF) $\psi(x,y)$ of the imaging system is set to a 2D Gaussian function:

$$\psi(x,y) = \frac{1}{\sqrt{2\pi\sigma_x\sigma_y}} \exp\left(-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}\right),$$

(13)

where $\sigma_x$ and $\sigma_y$ are the standard deviation (a measure of width) of the PSF in $x$ and $y$ direction respectively. For a given PSF, $\sigma_x$ and $\sigma_y$ are known parameters and set to $\sigma_x = \sigma_y$ in our study. We define the full width at half maximum (proportional to $\sigma_x$) of the PSF as Rayleigh length ($\text{rl}$) in our analysis.

The parameters of interest in this problem are the position and relative brightness of the $P$ emitters, i.e. $\theta = [x_1, \ldots, x_P, y_1, \ldots, y_P, b_1, \ldots, b_P]^T = [x, y, b]^T$. For the positions $[x,y]^T$, we use independent Gaussian distribution $\mathcal{N}$ as prior:

$$p(x,y) = \prod_i \mathcal{N}(x_i; \bar{x}_i, \sigma_{x_i})\mathcal{N}(y_i; \bar{y}_i, \sigma_{y_i}),$$

(14)

where for $1 \leq i \leq P$, $\bar{x}_i, \bar{y}_i, \sigma_{x_i}, \sigma_{y_i}$ are the mean and standard deviation of the position parameters $x_i$ and $y_i$ respectively.

For the brightness $b^0$ parameters a Dirichlet distribution is used as prior: $p(b) = \text{Dir}(b|\alpha)$, where $\alpha = [\alpha_1, \ldots, \alpha_P]^T$ are the hyper-parameters of the Dirichlet distribution. Thus, the overall prior is expressed as: $p(x,y,b) = p(x,y)p(b)$.

Now we have defined all relevant detail (i.e., photon state density operator, prior distribution) for the proposed adaptive sequential measurement scheme described in the previous section. As $p(x,y,b)$ is not a conjugate prior for the Poisson likelihood, we update the hyper-parameters of the prior distribution at $\tau^{\text{th}}$ iteration to derive the posterior, which assumes the role of the prior in the next $(\tau + 1)^{\text{th}}$ iteration. The prior hyper-parameters are: $h = [\bar{x}_1, \ldots, \bar{x}_P, \bar{y}_1, \ldots, \bar{y}_P, \sigma_{x_1}, \ldots, \sigma_{x_P}, \sigma_{y_1}, \ldots, \sigma_{y_P}, a_1, \ldots, a_P, \delta]^T = [\bar{x}, \bar{y}, \sigma_x, \sigma_y, \alpha, \delta]^T$. Here, $\delta$ is another hyper-parameter associated with the brightness prior distribution explained later.

To update the hyper-parameters of the prior position at the $(\tau + 1)^{\text{th}}$ iteration, we use the first- and the second-moments of the posterior distribution at the $\tau^{\text{th}}$ iteration:

$$\hat{a}_i^{(\tau+1)} = \int \alpha_i p(\theta^{(\tau)}|l^{(\tau)}; h^{(\tau)})d\theta^{(\tau)},$$

(15)

$$\hat{\sigma}_i^2(\tau+1) = \int [\alpha_i - \alpha_i^{(\tau+1)}]^2 p(\theta^{(\tau)}|l^{(\tau)}; h^{(\tau)})d\theta^{(\tau)},$$

(16)

where $\alpha$ represents $x$ or $y$ co-ordinate.

For the hyper-parameters $a^T$ of the brightness prior, expectation maximization approach is used. We first find the mean of the brightness vector as:

$$\hat{b}_i^{(\tau+1)} = \int b_i p(\theta^{(\tau)}|l^{(\tau)}; h^{(\tau)})d\theta^{(\tau)}.$$

(17)

Then, $a^T$ is updated such that $\hat{b}^{(\tau+1)}$ becomes the mode of the distribution:

$$a^{(\tau+1)} = \hat{b}^{(\tau+1)}\left[a_0^{(\tau)} + \delta^{(\tau)} - P\right] + 1 = \hat{b}^{(\tau+1)}[a_0^{(\tau)} - P] + 1,$$

(18)

where $a_0^{(\tau)} = \sum_i a_i^{(\tau)}$ and $a_0^{(\tau+1)} = a_0^{(\tau)} + \delta^{(\tau)}$. Roughly speaking, the larger the $a_0^{(\tau)}$, the smaller the total variance of the Dirichlet distribution. Adding $\delta^{(\tau)} \geq 0$ leads to $a_0^{(\tau+1)} \geq a_0^{(\tau)}$, such that the variance reduces monotonically with each iteration $\tau$. Note that the introduction of $\delta^{(\tau)}$ does not change the position of the mode in the distribution. We set $\delta^{(\tau)}$ to a constant for all $\tau$.

B. Simulation Results

We demonstrate the performance of the proposed adaptive sequential measurement scheme for 10 distinct realizations of 4 point emitter constellations. The position of the point emitters are generated inside a circle with radius of 0.25 rl (Rayleigh length) with the minimum emitter separation set to 0.1 rl. The relative brightness are chosen to be uniform. The total photon budget is set to $5 \times 10^5$ and each adaptive sequential step utilizes $10^4$ photons. The adaptive sequential scheme is initialized by employing 1000 photons for a direct imaging measurement followed by expectation maximization (EM) algorithm to estimate the initial model parameters. The remaining photons are detected by using Personick projection measurement in each adaptive sequential step.

For traditional direct imaging (baseline), Richardson-Lucy deconvolution algorithm is first used deconvolve the blurred image followed by the k-mean clustering algorithm to find the position and relative brightness of identified point emitters. If $P$ is known a priori, the k-mean clustering algorithm is used once for $P$; otherwise the algorithm is used for all $P_{\text{min}} \leq P \leq P_{\text{max}}$ and the model ($P$) that produces the smallest reconstruction error.

C. Estimation with $P$ known exactly

For each of the 10 constellations, we employ 100 Monte Carlo simulation. Fig. 1 shows an illustrative realization of the point emitter cluster and estimated location and brightness using the two measurement schemes.
Fig. 1. An illustrative example of a cluster of point emitter estimated with direct imaging and Personick projection measurements, when $P$ is known exactly. The black dots, blue circles and red squares correspond to the ground truth, estimates from direct imaging and Personick projection measurements respectively. The marker size is proportional to the point emitter brightness.

To obtain the average performance of the proposed measurement scheme, for each point emitter realization, we first pair the ground truth point emitter location with the estimated locations, such that the sum of the position errors, which is defined as: 

$$
\sum_{i=1}^{P} \sqrt{(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2}
$$

over all point-source matched pairs is minimized. The average (over all emitters) position error distribution of the point emitters is shown in Fig. 2. We observe that the proposed adaptive scheme outperforms the direct imaging. More specifically, the mean position error obtained by the adaptive scheme is four-fold lower than that of the direct imaging.

Fig. 2. Distribution of the point emitter position errors obtained with the two measurements.

D. Estimation with unknown $P$

When $P_{\text{max}} = 6$ is given as a prior, the estimation algorithm has to also estimate $P$. For the same set of constellations and using the same number of simulations, the distribution of number of point emitters estimated by the two measurement schemes in shown in Fig. 3. We observe that when $P$ is estimated correctly the proposed adaptive scheme maintains the significant performance advantage over direct imaging.

Fig. 3. Distribution of the number of point emitters estimated by the two measurements.

IV. CONCLUSION

Based on quantum estimation theory, we have developed a quantum adaptive Bayesian multi-parameter estimation scheme. We applied our proposed approach to super-resolving point emitters in an imaging application and demonstrated its superiority to the state-of-the-art direct imaging approach. This demonstrates that quantum estimation theory applied to sensing and imaging measurement design problems can yield significant gains by maximizing the photon information extraction with novel optical measurements.

REFERENCES

[1] S. D. Personick, “Application of quantum estimation theory to analog communication over quantum channels,” IEEE Trans. Inf. Theory, vol. 17, pp. 240–246, 1971.
[2] S. M. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory. Prentice Hall, 1997.

[3] J. Liu, H. Yuan, X.-M. Lu, and X. Wang, “Quantum Fisher information matrix and multiparameter estimation,” Journal of Physics A: Mathematical and Theoretical, vol. 53, no. 2, p. 023001, Dec 2019. [Online]. Available: https://doi.org/10.1088/1751-8121/ab5d4d

[4] J. Rubio and J. Dunningham, “Bayesian multiparameter quantum metrology with limited data,” Physical Review A, vol. 101, no. 3, p. 032114, 2020.

[5] A. Holevo, Probabilistic and Statistical Aspects of Quantum Theory. Edizioni della Normale, 2011.

[6] M. Tsang, R. Nair, and X.-M. Lu, “Quantum theory of superresolution for two incoherent optical point sources,” Phys. Rev. X, vol. 6, p. 031033, Aug 2016. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevX.6.031033

[7] M. DeGroot and M. DEGROOT, Optimal Statistical Decisions, ser. McGraw-Hill series in probability and statistics. McGraw-Hill, 1969. [Online]. Available: https://books.google.com/books?id=39UznQEACAAJ

[8] W. H. Richardson, “Bayesian-based iterative method of image restoration,” J. Opt. Soc. Am., vol. 62, no. 1, pp. 55–59, Jan 1972. [Online]. Available: http://www.osapublishing.org/abstract.cfm?URI=josa-62-1-55

[9] T. M. Kodinariya and P. R. Makwana, “Review on determining number of cluster in k-means clustering,” International Journal, vol. 1, no. 6, pp. 90–95, 2013.