arXiv:1612.04828v2 [quant-ph] 17 Jul 2017

Optimal Quantum Metrology of Distant Black Bodies

Mark E. Pearce, Earl T. Campbell, and Pieter Kok

Department of Physics & Astronomy, University of Sheffield, Sheffield S3 7RH, United Kingdom

July 18, 2017

Measurements of an object's temperature are important in many disciplines, from astronomy to engineering, as are estimates of an object's spatial configuration. We present the quantum optimal estimator for the temperature of a distant body based on the black body radiation received in the far-field. We also show how to perform separable quantum optimal estimates of the spatial configuration of a distant object, i.e. imaging. In doing so we necessarily deal with multi-parameter quantum estimation of incompatible observables, a problem that is poorly understood. We compare our optimal observables to the two mode analogue of lensed imaging and find that the latter is far from optimal, even when compared to measurements which are separable. To prove the optimality of the estimators we show that they minimise the cost function weighted by the quantum Fisher information—this is equivalent to maximising the average fidelity between the actual state and the estimated one.

1 Introduction

The quantum Fisher information (QFI) is a prevalent figure of merit in the field of quantum parameter estimation [5, 16, 18, 19, 21]. The quantum analogue of the Cramér-Rao bound provides a lower bound to the covariance matrix of the parameters to be estimated, although the attainability of this bound is not guaranteed as for classical statistics [12]. As in the classical case [17], the quantum Fisher information for states that follow a Gaussian distribution takes on a closed form that depends only on the first and second moments and their derivatives [7, 25, 34]. The fact that thermal states are so commonly found in nature and exhibit Gaussian statistics makes them the ideal testbed for Gaussian quantum estimation problems [23, 26].

Mark E. Pearce: mrkprc1@gmail.com
Earl T. Campbell: earltcampbell@gmail.com
Pieter Kok: p.kok@sheffield.ac.uk

The Cramér-Rao bound gives a lower bound on the covariance matrix of unbiased parameter estimates by the inverse of the Fisher information matrix [17]. Physically, this describes the relationship between the information obtained about a parameter via a measurement and the uncertainty in an estimate of the parameter from the measurement data. Similarly, the quantum Cramér-Rao bound gives a lower bound on estimation uncertainty via the inverse of the QFI [13]. Since the QFI is a property of the state alone and does not depend on a particular measurement scheme, the precision in parameter estimates is determined by the uncertainty in the state only [4], and is therefore fundamental in nature and cannot be reduced by improving measurement apparatus.

In classical metrology, the extension to multiple parameters does not entail a significant alteration to the basic theory. However, in the quantum theory, the possibility of observables being incompatible leads to additional complications [31]. The task of finding optimal observables under such circumstances is far from trivial and may involve collective measurements over many independent copies of the system, which is experimentally challenging. In this paper we consider only separable measurements (each system measured independently) and attempt to find the optimal observables among this class of observables.

Thermometry is an important method for interrogating the physical world. The temperature of astronomical objects reveals important clues about their nature, e.g. the cosmic microwave background [6] and estimates of effective stellar temperatures [1]; in both engineering and living systems, the temperature of components provides an essential diagnostic tool. The temperature of an object can be measured in multiple ways [24], for example via estimating the heat flow in direct-contact measurements, or by remotely measuring the radiation field emitted by the object. Typically, the thermal radiation field emitted is a black body spectrum. In this paper we consider the latter method, and present optimal quantum estimators for the temperature of black

body emitters.

Imaging provides us with important information about the spatial configuration of an object. Typically, imaging techniques are not studied for their optimality. Even in high resolution imaging, the focus tends to be towards increasing the resolution beyond the diffraction limit and rarely are optimal estimators considered. The main obstacle in applying the powerful techniques of metrology (both classical and quantum) to imaging is in identifying the parameters that constitute an image. Optimal metrology for an object of known parameterisation can be performed [20, 29, 32, 33] but for an object with an unknown spatial configuration it is not obvious what parameters we need to estimate. We show that the spatial configuration of a thermally radiating source is determined by spatial correlations in the far-field. In this paper we show how to optimally estimate these parameters, which completely determine the state, therefore mapping the problem of imaging to state estimation.

The density matrix of blackbody radiation is constructed from tensor products of many independent thermal $\mathbf{k}$-modes [22]:

$$\rho = \bigotimes_{\mathbf{k}} \rho_{\mathbf{k}} \,, \tag{1}$$

where the tensor product runs over all $\mathbf{k}$ of the form $\mathbf{k} = \Delta\mathbf{k}(x, y, z)$ where x, y, z are non-negative integers. Note we coarse grain $\mathbf{k}$-space rather than working with a continuum of modes. Physically, we motivate this coarse graining by appealing to the finite size of any realisable detector.

The spatial properties of the source are usually unknown, and without knowledge of them we typically cannot make accurate temperature measurements. A simple example helps to clarify the problem: consider the estimation of a star's temperature. If we simply measure the photon count at a single frequency in the far-field, we cannot distinguish between hotter sources and those with a larger angular size since an increase in either parameter will lead to a larger photon count. Therefore, if we wish to measure the temperature of the source, we must either know the angular size of the source or attempt to estimate it by measuring more than one frequency mode.

If we are interested in the temperature of the source alone, then the radiating area—or more precisely the solid angle of the source—is a nuisance parameter. Regardless of whether we are interested in the solid angle of the source, it is necessary to estimate it in order to provide optimal estimates of the temperature. The converse is also generally true. However, as we will show, it is not possible to make measurements of arbitrary spatial properties of the source by measuring spectral modes alone. In fact, it is necessary to consider spatially separated modes, where measurements of the correlations provide information about the spatial configuration of the source.

A blackbody emits at all frequencies with an intensity determined by the Planck distribution. This distribution can be considered as an infinite number of independent spectral modes with a spectral width inversely proportional to the time over which the state is observed [22]. It is therefore quite straightforward to determine the statistics of these independent modes and calculate the quantum Fisher information to determine the optimal measurements for simultaneous temperature and solid angle estimates. Considering multiple spatial modes is somewhat more cumbersome as the spatial modes are not generally independent. Here, by appealing to the fundamental quantum mechanical description of blackbody radiation, we demonstrate that it is the correlations between modes that convey the information about the spatial configuration of the source to the far-field.

It has been stated before that the optimal measurement for temperature estimates is photon number counting, [11, 27]. However, these results rely upon the assumption that we can measure a single thermal mode in the far-field, an assumption that we claim cannot be satisfied without knowledge of the exact spatial properties of the source. This is due to the transverse coherence area of the far-field radiation being determined by the spatial properties of the source [22], therefore to guarantee we are measuring exactly one mode, our detectors must be designed with these properties in mind. In this paper, the state ρ we consider is defined in Fig. 1. The state describes the radiation occupying a small volume $A_\rho \times c\tau$ in the far-field of a black body source, where A_ρ is the transverse area of the state and τ is the observation time.

An important assumption is that the transverse area of the state A_ρ is much smaller than the coherence area of the source. The effect is that we observe a fraction of each spectral mode, therefore ensuring that the transverse coherence is approximately equal to one across the entire area A_ρ. This allows us to coarse grain in the transverse direction, ignoring any effects due to the finite transverse coherence area of the radiation, within a single spatial mode.

In Sec. 2 we review the Gaussian state formalism, giving the definition of these states and establishing the notation we will use throughout this paper. Sec. 3 introduces the quantum Fisher information and gives an outline of the general principles of quantum metrology. In Sec. 4 we determine the density matrix for a blackbody state observed in the far-field and determine the optimal measurements for the parameters that govern

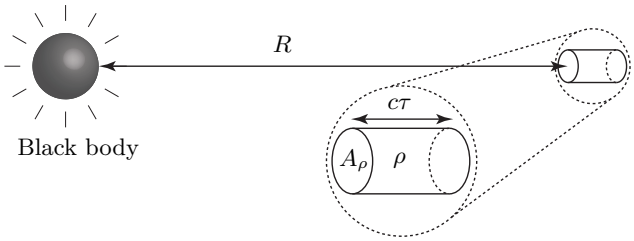


Figure 1: The definition of our state ρ. We assume that the transverse area over which the state is measured, A_ρ, is much smaller than the coherence area of the radiation. We also assume that the time over which the state is measured, τ, defines the spectral width $\Delta\nu = \tau^{-1}$.

these states. Unsurprisingly these turn out to be photon counting. In Sec. 5 we consider the effect of adding a secondary spatial mode, transversely separated from the first. We find it important to consider the coherence between the spatial modes, and in fact this becomes the essential parameter to estimate if we wish to glean spatial information about the source. We consider how to optimally obtain this spatial information and discover that the optimal observables do not commute. Therefore we are forced to find a measurement that can be considered optimal whilst being restricted by the inherent uncertainty of incompatible observables. A precisely optimal separable POVM is found that depends upon the values of the parameters. We also find a POVM which is independent of the parameter values and yet is close to optimal, which is of great practical importance.

2 Gaussian State Formalism

Gaussian states are defined to be states with a Gaussian characteristic function, which for an n-mode state defined by the bosonic mode operators $\boldsymbol{a} = (\hat{a}_1, \hat{a}_1^\dagger, \ldots, \hat{a}_n, \hat{a}_n^\dagger)^{\mathrm{T}}$, has the form [7]

$$\chi(\boldsymbol{\xi}) = \mathrm{Tr}[\rho \mathrm{e}^{-\boldsymbol{a}^{\mathrm{T}}\Omega\boldsymbol{\xi}}] = \exp\left(\frac{1}{2}(\Omega\boldsymbol{\xi})^{\mathrm{T}}\Sigma\Omega\boldsymbol{\xi} + \boldsymbol{\xi}\Omega\boldsymbol{\mu}\right), \tag{2}$$

where $\boldsymbol{\xi} = (\xi_1, \xi_1^*, \ldots, \xi_n, \xi_n^*)^{\mathrm{T}}$, $\Omega = \oplus_{i=1}^n i\sigma_y$, σ_y is the Pauli y matrix, and $\boldsymbol{\mu}$ and Σ are the first and second moments defined by

$$\mu^\alpha = \langle a^\alpha \rangle, \tag{3}$$

$$\Sigma^{\alpha\beta} = \frac{1}{2}\left(\langle \lambda^\alpha \lambda^\beta \rangle + \langle \lambda^\beta \lambda^\alpha \rangle\right), \tag{4}$$

and $\lambda^\alpha = a^\alpha - \mu^\alpha$. Often the quantities we wish to calculate can be given in closed form as a function of these moments. Due to the central limit theorem, Gaussian states are frequently encountered in systems where there are a large number of randomly fluctuating influences [14, 15].

3 Quantum Fisher Information

In classical parameter estimation, the Cramér-Rao bound states that the covariance matrix of unbiased estimates of the parameters, $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_d)^{\mathrm{T}}$, is bounded by the inverse of the Fisher information matrix [17]. That is,

$$\boldsymbol{\Sigma}_{\boldsymbol{\theta}} = \langle (\hat{\boldsymbol{\theta}} - \langle \hat{\boldsymbol{\theta}} \rangle)(\hat{\boldsymbol{\theta}} - \langle \hat{\boldsymbol{\theta}} \rangle)^{\mathrm{T}} \rangle \geq \mathbf{I}_C^{-1}, \tag{5}$$

where the hat over the $\boldsymbol{\theta}$ signifies that we are concerned with estimates of the parameters, and not the parameters themselves (and should not be confused with a quantum mechanical operator), and $\mathbf{I}_C$ is the classical Fisher information matrix:

$$\mathbf{I}_C = \sum_{\mathrm{x}} \frac{(\nabla_{\boldsymbol{\theta}} p(\mathrm{x}|\boldsymbol{\theta}))\,(\nabla_{\boldsymbol{\theta}} p(\mathrm{x}|\boldsymbol{\theta}))^{\mathrm{T}}}{p(\mathrm{x}|\boldsymbol{\theta})}, \tag{6}$$

where $\nabla_{\boldsymbol{\theta}} = (\frac{\partial}{\partial\theta_1}, \ldots, \frac{\partial}{\partial\theta_d})^{\mathrm{T}}$. The sum in Eq. (6) is performed over all possible outcomes, x, of the conditional probability distribution $p(\mathrm{x}|\boldsymbol{\theta})$.

The matrix inequality of Eq. (5) should be understood as implying that the matrix $\boldsymbol{\Sigma}_{\boldsymbol{\theta}} - \mathbf{I}_C^{-1}$ is positive semi-definite. It follows that the inequality

$$\mathrm{Tr}[G\,\boldsymbol{\Sigma}_{\boldsymbol{\theta}}] \geq \mathrm{Tr}[G\,\mathbf{I}_C^{-1}], \tag{7}$$

also holds for any positive definite weight matrix G. The weight matrix G allows us to assign a relative importance to different parameters.

The bound in Eq. (5) expresses a limitation of classical data processing. It states that estimates of a set of parameters, $\boldsymbol{\theta}$, are constrained in their variance by a quantity that is entirely determined by the probability distribution from which we sample data, $\{\mathrm{x}\}$. In quantum metrology, we consider how the covariance matrix of our estimates is constrained if the state, $\rho(\boldsymbol{\theta})$, is measured by some self-adjoint operator, $\hat{X} = \sum_{\mathrm{x}} \mathrm{x}\,\hat{\Pi}_{\mathrm{x}}$, giving rise to the probability distribution $p(\mathrm{x}|\boldsymbol{\theta}) = \mathrm{Tr}[\rho(\boldsymbol{\theta})\hat{\Pi}_{\mathrm{x}}]$. We assume that we are capable of measuring any self-adjoint operator, which allows us to reduce the variance by changing $p(\mathrm{x}|\boldsymbol{\theta})$. Finding the observable that minimises the variance is therefore the objective of quantum metrology.

In analogy with classical metrology, we can define a quantum Fisher information matrix, $\mathbf{I}_Q$, which leads to a quantum version of the Cramér-Rao bound. For a single parameter, the quantum Fisher information is

given by the expectation of the square of $\mathcal{L}$, the so-called symmetric logarithmic derivative (SLD):

$$\mathbf{I}_Q = \mathrm{Tr}[\rho\mathcal{L}^2]\,. \tag{8}$$

The SLD is implicitly defined by

$$\partial_\theta\rho = \frac{\rho\mathcal{L} + \mathcal{L}\rho}{2}\,. \tag{9}$$

In the multi-parameter case, this generalises to a quantum Fisher information matrix with elements [28]

$$\mathbf{I}_Q = \mathrm{Re}\left(\mathrm{Tr}[\rho\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{L}}^{\mathrm{T}}]\right)\,, \tag{10}$$

where we have defined $\boldsymbol{\mathcal{L}} = (\mathcal{L}_1, \ldots, \mathcal{L}_d)^{\mathrm{T}}$, which is the vector of SLD operators, and $\mathcal{L}_i$ is defined in analogy to the single parameter case $\frac{1}{2}(\rho\mathcal{L}_i + \mathcal{L}_i\rho) = \partial_{\theta_i}\rho$. For brevity, we write $\partial_i = \partial_{\theta_i}$ throughout the remainder of the paper.

In [7] Gao and Lee derived a closed form for the QFI of a Gaussian state in terms of the first and second moments, $\boldsymbol{\mu}$ and Σ. For blackbody states $\mu = 0$ and the expression found in [7] reduces to

$$\mathbf{I}_Q = \frac{1}{2}\mathfrak{M}^{-1}_{\alpha\beta,\gamma\kappa}\left(\nabla_{\boldsymbol{\theta}}\Sigma^{\gamma\kappa}\right)\left(\nabla_{\boldsymbol{\theta}}\Sigma^{\alpha\beta}\right)^{\mathrm{T}}\,, \tag{11}$$

where the matrix $\mathfrak{M} = \Sigma\otimes\Sigma + \frac{1}{4}\Omega\otimes\Omega$, and the summation convention is used for greek indices. Gao and Lee also provide an expression for the SLD, which is given in terms of $\mathfrak{M}$:

$$\mathcal{L}_i = \frac{1}{2}\mathfrak{M}^{-1}_{\gamma\kappa,\alpha\beta}\left(\partial_i\Sigma^{\alpha\beta}\right)\left(a^\gamma a^\kappa - \Sigma^{\gamma\kappa}\right)\,, \tag{12}$$

which we will make use of later to calculate the SLDs for blackbody states.

Eq. (10) gives the so-called SLD form of the quantum Fisher information. A quantum version of Eq. (5) can be determined, which, for a single parameter, is known to be attainable, at least in the asymptotic sense [3], by which we mean that it is possible to find an estimator where the variance of estimates tends towards the inverse of the QFI as the number of independent measurements tends to infinity.

To attain the bound requires finding a measurement for which the classical Fisher information is equal to the quantum Fisher information. For a single parameter, the optimal measurement is related to the SLD. However, in general the SLD will depend on the exact value of the parameter, which is assumed unknown before any measurements are made. The conventional method to circumvent this issue is to perform a suboptimal measurement first to obtain an initial estimate of the parameter, then use this estimate to measure an approximately optimal operator, preferably adaptively changing the measurement as more information is obtained about the parameters true value [2, 3, 9, 10].

For multiple parameters, the bound is not always attainable because the optimal observables may not commute and therefore may not be simultaneously measurable [12]. This introduces a new problem, independent to the problem of the optimal measurements depending on the true values of the parameters. Even if the SLDs do not commute, the quantum Cramér-Rao bound is asymptotically attainable if the following condition is satisfied [30]:

$$\mathrm{Tr}[\rho[\mathcal{L}_i, \mathcal{L}_j]] = 0. \tag{13}$$

However, to achieve the quantum Cramér-Rao bound in this context assumes the use of a collective measurement on multiple independent copies of the state. Although this is interesting, the implementation of such a measurement may well prove difficult. In Sec. 5 we encounter a problem with non-commuting SLDs and attempt to find the optimal separable measurement for our system.

4 Temperature Estimates of Far-Field Blackbody Sources

In this section we show how the quantum Fisher information can be calculated for temperature estimates of far-field blackbody radiation. We also find the optimal estimators for both temperature and solid angle and show that the quantum Cramér-Rao bound is attainable since the SLDs commute.

Blackbody radiation exhibits a Gaussian characteristic function and can therefore be analysed with the preceding formalism by Gao and Lee. In addition, blackbody states also have zero first moment, i.e., $\mu^\alpha = 0$ for all α. We assume that we observe the radiation from a blackbody state in the far-field, paraxial regime. Considering only a single spectral mode centred on the frequency ν, we find that the state has covariance matrix

$$\Sigma = \left(\langle n_\nu\rangle + \frac{1}{2}\right)\sigma_x\,, \tag{14}$$

where σ_x is the Pauli x operator. The covariance matrix depends only on $\langle n_\nu\rangle$ and we can therefore, estimate at most a single parameter.

From Planck's law, we can calculate the average number of photons for the arrangement given in Fig. 1,

$$\langle n_\nu\rangle = \frac{\nu^2 A_S A_\rho}{2\pi c^2 R^2}\frac{1}{\mathrm{e}^{\beta h\nu} - 1} = \nu^2\kappa\langle n_{\mathrm{th}}\rangle\,, \tag{15}$$

where A_S is the radiating area of the source, $\langle n_{\mathrm{th}}\rangle = [\exp(\beta h\nu) - 1]^{-1}$ is the average number of photons per

thermal mode, and κ contains all of the parameters regarding the arrangement of the source and its relation to the state ρ. We can also interpret $\nu^2\kappa$ as the number of modes of frequency ν that are captured by the state ρ. As mentioned above, we assume throughout that $\nu^2\kappa \ll 1$.

Given that we do not know the value of κ beforehand, the only way we can ensure that this condition is satisfied is by using appropriate values for the parameters that we can control, A_ρ and ν, the frequency we choose to observe. For example, using the area of a single CCD pixel as A_ρ, $\sim 10^{-11}\mathrm{m}^2$, we find that this condition is equivalent to $\cot\vartheta \gg \alpha\nu$, where ϑ is the angular size of the source and $\alpha = 10^{-14}\mathrm{s}^{-1}$. This condition should easily be satisfied in the far-field regime for frequencies up to X-rays.

Identifying $\theta_1 = T$ and $\theta_2 = \kappa$ we find that the QFI matrix for the spectral mode centred on ν is given by

$$\begin{aligned}\mathbf{I}_Q^{(\nu)} &= \frac{1}{\langle n_\nu\rangle + \langle n_\nu\rangle^2}\begin{pmatrix}(\partial_1\langle n_\nu\rangle)^2 & \partial_1\langle n_\nu\rangle\partial_2\langle n_\nu\rangle \\ \partial_2\langle n_\nu\rangle\partial_1\langle n_\nu\rangle & (\partial_2\langle n_\nu\rangle)^2\end{pmatrix} \\ &= \frac{1}{\langle n_\nu\rangle + \langle n_\nu\rangle^2}\nabla_{\boldsymbol{\theta}}\langle n_\nu\rangle(\nabla_{\boldsymbol{\theta}}\langle n_\nu\rangle)^{\mathrm{T}}, \end{aligned} \tag{16}$$

which is a rank 1 projector and therefore rank deficient. It is thus a singular matrix that does not have a well-defined inverse. Consequently, it is not possible to estimate both parameters with finite precision. We conclude that conforming to our expectation, we cannot find optimal unbiased estimators for both parameters from a single mode. However, if for some reason we do know one of the parameters precisely, we can make an estimate of the other. The SLD in that case is given by

$$\mathcal{L}_i = \frac{\partial_i\langle n_\nu\rangle}{\langle n_\nu\rangle + \langle n_\nu\rangle^2}\hat{a}^\dagger\hat{a} - \frac{\langle n_\nu\rangle\partial_i\langle n_\nu\rangle}{\langle n_\nu\rangle + \langle n_\nu\rangle^2}\mathbb{1}\,, \tag{17}$$

and the quantum Cramér-Rao bound for the single parameter θ_i is $\langle(\Delta\theta_i)^2\rangle \geq [\langle n_\nu\rangle + \langle n_\nu\rangle^2]/(\partial_i\langle n_\nu\rangle)^2$. This result was obtained in Ref. [11] and shows that the optimal measurement for temperature estimation is photon counting measurements. This follows from the form of the SLD in Eq. (17) and the observation that the optimal estimator for θ_i is given by [28]

$$X = \frac{\mathcal{L}_i}{\mathrm{I}_Q}\,. \tag{18}$$

Measurements of X are therefore achieved by measuring in the basis of the SLD $\mathcal{L}_i$. Since this is the number basis we can perform photon number counting and postprocess the outcomes to find the value of θ_i.

When we are ignorant of both parameters, which will generally be the case, we are required to make use of at least two spectral modes. The covariance matrix then becomes

$$\Sigma = \bigoplus_{i=1}^{M}\left(\langle n_{\nu_i}\rangle + \frac{1}{2}\right)\sigma_x\,, \tag{19}$$

where M is the total number of spectral modes we consider. Since each spectral mode is independent, the QFI matrix is given by

$$\mathbf{I}_Q = \sum_{i=1}^{M}\mathbf{I}_Q^{(\nu_i)}\,, \tag{20}$$

where $\mathbf{I}_Q^{(\nu_i)}$ are given by matrices of the form Eq. (16). Despite the matrices $\mathbf{I}_Q^{(\nu_i)}$ being singular, the matrix $\mathbf{I}_Q$ is non-singular as long as $M \geq 2$. The SLD for θ_i is given by

$$\mathcal{L}_i = \sum_{j=1}^{M}\frac{(\partial_i\langle n_{\nu_j}\rangle)(\langle n_{\nu_j}\rangle - \hat{n}_{\nu_j})}{\langle n_{\nu_j}\rangle + \langle n_{\nu_j}\rangle^2}\,. \tag{21}$$

From Eq. (21) we see that $[\mathcal{L}_1, \mathcal{L}_2] = 0$ and therefore it is possible to measure in the eigenbasis of both SLDs simultaneously. Therefore, photon counting in each of the M independent spectral modes provides an optimal estimate of both parameters simultaneously. In practice this is typically how stellar temperatures are measured, although often more sophisticated models are used to allow for deviations away from exact black body behaviour.

We stress that photon counting must be performed on at least two spectral modes in order to obtain estimates for both parameters. The variance for the parameter θ_i is

$$\langle(\Delta\theta_i)^2\rangle \geq \frac{\sum_{l=1}^{M}\left[\mathbf{C}^{(\nu_l)}\right]_{ii}}{\sum_{l,k=1}^{M}\left(\left[\mathbf{I}_Q^{(\nu_l)}\right]_{11}\left[\mathbf{I}_Q^{(\nu_k)}\right]_{22} - \left[\mathbf{I}_Q^{(\nu_l)}\right]_{12}\left[\mathbf{I}_Q^{(\nu_k)}\right]_{21}\right)}\,, \tag{22}$$

where $\mathbf{C}^{(\nu_l)}$ is the cofactor matrix of $\mathbf{I}_Q^{(\nu_l)}$, and $\mathbf{I}_Q^{(\nu_l)}$ is

$$\mathbf{I}_Q^{(\nu_l)} = \frac{(\nabla_{\boldsymbol{\theta}}\langle n_{\nu_l}\rangle)(\nabla_{\boldsymbol{\theta}}\langle n_{\nu_l}\rangle)^{\mathrm{T}}}{\langle n_{\nu_l}\rangle + \langle n_{\nu_l}\rangle^2}\,. \tag{23}$$

Assuming that we are interested in estimating T and treating κ as a nuisance parameter, we plot in Fig. 2 the variance in estimates of T for two spectral modes as a function of the frequency of each mode. In Fig. 2 we see that there is a well defined optimum for the frequencies ν_1 and ν_2, the exact value of which we would expect to depend upon the precise values of the parameters to be estimated.

Numerically we find that the optimal values of ν_1 and ν_2 are independent of κ and linearly dependent

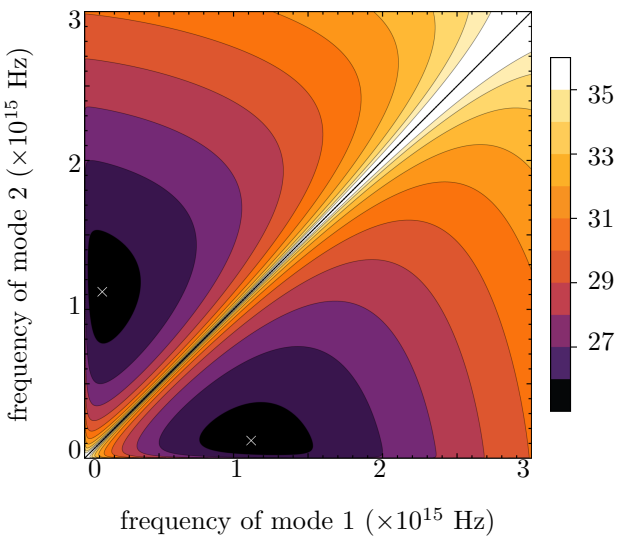


Figure 2: The natural logarithm of the variance of T (in units of Kelvin) as a function of the frequencies ν_1 and ν_2. The exact location of the optimal frequencies is dependent on the value of T. In this plot we have used $T = 10000$ K and $\kappa = 10^{-32}\,s^2$, which ensures that $N \ll 1$ over the range of frequencies plotted. It is not surprising that the values on the bar are so large when we consider that the variance quoted is for a single round of measurement and we are using only two spectral modes. Also, at 10000 K the average number of photons per measurement is always less than 0.0003. The white crosses show the location of the minimum.

on T, approximately $\nu_1^{(\min)} = (1.188 \times 10^{10}\,\mathrm{Hz\,K^{-1}})T$, $\nu_2^{(\min)} = (1.118 \times 10^{11}\,\mathrm{Hz\,K^{-1}})T$. This is reminiscent of Wien's displacement law, which gives the peak frequency of blackbody radiation as $\nu^{(\max)} = (5.88 \times 10^{10}\,\mathrm{Hz\,K^{-1}})T$. In the absence of an initial estimate of T, the best we can do is to perform photon counting for two frequencies and adjust these correspondingly as information is obtained about the temperature. If we happen to know of an a priori distribution for the temperature (for example, when estimating the temperature of stars an a priori distribution can be obtained using observational data about the relative frequency of stellar temperatures), we can choose the initial frequencies based upon minimisation of the averaged variance, that is

$$\min_{\nu_1,\nu_2} \int \mathrm{d}T\, p(T)\, [\mathbf{I}_Q^{-1}]_{11}\,, \tag{24}$$

where $p(T)$ is our prior distribution over T.

Generally we may want to estimate both parameters. A higher precision can be achieved in our estimates by increasing the number of spectral modes observed. Assuming that we are only capable of photon counting in M spectral modes simultaneously, an optimal measurement can be found given these limited resources.

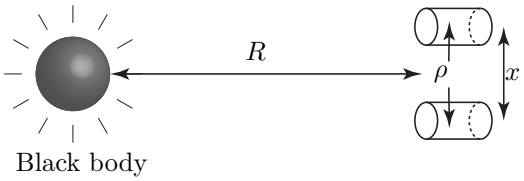


Figure 3: The addition of a secondary spatial mode, the modes are transversely separated by a distance $x = |\mathbf{x}_1 - \mathbf{x}_2|$. The state ρ is now a composite system consisting of two spatially separated modes. We again assume that the transverse area of each spatial mode is less than the coherence area of the radiation, which means that each mode captures a fraction of a spectral mode. The van Cittert-Zernike theorem allows us to calculate the far-field complex degree of coherence between the two modes, $\gamma(\mathbf{x}_1, \mathbf{x}_2) = \gamma_{12}$.

Defining $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_M)^{\mathrm{T}}$, an optimal measurement is given by

$$\min_{\boldsymbol{\nu}} \mathrm{Tr}[G\, \mathbf{I}_Q^{-1}(\boldsymbol{\nu})]\,. \tag{25}$$

The solution of this minimisation will depend on the actual values of the parameters T, and κ, and the chosen weight matrix G. Therefore, without prior knowledge about the parameters, an optimal measurement for a given weight matrix is one that minimises

$$\min_{\boldsymbol{\nu}} \int \mathrm{d}\kappa\, \mathrm{d}T\, p(\kappa, T) \mathrm{Tr}[G\, \mathbf{I}_Q^{-1}(\boldsymbol{\nu})]\,, \tag{26}$$

where $p(\kappa, T)$ is a prior distribution over the parameters, which will depend strongly upon the exact application.

We are now naturally led to the question: what measurement can we perform to determine the spatial configuration of the source? Most single parameter problems can be treated by estimating κ, which can then be used to determine this single parameter. For example if the source is simply planar circular, then the estimation of the parameter $\kappa \propto A_S/R^2 = 2\pi r^2/R^2 = 2\pi\vartheta^2$ is equivalent to estimating the angular radius ϑ. In general however, since the QFI depends only upon the average number of photons in each frequency mode, which in turn depends only on the source temperature and κ, we cannot estimate general spatial properties of the source, no matter how many spectral modes we measure.

Since the spatial properties are all contained within the parameter κ, any attempt to estimate more than a single spatial parameter results in a singular QFI matrix. In order to allow the estimation of arbitrary spatial properties, we need to consider the effect of adding additional spatial modes, which are spatially separated from the first mode, see Fig. 3.

5 Estimation of Spatial Parameters

In the previous section we showed how to optimally estimate the temperature of black body sources. In this section we show that the addition of spatial modes allows for the determination of the spatial properties of the source. We compare different POVMs, finding the optimal separable POVM, as well as a POVM which is close to optimal for a large range of parameters. This near optimal POVM has the advantage that its implementation does not require knowledge of the parameters, and therefore it can be implemented at the outset and without the need to adaptively change the POVM.

In order to compare various schemes we use as a quantifier the cost function

$$V_M(\mathbf{I}_Q) = \mathrm{Tr}[\mathbf{I}_Q \mathbf{I}_C^{-1}(M)]\,, \tag{27}$$

where $\mathbf{I}_Q$ is the quantum Fisher information and $\mathbf{I}_C^{-1}$ is the inverse of the classical Fisher information matrix for a given POVM, M. In general any positive matrix can be used as the weight matrix in the cost function. The reason for this particular choice of weight matrix is that it maximises the average fidelity between the estimated state and the actual state [2] and is therefore the natural choice for state estimation. We would expect that a measurement which is simultaneously optimal for all parameters should achieve $V(\mathbf{I}_Q) = d$, where d is the number of parameters. However, when the SLDs do not commute and we restrict ourselves to separable measurements, we find $V(\mathbf{I}_Q) > d$.

5.1 Two Spatial Mode State

The state ρ now describes the composite system of two spatially separated volumes, see Fig. 3. The addition of a secondary spatial mode changes the covariance matrix in the following way:

$$\Sigma = \bigoplus_{i=1}^{M} \begin{pmatrix} 0 & c_1^{(i)} & 0 & b_i \\ c_1^{(i)} & 0 & b_i^* & 0 \\ 0 & b_i^* & 0 & c_2^{(i)} \\ b_i & 0 & c_2^{(i)} & 0 \end{pmatrix}, \tag{28}$$

where $b_i = \langle \hat{a}_2^{(i)\dagger} \hat{a}_1^{(i)} \rangle$, $c_j^{(i)} = \langle \hat{a}_j^{(i)\dagger} \hat{a}_j^{(i)} \rangle + 1/2$, and $\hat{a}_j^{(i)}$ is the annihilation operator in the spatial mode j and the spectral mode i. To determine the elements b_i and b_i^* we first note that the complex degree of coherence between the two spatial modes is defined as

$$\gamma_{12}^{(i)} = \frac{\langle \hat{a}_2^{(i)\dagger} \hat{a}_1^{(i)} \rangle}{\left[\langle n_1^{(i)} \rangle \langle n_2^{(i)} \rangle \right]^{\frac{1}{2}}}\,, \tag{29}$$

which allows us to write $b_i = [\langle n_1^{(i)} \rangle \langle n_2^{(i)} \rangle]^{\frac{1}{2}} \gamma_{21}^{(i)}$ and $b_i^* = [\langle n_1^{(i)} \rangle \langle n_2^{(i)} \rangle]^{\frac{1}{2}} \gamma_{12}^{(i)}$. We now make use of the van Cittert-Zernike theorem, which states that the far-field complex degree of coherence of a spatially incoherent source is proportional to the Fourier transform of the intensity distribution [22]. Therefore the correlations between the two modes provide information about the Fourier transform of the source distribution.

In Section 4 we saw that the variables $\langle n_j^{(i)} \rangle$ can convey only information about the temperature of the source and κ. In contrast, the variables $\gamma_{12}^{(i)}$ convey detailed information about the distribution of the source. Since the Fourier transform is an injective mapping, knowledge about the Fourier transform over the entire far-field plane can be used to exactly reconstruct the source intensity distribution. Restricting the detection area of the state will result in incomplete information about the Fourier transform and therefore limits the resolution of a reconstructed intensity distribution. In this paper we restrict our attention to the optimal estimation of the state consisting of two spatial modes for simplicity.

We make an additional simplification and assume that the average photon number in both spatial modes is approximately equal, $\langle n_1^{(i)} \rangle = \langle n_2^{(i)} \rangle$. We expect this assumption to hold in the far-field of an isotropic emitter. For brevity we will also drop the superscript (i) since each frequency mode is independent and therefore can be optimised separately.

5.2 Optimal Estimators of Spatial Parameters

Under this assumption we calculate the SLDs for the parameters $\boldsymbol{\theta} = (\langle n \rangle, |\gamma|, \phi)^{\mathrm{T}}$, where $\gamma_{12} = |\gamma| \exp(i\phi)$, which are found to be of the form

$$\mathcal{L}_j = P_j \hat{n}_{\mathrm{tot}} + Q_j \hat{a}_1^\dagger \hat{a}_2 + Q_j^* \hat{a}_2^\dagger \hat{a}_1 + R_j \mathbb{1}, \tag{30}$$

where $\hat{n}_{\mathrm{tot}} = \hat{n}_1 + \hat{n}_2$ and explicit expressions for P_j, Q_j, and R_j are given in App. A. As we show in App. A, the commutators for these operators are $[\mathcal{L}_1, \mathcal{L}_2] = 0$, $[\mathcal{L}_1, \mathcal{L}_3] \neq 0$, $[\mathcal{L}_2, \mathcal{L}_3] \neq 0$ and therefore we cannot find a simultaneous eigenbasis for all three operators. This rules out the possibility of simultaneously measuring in the eigenbases of the SLDs to achieve a simultaneously optimal estimate of each parameter. Since they do however satisfy the condition Eq. (13) (see App. A), it is possible that a collective measurement exists which attains the QCRB asymptotically [8].

In this paper we will not consider collective measurements due to the immense technical obstacles. Instead we consider only the class of measurements which are separable. To find the optimal POVM, we first determine the operators $\boldsymbol{X} = \mathbf{I}_Q^{-1} \boldsymbol{\mathcal{L}}$, where we have defined $\boldsymbol{X} = (X_1, X_2, X_3)^{\mathrm{T}}$ and $\boldsymbol{\mathcal{L}} = (\mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3)^{\mathrm{T}}$. This is the multi-parameter extension of Eq. (18) and has the desirable property that $\mathrm{Var}(X_i) = [\mathbf{I}_Q^{-1}]_{ii}$.

Determining the commutators of these operators, we

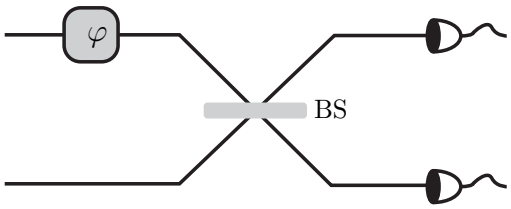


Figure 4: Measurement scheme with variable phase shift φ. Choosing different values for the phase shift allows optimal estimation of the parameters $\langle n\rangle$, $|\gamma|$, and ϕ.

find $[X_1, X_2] = [X_1, X_3] = 0$, $[X_2, X_3] \neq 0$ (see App. A) and therefore X_2, X_3 cannot be simultaneously measured. Next we calculate the quantum Fisher information from the covariance matrix Eq. (28) and determine the classical Fisher information for measurements of the operators X_2, X_3. We find,

$$\mathbf{I}_Q = \begin{pmatrix} [\mathrm{I}_Q]_{11} & [\mathrm{I}_Q]_{12} & 0 \\ [\mathrm{I}_Q]_{12} & [\mathrm{I}_Q]_{22} & 0 \\ 0 & 0 & [\mathrm{I}_Q]_{33} \end{pmatrix}, \tag{31}$$

$$\mathbf{I}_C(X_2) = \begin{pmatrix} [\mathrm{I}_Q]_{11} & [\mathrm{I}_Q]_{12} & 0 \\ [\mathrm{I}_Q]_{12} & [\mathrm{I}_Q]_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{32}$$

$$\mathbf{I}_C(X_3) = \begin{pmatrix} [\mathrm{I}_Q]_{11} & \delta_1 & 0 \\ \delta_1 & \delta_2 & 0 \\ 0 & 0 & [\mathrm{I}_Q]_{33} \end{pmatrix}, \tag{33}$$

where $\delta_1 \ll [\mathrm{I}_Q]_{12}$ and $\delta_2 \ll [\mathrm{I}_Q]_{22}$.

We notice that a slight asymmetry exists between the operators X_2 and X_3. Whilst measurements of X_2 do not provide information about ϕ (since $[\mathbf{I}_C(X_2)]_{33} = 0$), measurements of X_3 do provide a small amount of information about $|\gamma|$ ($\delta_2, \delta_3 \neq 0$). In App. B we argue that the optimal measurement scheme is given by the solution to the following minimization:

$$\min_p \mathrm{Tr}[\mathbf{I}_Q(p\,\mathbf{I}_C(X_2) + (1-p)\mathbf{I}_C(X_3))^{-1}]. \tag{34}$$

This scheme can be interpreted as probabilistically choosing to measure X_2 or X_3 with probabilities p and $1-p$ respectively. The minimisation in Eq. (34) ensures that we choose the optimal weighting with respect to X_2 and X_3. We refer to this scheme as the weighted measurement scheme.

Due to the simple quadratic forms of the SLDs, and therefore also the operators $\boldsymbol{X}$, we can determine the eigenmodes of X_2 and X_3. We find that both operators can be measured using the scheme shown in Fig. 4, where the value of φ is different for X_2 and X_3. To measure X_2 requires the choice $\varphi = \phi$ and X_3 can be achieved by setting $\varphi = \phi - \pi/2$. An optimal measurement therefore requires switching between the two distinct phase settings. As discussed above, the probability with which each measurement is made should also be optimised. The exact value of the probability p will depend on the exact values of the parameters. Since this scheme clearly requires knowledge of the parameters we are attempting to measure, it can only be implemented in an adaptive scenario. However, since we now know the optimal measurement, we can use this to compare other, simpler schemes, with the aim to find a scheme that is close to optimal but is independent of the parameters.

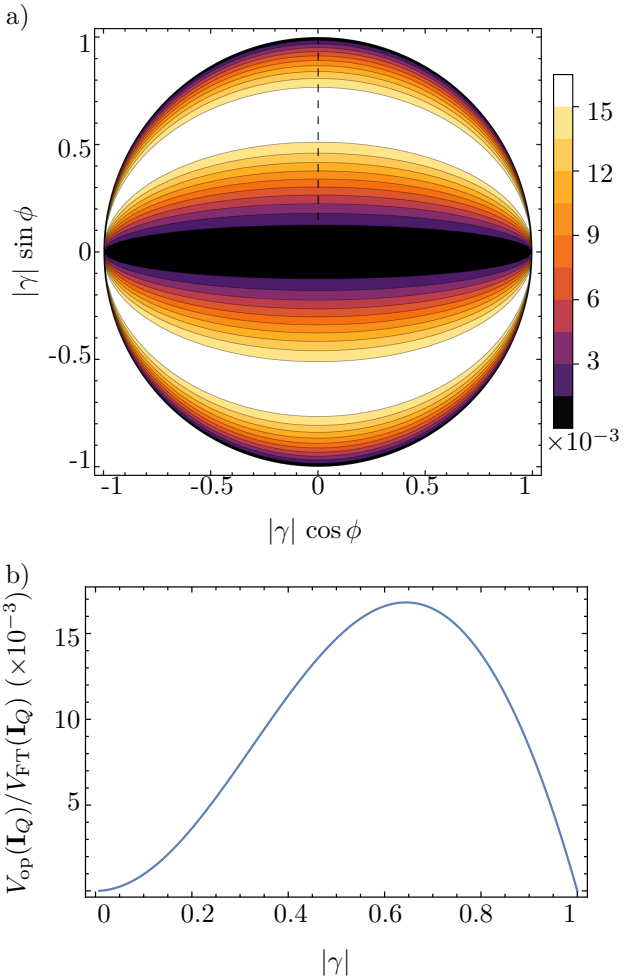


Figure 5: $V_{\mathrm{op}}(\mathbf{I}_Q)/V_{\mathrm{FT}}(\mathbf{I}_Q)$ for the measurement scheme involving only a Fourier transform and photon counting. a) We see that this scheme is suboptimal for the entire parameter range $0 \leq |\gamma| \leq 1$, $0 \leq \phi < 2\pi$ and in fact never achieves more than 1.7 per cent of the performance of the optimal estimator. b) Shows a cut through fig a) along the dashed line. In this plot we set $\langle n\rangle = 0.01$.

5.3 Fixed Phase Scheme

We first consider how well the parameters can be estimated using a Fourier transform and photon counting on the two input modes. This is of interest for optics because it closely approximates the action of lensing and intensity measurements, which is by far the most commonly used method in imaging. The photon count probability distributions for both the state ρ and the state after the action of the Fourier transform are derived in App. C.

We find that without the Fourier transform the photon count distribution is independent of the phase ϕ and therefore cannot be used to estimate ϕ. However, once the Fourier transform has been applied to the input modes all three parameters can be estimated. In Fig. 5 we plot the ratio of the cost function for the Fourier transform scheme, $V_{\rm FT}(\mathbf{I}_Q)$, to the optimal scheme for the entire range of the parameters $0 \leq |\gamma| \leq 1$, and $0 \leq \phi < 2\pi$, which follows from the definition of γ_{12} and the restriction $0 \leq |\gamma_{12}| \leq 1$. We see that this scheme performs far below the optimal scheme and is therefore not an efficient measurement for determining the spatial configuration of the source.

5.4 Random Phase Scheme

The second scheme we consider is a scheme where we randomly select a phase shift, uniformly over the range $\varphi \in [0, 2\pi)$, then pass the modes through a beamsplitter and measure in the number basis. We call this scheme the random phase (RP) scheme. In Fig. 6 we plot the ratio $V_{\rm op}(\mathbf{I}_Q)/V_{\rm RP}(\mathbf{I}_Q)$ to compare the random phase scheme to the optimal measurement scheme. We find that the random phase scheme is very close to optimal for most of the parameter range $0 \leq |\gamma| \leq 1$, $0 \leq \phi < 2\pi$. The advantage of the random phase scheme is that, in implementing this scheme, we do not require the values of the parameters. This therefore means that this POVM can be implemented before we have acquired any information about the parameters. Surprisingly, even without this information, the random phase scheme performs very efficiently.

As discussed above, the truly optimal POVM requires us to choose measurements of X_2 with probability p and X_3 with probability $(1-p)$. Although this achieves the optimal performance it requires knowledge of the parameter values and therefore can only be implemented in an adaptive way. We expect that an approximately optimal scheme can be achieved by first making measurements using the random phase scheme, and when the estimates of the parameters have reached a sufficient precision the optimal scheme can be used to further enhance the measurement precision.

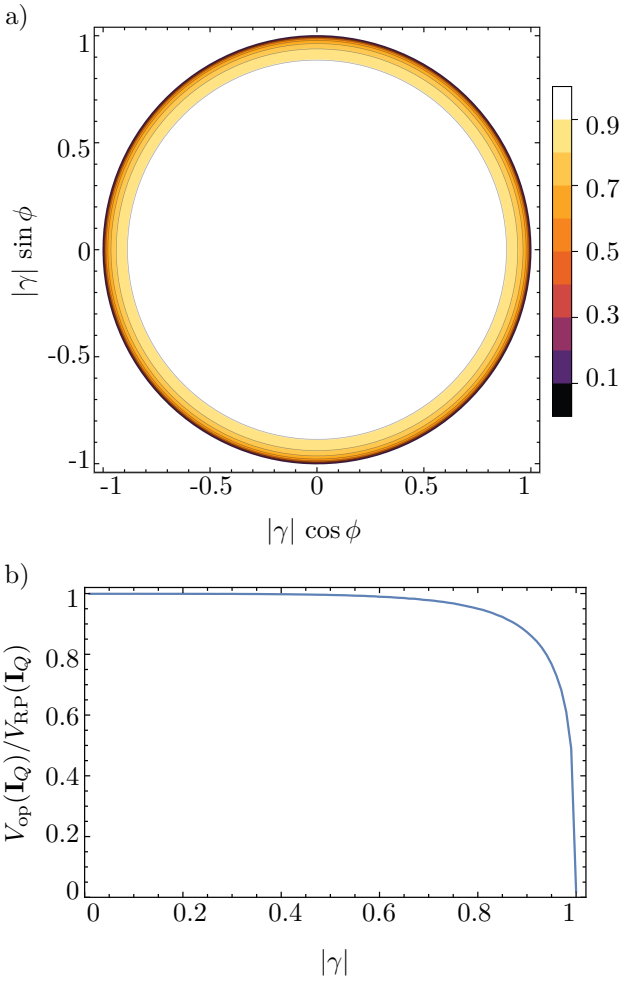


Figure 6: The ratio of $V_{\rm op}(\mathbf{I}_Q)$ to $V_{\rm RP}(\mathbf{I}_Q)$. The random phase scheme is seen to be close to optimal unless $|\gamma|$ is close to one. For this figure we randomly sampled 1000 phases φ and determined the inverse of the classical Fisher information for such a set of measurements. Since the exact performance is affected by the exact values of φ we averaged each point for over 400 trials to remove most of the statistical noise. The inset shows the performance over the whole range, which shows that the performance of the random phase scheme is independent of ϕ.

6 Discussion and Conclusions

In this paper we have derived a complete description of states originating in far-field blackbody sources and observed over a finite detection area in the far-field. This description is complete in the sense that it encompasses all parameters that determine the state, including nuisance parameters that must be estimated in order to estimate the temperature of the blackbody and its spatial configuration. We found that it is necessary to at-

tempt to estimate these nuisance parameters in order to estimate the temperature and to do so requires photon counting to be performed on at least two spectral modes.

In order to derive our results, it was necessary to make certain assumptions about the arrangement of the source and its relation to the state ρ. In the paper we assumed that the transverse area of the source is much less than one coherence area. We demonstrated that this gave a bound on the frequencies we can measure, $10^{-14}\nu \ll \cot\vartheta$, where ϑ is the angular size of the source. We also found a linear relationship between the optimal pair of frequencies and the temperature. Putting these results together allows us to find the following bound $\cot\vartheta \gg \epsilon T$, where $\epsilon \approx 10^{-3}\mathrm{K}^{-1}$ which shows that the optimal frequencies for very hot sources ($\sim 10^5 - 10^6\mathrm{K}$) should fall in the range of acceptable frequencies for far-field objects.

Interestingly, we found that the estimation of general spatial parameters requires that the states are observed over at least two spatially separated modes. The introduction of a second spatial mode introduces a complex parameter which determines the coherence between the spatial modes and can be identified as the Fourier transform of the intensity distribution of the source. We then identified the optimal separable measurement for the resulting three parameter estimation problem.

The optimal strategy depends upon the exact value of the parameters we wish to estimate, and therefore can only be implemented in an adaptive setting. However, we also identified another strategy, the random phase scheme, which performs close to optimal for a wide range of the parameters (see Fig. 6). The implementation of the random phase scheme is independent of the parameter values and therefore has the advantage of not requiring an adaptive arrangement. We can then suppose that an asymptotically optimal scheme can be achieved by starting with the random phase scheme, using this to find a reasonable estimate of the parameters, and then implementing the optimal scheme to further increase the precision of the estimates.

We also considered the performance of the measurement scheme consisting of simply a Fourier transform and photon counting. This is actually a special case of the measurement scheme in Fig. 4, with φ fixed at a constant value (namely zero). Our motivation is that this scheme most closely resembles a typical imaging scheme, with the role of the Fourier transform being provided by a lens and photon counting provided by a CCD or other photosensitive surface. We find that this scheme performs poorly over the entire parameter range.

Our results expose simple imaging schemes as far from optimal, even in the simplistic two mode settings. We have shown how spatial information is conveyed to the far-field, and how we can optimally extract this information for the simplest case of two modes. We expect that this work will enable further research into quantum optimal imaging and helps to answer the two open questions: 1) What are the optimal measurements for measuring spatial features of radiating sources? 2) How well do standard imaging techniques compare to optimal schemes?

Acknowledgements

The authors would like to thank Mark Howard, Madalin Guta, and Mankei Tsang for helpful discussions relating to the issues addressed in this paper. MEP and PK acknowledge EPSRC for funding via the Quantum Communications Hub. ETC is supported by the EPSRC (EP/M024261/1).

A Symmetric Logarithmic Derivatives of $\langle n\rangle$, $|\gamma|$, and ϕ

Here we give the explicit forms of the variables P_j, Q_j, and R_j, defined in Eq. (30), which determine the SLDs for the three parameters $\boldsymbol{\theta} = (\langle n\rangle, |\gamma|, \phi)^{\mathrm{T}}$. They are:

$$P_{\langle n\rangle} = \frac{\langle n\rangle + 1}{\langle n\rangle\left[\langle n\rangle^2|\gamma|^2 - (\langle n\rangle - 1)^2\right]} \tag{35}$$

$$Q_{\langle n\rangle} = \frac{|\gamma|\mathrm{e}^{-i\phi}}{\langle n\rangle\left[\langle n\rangle^2|\gamma|^2 - (\langle n\rangle - 1)^2\right]} \tag{36}$$

$$R_{\langle n\rangle} = \frac{2\langle n\rangle(|\gamma|^2\langle n\rangle - \langle n\rangle - 1)}{\langle n\rangle\left[\langle n\rangle^2|\gamma|^2 - (\langle n\rangle - 1)^2\right]} \tag{37}$$

$$P_{|\gamma|} = \frac{2\langle n\rangle + 1}{(|\gamma|^2 - 1)\left[\langle n\rangle^2|\gamma|^2 - (\langle n\rangle - 1)^2\right]} \tag{38}$$

$$Q_{|\gamma|} = \frac{\mathrm{e}^{-i\phi}(1 + \langle n\rangle + |\gamma|^2\langle n\rangle^2)}{(|\gamma|^2 - 1)\left[\langle n\rangle^2|\gamma|^2 - (\langle n\rangle - 1)^2\right]} \tag{39}$$

$$R_{|\gamma|} = \frac{2|\gamma|\left[\langle n\rangle^2(|\gamma|^2 - 1)\right]}{(A^2 - 1)\left[\langle n\rangle^2|\gamma|^2 - (\langle n\rangle - 1)^2\right]} \tag{40}$$

$$P_\phi = 0 \tag{41}$$

$$Q_\phi = i|\gamma|\mathrm{e}^{-i\phi} \tag{42}$$

$$R_\phi = 0\,. \tag{43}$$

Taking the commutator of each of the SLDs pairwise, we find $[\mathcal{L}_{\langle n\rangle}, \mathcal{L}_{|\gamma|}] = 0$, and $[\mathcal{L}_{\langle n\rangle}, \mathcal{L}_\phi], [\mathcal{L}_{|\gamma|}, \mathcal{L}_\phi] \propto \hat{n}_1 - \hat{n}_2$. As we show in the main paper, the observables that provide optimal information about each parameter independently are the set $\boldsymbol{X}$, given by

$$X_i = \left[\mathbf{I}_Q^{-1}\boldsymbol{\mathcal{L}}\right]_i\,. \tag{44}$$

From Eqs. (35) to (43) we can determine the commutators of the $\boldsymbol{X}$ operators. These are $[X_{\langle n\rangle}, X_{|\gamma|}] = [X_{\langle n\rangle}, X_\phi] = 0$, and $[X_{|\gamma|}, X_\phi] \propto \hat{n}_1 - \hat{n}_2$. Therefore we cannot measure $X_{|\gamma|}$ and X_ϕ simultaneously.

B Optimality of Weighted Measurement Scheme

Here we present evidence of the optimality of the weighted measurement scheme presented in Sec. 5. We appeal to the results of [9] where the authors show

$$\mathrm{Tr}[\mathbf{I}_Q^{-1}\mathbf{I}_C] \leq D - 1\,, \tag{45}$$

where D is the dimension of the Hilbert space. For a d parameter problem, it immediately follows that

$$\mathrm{Tr}[\mathbf{I}_Q\mathbf{I}_C^{-1}] \geq \frac{d^2}{D-1}\,. \tag{46}$$

In our setting, the dimension is infinite. However, since the state consists of thermal modes, which typically have very small average photon numbers [22], we can truncate the state to the 0, 1 photon basis with low truncation error. Using this $D = 3$ approximation (the basis of $|0,0\rangle$, $|0,1\rangle$, $|1,0\rangle$) suggests that 4.5 is a good lower bound on $\mathrm{Tr}[\mathbf{I}_Q\mathbf{I}_C^{-1}]$. However, there is no reason to believe it is attainable. In the main text, we reported that we can achieve just below 5. Our approach is to choose between the optimal measurements with some probability. Below we show that for qubits, such an approach is optimal and saturates the lower bound of Eq. (46). However, for qutrits and low photon number Gaussian states, our numerical investigation indicate that it is not possible to saturate Eq. (46), though we come very close.

The quantity $\mathrm{Tr}[\mathbf{I}_Q\mathbf{I}_C^{-1}]$ is invariant under reparameterisation and we can always work in parameterisation where $\mathbf{I}_Q$ is diagonal [9]. If we make the optimal measurement for parameter i, we obtain a classical Fisher information $\mathbf{I}_C^{(i)}$ such that $[\mathbf{I}_C^{(i)}]_{i,i} = [\mathbf{I}_Q]_{i,i}$. We may obtain information about other parameters and so

$$\mathbf{I}_C^{(i)} \geq \tilde{\mathbf{I}}_C^{(i)}, \tag{47}$$

where $\tilde{\mathbf{I}}_C^{(i)} = \mathrm{diag}(0, \ldots, [\mathbf{I}_Q]_{ii}, \ldots, 0)$. For a d parameter problem, we propose a scheme were each optimal observable X_i is measured on a fraction $1/d$ of the samples. Then we obtain

$$\mathbf{I}_C = \sum_j \frac{1}{d}\mathbf{I}_C^{(j)} \geq \sum \frac{1}{d}\tilde{\mathbf{I}}_C^{(i)} = \frac{1}{d}\mathbf{I}_Q. \tag{48}$$

Therefore,

$$\mathrm{Tr}[\mathbf{I}_Q\mathbf{I}_C^{-1}] \leq d\,\mathrm{Tr}[\mathbb{1}] = d^2. \tag{49}$$

For a qubit, Eq. (46) gives a lower bound of d^2 and so we see such a scheme is optimal. For a general qutrit problem, we have

$$\frac{d^2}{2} \leq \mathrm{Tr}[\mathbf{I}_Q\mathbf{I}_C^{-1}] \leq d^2 \tag{50}$$

with the actual optimal depending on the exact nature of the problem. For a three parameter problem, we have a lower bound of 4.5, and the above scheme achieves 9 or better. In the problem of interest to us, some of our optimal measurements commute (see App. A) and so we expect to do much better than 9.

Inspired by the above, to find the exact optimum for our problem we allow for minimisation with respect to the relative probability that we measure X_A and X_ϕ. Denoting the probability that we measure X_A as p and using the approximation $\delta_1 = \delta_2 = 0$, we find

$$\min_p\ \mathrm{Tr}[\mathbf{I}_Q(p\mathbf{I}_C^{(2)} + (1-p)\mathbf{I}_C^{(3)})^{-1}] = \min_p\ 1 + \frac{1}{p} + \frac{1}{1-p} = 5\,, \tag{51}$$

where the value of p that minimises this expression is $p = 1/2$. When $\delta_1, \delta_2 \neq 0$, the value of p that minimises is slightly less than one half and the cost is just less than 5.

To further support our claim that this scheme is indeed optimal, we perform a numerical optimisation over a subset of POVMs. In order to reduce the number of parameters to optimise, we search over the truncated Hilbert space of $|0,0\rangle, |0,1\rangle, |1,0\rangle$. We search over the set of POVMs with six elements corresponding to two sets of three orthogonal components. To perform the optimisation, we apply two unique three dimensional unitaries to the POVM with elements $\{|0,0\rangle\langle 0,0|, |0,1\rangle\langle 0,1|, |1,0\rangle\langle 1,0|\}$, giving two POVMs with three elements each, M_1 and M_2. A six element POVM is constructed by taking $M = pM_1 + (1-p)M_2$, for which we calculate the classical Fisher information. We then search for the minimum with respect to the cost function $\mathrm{Tr}[\mathbf{I}_Q\mathbf{I}_C^{-1}]$, where $\mathbf{I}_C$ is the classical Fisher information associated with measurements of this POVM. Performing the optimisation across the range $0 \leq |\gamma| \leq 1$, the results are in good agreement with the values obtained for the weighted measurement scheme.

C Photon Count Probability Distributions

To determine the photon count probability distributions we first notice that the state with covariance matrix

Eq. (28) can be obtained by the action of the unitary

$$U(\phi) = \frac{1}{\sqrt{2}} \begin{pmatrix} -\mathrm{e}^{-i\phi} & \mathrm{e}^{-i\phi} \\ 1 & 1 \end{pmatrix}, \tag{52}$$

acting on the mode space of the two mode, thermal state with covariance matrix

$$\begin{pmatrix} 0 & x_1 + \frac{1}{2} \\ x_1 + \frac{1}{2} & 0 \end{pmatrix} \oplus \begin{pmatrix} 0 & x_2 + \frac{1}{2} \\ x_2 + \frac{1}{2} & 0 \end{pmatrix}, \tag{53}$$

where $x_1 = \langle n \rangle (1 - |\gamma|)$, $x_2 = \langle n \rangle (1 + |\gamma|)$. Since the modes are independent, the photon count probability distribution is simply given by

$$p_{\mathrm{in}}(n_1, n_2) = \frac{x_1^{n_1}}{(1+x_1)^{n_1+1}} \frac{x_2^{n_2}}{(1+x_2)^{n_2+1}}, \tag{54}$$

and the state can be expressed in the number basis as

$$\rho_{\mathrm{in}} = \sum_{n_1, n_2 = 0}^{\infty} p_{\mathrm{in}}(n_1, n_2) |n_1, n_2\rangle \langle n_1, n_2|. \tag{55}$$

Making use of the relation $|n_i\rangle = n_i^{-\frac{1}{2}} (\hat{a}_i^\dagger)^{n_i} |0\rangle$, and taking the transformation $[\hat{a}_{\mathrm{in}}]_i = U_{ij}^\dagger(\phi) [\hat{a}_{\mathrm{out}}]_j$, we find the photon counting probability distribution is given by

$$p_{\mathrm{out}}(m_1, m_2) = \sum_{n_1=0}^{m_1+m_2} \frac{p_{\mathrm{in}}(n_1, m_1 + m_2 - n_1)}{n_1!(m_1 + m_2 - n_1)!} \frac{m_1! m_2!}{2^{m_1+m_2}} \times \left| \sum_{j=0}^{m_1} (-1)^j \binom{n_1}{j} \binom{m_1 + m_2 - n_1}{m_1 - j} \right|^2, \tag{56}$$

where we have made use of the binomial theorem. We notice that this distribution is independent of ϕ and therefore cannot be used to obtain estimates of ϕ.

To obtain the distribution after the application of the Fourier transform, we note that the discrete 2D-Fourier transform is given by $U(0)$, we therefore use the transformation $[\hat{a}_{\mathrm{in}}]_i = [(U(0)U(\phi))^\dagger]_{ij} [\hat{a}_{\mathrm{out}}]_j$ to obtain the distribution

$$p_{\mathrm{out}}(m_1, m_2) = \sum_{n_1=0}^{m_1+m_2} \frac{p_{\mathrm{in}}(n_1, m_1 + m_2 - n_1)}{n_1!(m_1 + m_2 - n_1)!} \frac{m_1! m_2!}{4^{m_1+m_2}} \left| \sum_{j=0}^{m_1} (-1)^j \binom{n_1}{j} \binom{m_1 + m_2 - n_1}{m_1 - j} (1 - \mathrm{e}^{-i\phi})^{m_1+n_1-2j} (1 + \mathrm{e}^{-i\phi})^{m_1-n_1+2j} \right|^2. \tag{57}$$

The presence of ϕ in Eq. (57) means that, after the application of the Fourier transform, photon counting can be used to determine ϕ. The classical Fisher information for these distributions is simply given by

$$\mathbf{I}_C = \sum_{\substack{m_1, m_2 \\ =0}}^{\infty} \frac{(\nabla_{\boldsymbol{\theta}} p_{\mathrm{out}}(m_1, m_2)) (\nabla_{\boldsymbol{\theta}} p_{\mathrm{out}}(m_1, m_2))^{\mathrm{T}}}{p_{\mathrm{out}}(m_1, m_2)}, \tag{58}$$

where $\nabla_{\boldsymbol{\theta}} = (\frac{\partial}{\partial \theta_1}, \frac{\partial}{\partial \theta_2}, \ldots, \frac{\partial}{\partial \theta_l})^{\mathrm{T}}$.

References

[1] A. Alonso, S. Arribas, and C. Martínez-Roger. *Astron. Astrophys. Suppl. Ser.*, 139(2):335–358, 1999.

[2] E. Bagan, M. A. Ballester, R. D. Gill, A. Monras, and R. Muñoz-Tapia. *Physical Review A*, 73(3): 1–19, 2006.

[3] O. E. Barndorff-Nielsen and R. D. Gill. *Journal of Physics A: Mathematical and General*, 33(24): 4481, 2000.

[4] Samuel L. Braunstein and Carlton M. Caves. *Phys. Rev. Lett.*, 72:3439–3443, 1994.

[5] B. M. Escher, L. Davidovich, N. Zagury, and R. L. de Matos Filho. *Phys. Rev. Lett.*, 109:190404, 2012.

[6] D. J. Fixsen. *The Astrophysical Journal*, 707(2): 916–920, 2009.

[7] Yang Gao and Hwang Lee. *The European Physical Journal D*, 68(11):347, 2014.

[8] Richard D. Gill and Mădălin I. Guţă. DOI: 10.1214/12-IMSCOLL909.

[9] Richard D. Gill and Serge Massar. *Physical Review A*, 61(4):042312, 2000.

[10] Vittorio Giovannetti, Seth Lloyd and Lorenzo Maccone. *Nature Photonics*, 5(4):222-229, 2011.

[11] C. Helstrom. *IEEE Transactions on Information Theory*, 14(2):234–242, 1968.

[12] Carl W. Helstrom. *Quantum Detection and Estimation Theory.* Academic Press, 1976.

[13] C.W. Helstrom. *Physics Letters A*, 25(2):101–102, 1967.
[14] Alexander Holevo. *Probabilistic and Statistical Aspects of Quantum Theory*. Edizioni Della Normale, 2011.
[15] R. L. Hudson. *Journal of Applied Probability*, 10 (3):502–509, 1973.
[16] Marcin Jarzyna and Rafał Demkowicz-Dobrzański. *New Journal of Physics*, 17(1):013010, 2015.
[17] Steven M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993.
[18] P. A. Knott, T. J. Proctor, A. J. Hayes, J. P. Cooling, and J. A. Dunningham. *Phys. Rev. A*, 93: 033859, 2016.
[19] Pieter Kok, Jacob Dunningham, and Jason F. Ralph. *Physical Review A, 95, 012326, 2017.*
[20] Cosmo Lupo and Stefano Pirandola *Phys. Rev. Lett.*, 117:190802, 2016.
[21] Katarzyna Macieszczak, Martin Fraas, and Rafal Demkowicz-Dobrzański. *New Journal of Physics*, 16, 2014.
[22] Leonard Mandel and Emil Wolf. *Optical Coherence and Quantum Optics*. Cambridge University Press, 1995.
[23] Paulina Marian and Tudor A. Marian. *Phys. Rev. A*, 93:052330, 2016.
[24] Michael R. Moldover, Weston L. Tew, and Howard W. Yoon. *Nature Publishing Group*, 12 (1):7–11, 2016.
[25] Alex Monras. arXiv:1303.3682, 2013.
[26] Alex Monras and Fabrizio Illuminati. *Physical Review A - Atomic, Molecular, and Optical Physics*, 83(1):1–12, 2011.
[27] Ranjith Nair and Mankei Tsang. *The Astrophysical Journal*, 808(2):125, 2015.
[28] Matteo G. A. Paris. *International Journal of Quantum Information*, 07:125–137, 2009.
[29] M. E. Pearce, T. Mehringer, J. von Zanthier, and P. Kok. *Physical Review A*, 92(4):043831, 2015.
[30] Sammy Ragy, Marcin Jarzyna, and Rafał Demkowicz-Dobrzański. *Phys. Rev. A*, 94:052108, 2016.
[31] Magdalena Szczykulska, Tillmann Baumgratz, and Animesh Datta. *Advances in Physics: X*, 1(4):621–639, 2016.
[32] Mankei Tsang. *Optica*, 2(7):646–653, 2015.
[33] Mankei Tsang, Ranjith Nair, and Xiao-Ming Lu. *Phys. Rev. X*, 6:031033, 2016.
[34] Dominik Šafránek, Antony R Lee, and Ivette Fuentes. *New Journal of Physics*, 17(7):073016, 2015.