OPEN ACCESS

University of BRISTOL

Liu, N., & Cable, H. (2017). Quantum-enhanced multi-parameter estimation for unitary photonic systems. Quantum Science and Technology, 2, [025008]. DOI: 10.1088/2058-9565/aa6fea

**University of Bristol - Explore Bristol Research**

**General rights**

You may also be interested in:

# Quantum Science and Technology

CrossMark

# Quantum-enhanced multi-parameter estimation for unitary photonic systems

## Nana Liu[1,2,3,5] and Hugo Cable[4]

1    Clarendon Laboratory, Department of Physics, University of Oxford, Oxford OX1 3PU, United Kingdom
2    Singapore University of Technology and Design, 8 Somapah Road, 487372 Singapore
3    Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, 117543 Singapore
4    Quantum Engineering Technology Labs, H. H. Wills Physics Laboratory and Department of Electrical and Electronic Engineering, University of Bristol, BS8 1FD, United Kingdom
5    Author to whom any correspondence should be addressed.

**E-mail:** Nana.Liu@physics.ox.ac.uk, Nana.Liu@quantumlah.org and Hugo.Cable@bristol.ac.uk

## Abstract

Precise device characterisation is a fundamental requirement for a large range of applications using photonic hardware, and constitutes a multi-parameter estimation problem. Estimates based on measurements using single photons or classical light have precision which is limited by shot noise, while quantum resources can be used to achieve sub-shot-noise precision. However, there are many open questions with regard to the best quantum protocols for multi-parameter estimation, including the ultimate limits to achievable precision, as well as optimal choices for probe states and measurements. In this paper, we develop a formalism based on Fisher information to tackle these questions for setups based on linear-optical components and photon-counting measurements. A key ingredient of our analysis is a mapping for equivalent protocols defined for photonic and spin systems, which allows us to draw upon results in the literature for general finite-dimensional systems. Motivated by the protocol in Zhou, *et al* Optica 2, 510 (2015), we present new results for quantum-enhanced tomography of unitary processes, including a comparison of Holland-Burnett and NOON probe states.

## 1. Introduction

Advances in precision measurement are playing an ever more important role in technological development. From biological imaging [1], quantum clocks [2, 3] quantum computing [4], thermometry [5], to the recent detection of gravitational waves [6], there is an increasing demand for higher precision in parameter estimation schemes using light. Quantum resources have been shown to serve a crucial role in pushing beyond the precision limits available to classical probes, especially in single-parameter estimation [7]. The quantum advantage in multi-parameter estimation is less well studied. However, applications like imaging, linear-optical quantum computing and characterisation of optical fibres for quantum communication require the simultaneous estimation of multiple parameters. Thus, a question of chief interest for future applications is to find the quantum resources for multi-parameter problems that offer optimal precision while remaining experimentally accessible.

One primary resource in parameter estimation is particle number $N$. Parameter estimation using classical resources is constrained by the shot-noise limit [7], for which precision (mean-square error) scales as $1/N$ [7]. This is true in conventional multi-parameter estimation using single-photon probes (process tomography) [8]. There are situations, however, such as when probing delicate samples, where the use of high $N$ is undesirable [1].

It is well known for the idealised case, when the effects of particle losses and decoherence can be ignored, that quantum resources enable up to quadratic improvement in precision for both single-parameter estimation and multi-parameter estimation. This is called Heisenberg scaling, i.e., precision scales like $\mathcal{O}(1/N^2)$ [9]. In single-parameter estimation, NOON states uniquely achieve the optimal precision (Heisenberg limit) using

photon-number-counting measurements when again losses and decoherence are neglected [10]. In contrast, for the task of estimating a set of parameters that fully characterise any linear-optical process, it has not previously been shown what conditions must be satisfied to achieve optimal precision, or even what this optimal precision is.

One of the first experimental demonstrations of Heisenberg scaling for a general linear-optical process was recently performed [11] (for $N = 4$). It was based on a new protocol for characterising an unknown two-mode linear-optical process, using Holland-Burnett states [12] and photon-number-counting measurements. This is equivalent to estimating the independent parameters of an unknown $SU(2)$ matrix. The analysis of [11] uses process fidelity to quantify precision, as is typical for process tomography. However, using process fidelity, it is difficult to establish the optimality of the protocol, which is essential to enable comparisons between alternative choices of probe states and measurement schemes.

A theoretical tool very well suited to this analysis is the Fisher information formalism, which is already much used in estimation theory [7, 13]. For photonic systems, Fisher information has already been exploited for some very specific cases of multi-parameter estimation where substantial simplifications occur. For example, when parameters are associated with commuting operations [14–16] or single-parameter estimation with environment interaction [17, 18]. There has been active research on spin systems in related contexts. A succession of theoretical studies, using fidelity measures to characterise precision [19–22] and using Fisher information [23–25] have demonstrated how Heisenberg scaling can be achieved. Theoretical studies of $SU(2)$-estimation using the Fisher information matrix have provided mathematical conditionals for achieving optimal precision [26, 27].

In this paper, we develop theoretical machinery to explore $SU(2)$-estimation protocols for linear-optical setups using quantum states. We provide simple conditions to test which photonic states are optimal (extending results in [24] for spin systems). We also interpret these results in terms of optical interferometry. As two important examples, we show that both Holland-Burnett states and NOON states are optimal, but we find that neither are optimal using photon-number-counting measurements. Furthermore, we demonstrate that the precision in this context is dependent on the unitary itself. In addition to having important practical implications, these results show that multi-parameter estimation cannot be considered a simple generalisation of single-parameter estimation.

We also introduce a mapping between photonic and spin states that allow us to translate the results for photonic systems to spin systems and vice versa. This analogy allows us to compare the spin analogue of our optical protocol to those of existing multi-parameter estimation schemes using spin systems. For instance, this makes it possible to prove that a probe state used in [25] is indeed optimal, which was suspected but not proved. We also show that the spin analogue of our protocol contains a larger class of optimal states compared to [25].

After establishing our mapping from photonic to spin states and processes in section 2.1, we turn to a brief introduction to the experimental protocol in [11] in section 2.2. We then introduce the basics of the Fisher information formalism in section 2.3. In section 3 we extend previous results in the literature to find the optimal precision and the conditions for optimal states in our protocol. We also discuss the implications of these results for spin systems. In section 4, we apply our results to study the special cases of Holland-Burnett states and NOON states under photon-number-counting measurements before summarising our main results and future directions in section 5.

# 2. Background

## 2.1. Equivalent protocols for photonic and spin systems

We now establish a mapping between an $N$-particle two-mode linear-optical process and a process with $N$ spin-1/2 particles. This is important for making a formal analogy for protocols specified for photonic and spin systems. It can be used for translating results on multi-parameter estimation using spin systems to the context of photonic multi-parameter estimation. For standard mappings between bosonic and spin states and operators, see [28, 29].

We begin with an $N$-particle two-mode photonic state $|M, N - M\rangle$, where $M$ is an integer $0 \leqslant M \leqslant N$. There is a one-to-one correspondence between this state and an $N$-particle spin-1/2 state that remains invariant with respect to any particle exchange (i.e., symmetric):

$$|M, N - M\rangle \longleftrightarrow \frac{1}{\sqrt{\binom{N}{M}}} \sum_j \Pi_j (|\uparrow\rangle^{\otimes M} \otimes |\downarrow\rangle^{\otimes(N-M)}) \equiv |\xi_0\rangle_{\text{spin}}. \tag{1}$$

The summation $\sum_j \Pi_j$ is over all the possible permutations of the product states and $|\xi_0\rangle_{\text{spin}}$ is also known as a Dicke state [30]. For concreteness, we choose $|\uparrow\rangle, |\downarrow\rangle$ to be the spin-up and spin-down eigenstates of Pauli matrix $\sigma_z$. We denote the creation operators for the two photonic modes by $a^\dagger$ and $b^\dagger$, obeying commutation relations $[a, a^\dagger] = [b, b^\dagger] = 1$. The creation operators corresponding to the up and down spin states are

**Figure 1.** Linear-optical process and corresponding evolution of spin system. Mapping between a two-mode linear-optical process $U$ acting on $N$ photons, and the analogous process with $N$ spin-$\frac{1}{2}$ particles each undergoing evolution $U_s$. Photonic superposition states correspond to superpositions of spin states that are symmetric under particle exchange. Measurement of $M'$ (or $N - M'$) photons in the first (or second) mode correspond to measurement of $M'$ spin-up (or $N - M'$ spin-down) particles in the spin picture (where the ordering is ignored).

represented by $a_\uparrow^\dagger$ and $a_\downarrow^\dagger$. These satisfy the anticommutation relations $\{a_\uparrow^\dagger, a_\uparrow\} = \mathbf{1} = \{a_\downarrow^\dagger, a_\downarrow\}$ and where all other anticommutation relations vanish. In the single-particle case, we can make the correspondence $a^\dagger|0, 0\rangle \longleftrightarrow a_\uparrow^\dagger|0\rangle = |\uparrow\rangle$ and $b^\dagger|0, 0\rangle \longleftrightarrow a_\downarrow^\dagger|0\rangle = |\downarrow\rangle$. Extending this to $N$ particles gives the mapping between photonic states and spin states that we require.

This mapping also allows us to describe a transformation of the two-mode photonic state under unitary operator $U$ in terms of the evolution of $N$ spin-1/2 particles, described by the unitary operator $U_s$ (see figure 1). It can be shown that the following correspondence holds (for a derivation, see appendix A)

$$U|M, N - M\rangle \longleftrightarrow U_s^{\otimes N}|\xi_0\rangle_{\text{spin}}. \tag{2}$$

Each two-mode linear-optical unitary corresponds to a two-by-two unitary matrix. In the remainder of this paper, we will disregard any global phase for $U$, which can then be characterised by three parameters, and represented by an $SU(2)$ matrix.

We can complete the analogy by mapping photon-number-counting measurement to measurements of the spin system. We note that a projective measurement $|M', N - M'\rangle\langle M', N - M'|$ on the photonic state (where $M'$ is an integer $0 \leqslant M' \leqslant N$) corresponds in the spin picture to measurement of $M'$ spin-up particles and $N - M'$ spin-down particles, where there are $N!/(M'!(N - M')!)$ equivalent measurement patterns.
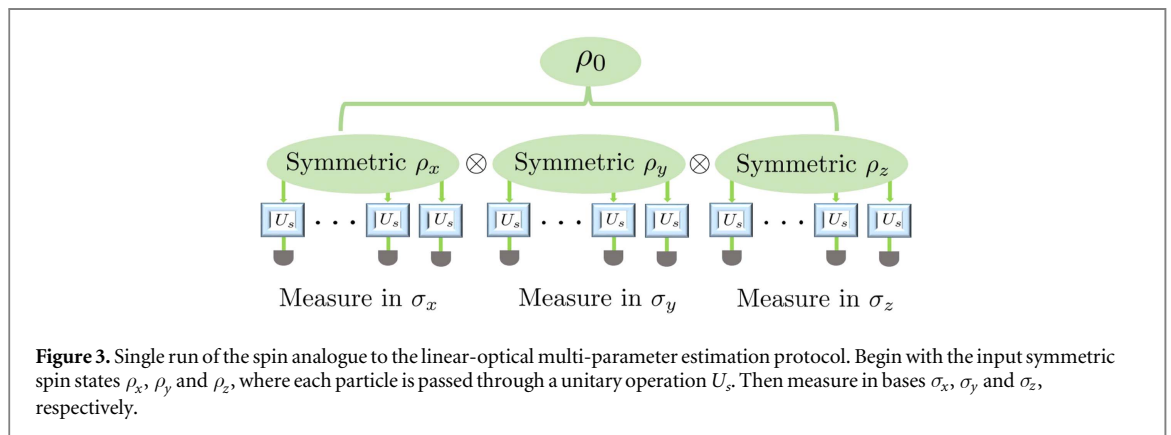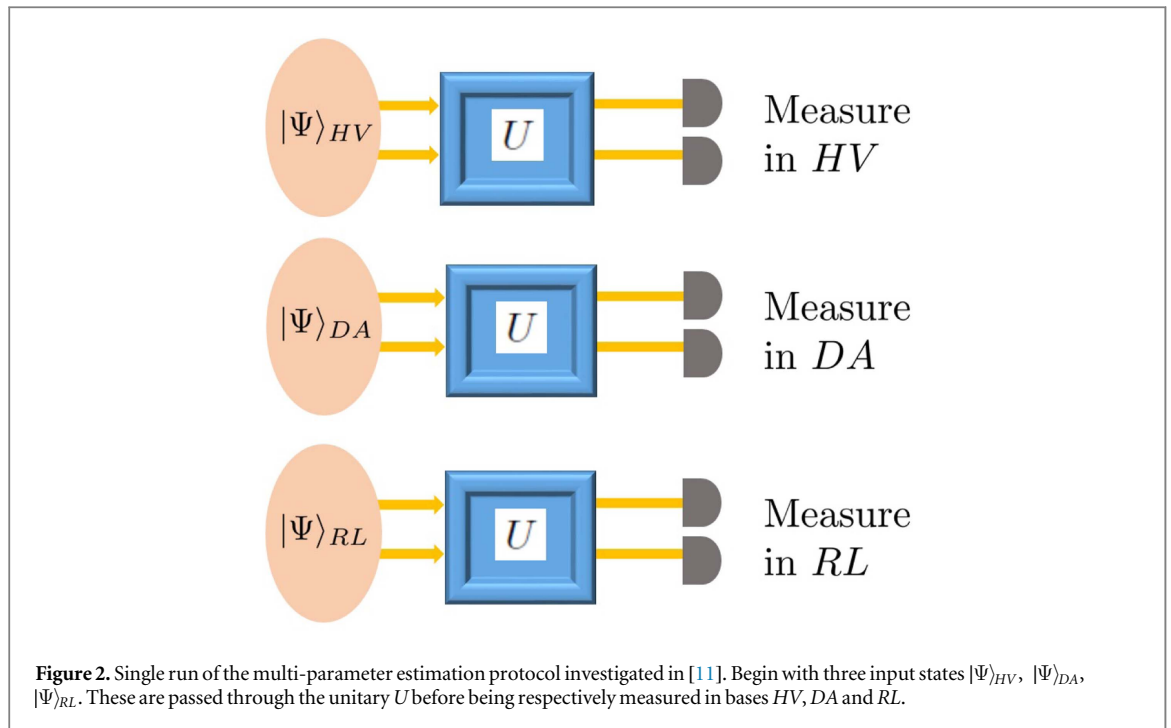
We can use this mapping to show that the $N$-particle NOON states $(|N, 0\rangle + |0, N\rangle)/\sqrt{2}$ map to $N$-particle GHZ [31] states $(|0\rangle^{\otimes N} + |1\rangle^{\otimes N})/\sqrt{2}$. This explains why both NOON states and GHZ states have been found to be optimal states (i.e., achieving optimal precision) in single-parameter estimation [32], despite being used in different types of setups. Another example is the correspondence between $N$-particle Holland-Burnett states $|N/2, N/2\rangle$ and symmetric Dicke states with $N/2$ excitations [32].

## 2.2. Quantum-enhanced protocol for unitary estimation

To identify an unknown optical process, process tomography is traditionally used. It relies on single-photon probes or classical light and is shot-noise limited. However, when using non-classical multi-photon probe states, much greater precision per number of photons can be achieved. In this section, we briefly describe the multi-photon probe scheme for $SU(2)$-estimation recently performed in [11].

We begin with a two-mode $N$-particle photonic state $|\Psi\rangle_{HV} = |M, N - M\rangle_{HV}$ in the $HV$ (horizontal and vertical) polarisation basis. This state is passed through a linear-optical process, which can be characterised by a $SU(2)$ matrix. It is possible to recover the probability distributions of photon numbers in each of the two modes after passing through the unknown unitary by using photon-number-counting measurements with respect to the $HV$ basis. This procedure can be repeated with respect to the $DA$ (diagonal and anti-diagonal) and $RL$ (right and left-circular) polarisation bases. By definition $|1, 0\rangle_{DA} = (1/\sqrt{2})(|1, 0\rangle_{HV} + |0, 1\rangle_{HV})$, $|0, 1\rangle_{DA} = (1/\sqrt{2})(|1, 0\rangle_{HV} - |0, 1\rangle_{HV})$ and $|1, 0\rangle_{RL} = (1/\sqrt{2})(|1, 0\rangle_{HV} + i|0, 1\rangle_{HV})$, $|0, 1\rangle_{RL} = (1/\sqrt{2})(|1, 0\rangle_{HV} - i|0, 1\rangle_{HV})$. A *single run* of the $SU(2)$-estimation protocol is defined as the procedure above performed with respect to all three $HV$, $DA$ and $RL$ polarisation bases. The three independent parameters of the unknown $SU(2)$ matrix can then be fully recovered from the set of three photon number probability distributions (see [11]). See figure 2 for a diagram representing a single run of this protocol.

When this protocol is performed $\nu$ times, the total number $n$ of input photons used is thus $3N\nu$. It can be experimentally shown [11] that, for up to $N = 4$, the precision for estimating each of a full set of parameters which characterise the unknown unitary using Holland-Burnett input states $|\Psi\rangle = |N/2, N/2\rangle$ is $\mathcal{O}(1/(\nu N^2))$. This is Heisenberg scaling. This is contrasted with the shot-noise limited precision using

**Figure 2.** Single run of the multi-parameter estimation protocol investigated in [11]. Begin with three input states $|\Psi\rangle_{HV}$, $|\Psi\rangle_{DA}$, $|\Psi\rangle_{RL}$. These are passed through the unitary $U$ before being respectively measured in bases $HV$, $DA$ and $RL$.



**Figure 3.** Single run of the spin analogue to the linear-optical multi-parameter estimation protocol. Begin with the input symmetric spin states $\rho_x$, $\rho_y$ and $\rho_z$, where each particle is passed through a unitary operation $U_s$. Then measure in bases $\sigma_x$, $\sigma_y$ and $\sigma_z$, respectively.

single-photon probes, which has scaling $\mathcal{O}(1/(\nu N))$. In [11], the performance of the protocol for different choices of unitary and probe states was compared using process fidelity.

We note that the HV, DA and RL polarisation bases respectively map to the $z$, $x$ and $y$ bases in the spin case. We denote here $\rho_x$, $\rho_y$ and $\rho_z$ as the symmetric spin state analogues to the photonic input states in bases $HV$, $DA$ and $RL$. The probe state $\rho_0$ for a single run of the analogous spin system protocol (i.e., $\nu = 1$) is of the form $\rho_0 = \bigotimes_{k=x,y,z} \rho_k$. See figure 3 for a spin-system analogue of the linear-optical unitary estimation protocol.

### 2.3. Fisher information

*2.3.1. Cramér-Rao inequality for single and multi-parameter estimation*
In any scheme to estimate unknown parameters, it is useful to bound the variance of those parameters as a way of characterising the precision. The inverse of a quantity known as the Fisher information provides a means to bound variances in local parameter estimation, i.e., beginning from a rough estimate of the parameters and making this estimate more precise. This type of bound is provided by the Cramér-Rao inequality. It is this bound that makes Fisher information so crucial and the reason we choose this measure to characterise precision.

In any given process, the goal is to find the initial state and measurement maximising the Fisher information while being subject to some given constraints of one's resources. Here we take this resource to be the total number of particles used in probe states over all runs of the estimation protocol.

To estimate a single unknown parameter $\theta$, the Cramér-Rao inequality states [33] that the variance of the unbiased estimators of $\theta$ achievable is

$$\delta\theta^2 \geqslant \frac{1}{F_\theta}. \tag{3}$$

By an unbiased estimator we mean that the average value is exactly $\theta$. The Fisher information is denoted $F_\theta$ and is defined by

$$F_\theta = \sum_{M_d} \frac{1}{P(M_d; \theta)}\left(\frac{\partial P(M_d; \theta)}{\partial\theta}\right)^2. \tag{4}$$

The probability the final state of the protocol $\rho_\theta$ has the (detected) measurement outcome $M_d$ is denoted $P(M_d; \theta)$.

This bound can be made tighter by introducing quantum Fisher information $I_\theta$. This is defined as the maximum Fisher information with respect to all possible final measurements. It can be shown that $F_\theta \leqslant I_\theta = \mathrm{tr}(\rho_\theta \lambda_\theta^2)$, where $\lambda_\theta$ is the symmetric logarithmic derivative (SLD) [34] and is defined by $\partial_\theta \rho_\theta = (1/2)(\lambda_\theta \rho_\theta + \rho_\theta \lambda_\theta)$. This introduces the quantum Cramér-Rao inequality

$$\delta\theta^2 \geqslant \frac{1}{F_\theta} \geqslant \frac{1}{I_\theta}, \tag{5}$$

where the equality can always be saturated and the optimal precision is achievable by known initial states and final measurements [34].

The optimal achievable precision from the quantum Cramér-Rao inequality is $\delta\theta^2 = 1/(\nu N^2)$. This is uniquely achieved using the $N$-particle NOON state. The optimal measurements, which include photon-number-counting measurements, can be expressed in terms of the eigenvectors of the SLD. Another state that can achieve Heisenberg scaling (i.e., $\delta\theta^2 = \mathcal{O}(1/(\nu N^2))$), but with sub-optimal precision, is the Holland-Burnett state [10]. Other sub-optimal states with Heisenberg scaling include Yurke states [35–37], amongst many others [10, 35].

In the estimation of *multiple* parameters $\{\theta_\alpha\}$, precision is captured by the covariance matrix $C$, which is defined by $C_{\alpha\beta} \equiv \langle\theta_\alpha\theta_\beta\rangle - \langle\theta_\alpha\rangle\langle\theta_\beta\rangle$ and $\langle\cdot\rangle$ denotes an average over all measurement outcomes. Note that the covariance matrix reduces to the variance in the case of a single parameter. The covariance matrix can be bounded by the inverse of the Fisher information matrix $F$, defined by

$$F_{\alpha\beta} = \sum_{M_d}(1/P(M_d; \{\theta_\gamma\}))(\partial P(M_d; \{\theta_\gamma\})/\partial\theta_\alpha)(\partial P(M_d; \{\theta_\gamma\})/\partial\theta_\beta). \tag{6}$$

$F_{\alpha\beta}$ is in turn upper bounded by the quantum Fisher information matrix $I$, which defined by [38–40]

$$I_{\alpha\beta} = \mathrm{Re}\,\mathrm{tr}(\rho_\theta \lambda_\alpha \lambda_\beta), \tag{7}$$

where for pure states $\lambda_\alpha = 2\partial\rho_\theta/\partial\theta_\alpha$. A multi-parameter quantum Cramér-Rao inequality can thus be written as

$$\mathrm{tr}(C) \geqslant \mathrm{tr}(F^{-1}) \geqslant \mathrm{tr}(I^{-1}). \tag{8}$$

This follows from the inequalities $C \succcurlyeq F^{-1} \succcurlyeq I^{-1}$, where $C - I^{-1} \succcurlyeq 0$ denotes $C - I^{-1}$ is a positive semi-definite matrix [38–40]. When the equality in $C - I^{-1} \succcurlyeq 0$ is satisfied, this is known as saturating the multi-parameter quantum Cramér-Rao inequality and implies $\mathrm{tr}(C) = \mathrm{tr}(F^{-1}) = \mathrm{tr}(I^{-1})$.

However, unlike in single-parameter estimation, it is not always possible to always saturate the multi-parameter quantum Cramér-Rao inequality. This is because the operators corresponding to optimal measurements (constructed from the eigenvectors of $\lambda_\alpha$) corresponding to different parameters may not commute. Hence it is not possible to implement them simultaneously and thus optimise precision with respect to all the parameters. However, there are cases where saturation is possible and for pure states this is investigated in section 3. For a discussion of differences between cases with pure and mixed states, see [41].

For recent reviews on this subject, we refer interested readers to [32, 40, 42] and references therein.

### 2.3.2. Quantum Fisher information matrix

Now we derive the general form for protocols using pure probe states. We consider a general pure probe state $\rho_0$, unitary transformation $\tilde{U}$, and final state $\rho_\theta = \tilde{U}\rho_0\tilde{U}^\dagger$. It is convenient to define $\tau_\alpha = i\tilde{U}^\dagger \partial\tilde{U}/\partial\theta_\alpha$, which is the generator of $\tilde{U}$ and which must be Hermitian. Then the quantum Fisher information matrix is given by

$$I_{\alpha\beta}(\rho_0) = 4\,\mathrm{Re}\left(\mathrm{tr}\left(\rho_\theta \frac{\partial\rho_\theta}{\partial\theta_\alpha}\frac{\partial\rho_\theta}{\partial\theta_\beta}\right)\right). \tag{9}$$

Using the unitarity of $\tilde{U}$, the cyclic-permutation invariance of the trace and purity $\rho_0^2 = \rho_0$, it is straightforward to show $\mathrm{tr}(\rho_\theta \lambda_\alpha \lambda_\beta) = 4[\mathrm{tr}(\rho_0 \tau_\alpha \tau_\beta) - \mathrm{tr}(\rho_0 \tau_\alpha \rho_0 \tau_\beta)]$. Since $\rho_0$ is again pure, $\mathrm{tr}(\rho_0 \tau_\alpha \rho_0 \tau_\beta)] = \mathrm{tr}(\rho_0 \tau_\alpha)\mathrm{tr}(\rho_0 \tau_\beta)$ and

$$I_{\alpha\beta}(\rho_0) = 4\,\mathrm{Re}[\mathrm{tr}(\rho_0\tau_\alpha\tau_\beta)] - 4\mathrm{tr}(\rho_0\tau_\alpha)\mathrm{tr}(\rho_0\tau_\beta). \tag{10}$$

We turn now to the spin protocol considered in this paper, where we saw $\rho_0$ is an $n$-particle pure state undergoing unitary transformation $\tilde{U} = U_s^{\otimes n}$. Imposing no other restrictions on $\rho_0$ other than purity, we want to express $I_{\alpha\beta}$ in terms of the generators of $U_s$, which are $t_\alpha \equiv iU_s^\dagger \partial U_s/\partial\theta_\alpha$. We observe the relations $\tau_\alpha = t_\alpha \otimes \mathbf{1}^{\otimes(n-1)} + \mathbf{1} \otimes t_\alpha \otimes \mathbf{1}^{\otimes(n-2)} + \ldots + \mathbf{1}^{\otimes(n-1)} \otimes t_\alpha$ (which has $n$ terms altogether) and $\tau_\alpha\tau_\beta = t_\alpha t_\beta \otimes \mathbf{1}^{\otimes(n-1)} + t_\alpha \otimes t_\beta \otimes \mathbf{1}^{\otimes(n-2)} + t_\alpha \otimes \mathbf{1}^{\otimes(n-1)} \otimes t_\beta + \ldots + \mathbf{1}^{\otimes(n-1)} \otimes t_\alpha t_\beta$ (which has $n^2$ terms altogether). These relations together with standard trace equalities yield

$$I_{\alpha\beta}(\rho_0) = 4\,\mathrm{Re}\sum_{i=1}^{n}\mathrm{tr}(\mathrm{tr}_{[i]}(\rho_0)t_\alpha t_\beta) + 4\sum_{i\neq j=1}^{n}\mathrm{tr}(\mathrm{tr}_{[i,j]}(\rho_0)(t_\alpha\otimes t_\beta)) - 4\left(\sum_{i=1}^{n}\mathrm{tr}(\mathrm{tr}_{[i]}(\rho_0)t_\alpha)\right)\left(\sum_{j=1}^{n}\mathrm{tr}(\mathrm{tr}_{[j]}(\rho_0)t_\beta)\right). \tag{11}$$

Here the notation $\mathrm{tr}_{[i,j]}$ denotes a trace over all particles *except* those particles labelled by positions $i$, $j$. The ordering of $i$, $j$ matters so that the reduced state $\mathrm{tr}_{[j,i]}(\rho_0)$ corresponds to $\mathrm{tr}_{[i,j]}(\rho_0)$ with a swap of the two subsystems $i$ and $j$. Interestingly, this means that the total Fisher information is dependent only on the one-particle and two-particle reduced density matrices of $\rho_0$.

Our equation (11) holds for any unitary $U_s$. In a special instance when $U_s = \exp(iH)$, where $H = \sum_\alpha \theta_\alpha h_\alpha$ and $h_\alpha$ is independent of unknown parameters $\theta_\alpha$, equation (11) agrees with [25]. We also note that [24] considers a closely-related case where the protocol incorporates an ancilla and $\rho_0$ is symmetric.

### 2.3.3. Parameterisation

In section 2.1, we saw that the unitary operator $U_s$ is sufficient to describe a two-mode linear-optical process and can be represented by a $SU(2)$ matrix. We now need to define a parameterisation of this $SU(2)$ matrix. This helps us compute the quantum Fisher information and Fisher information matrices in a way which allows us to meaningfully quantify the corresponding notions of statistical information about $U_s$. To this end, we can express this parameterisation as a restriction on the generators $\{t_\alpha\}$ in the following way.

Since we are performing a local parameter estimation of $U_s$, we can define $U_s$ in terms of a local expansion about some known unitary operator $U_s^{(0)}$. The higher precision to which this expansion is known, the better estimate we have of $U_s$. $U_s$ is dependent only on the three *unknown* parameters $\{\theta_\alpha\}$. Let $U_s^{(0)}$ be defined only in terms of the three *known* parameters $\{\theta_\alpha^{(0)}\}$. To linear order, a Taylor expansion about $U_s^{(0)}$ can be written as $U_s(\{\theta_\alpha\}) \equiv U_s^{(0)}(\{\theta_\alpha^{(0)}\})\left(\mathbf{1} - i\sum_m t_\alpha|_{\theta_\alpha=\theta_\alpha^{(0)}}(\theta_\alpha - \theta_\alpha^{(0)})\right)$.

One good parameterisation is where precision is independent of the particular $U_s^{(0)}$, i.e., the value of the initial guess for $U_s$. This is achieved when $t_\alpha|_{\theta_\alpha=\theta_\alpha^{(0)}}$ is independent of $\{\theta_\alpha^{(0)}\}$, which means $t_\alpha$ is independent of $\{\theta_\alpha\}$. A natural decision is choosing $\{t_\alpha\}$ to be proportional to the Pauli spin matrices $\{\sigma_\alpha\}$ (we note that this parameterisation was also adopted in [24] and provides simplification of the analysis). We will call parameters $\{\theta_\alpha\}$ satisfying $t_\alpha = \sigma_\alpha/\sqrt{2}$ *locally independent parameters*. We adopt this terminology since in every local region about some $U_s^{(0)}$, the generators $\{t_\alpha\}$ are independent of $U_s^{(0)}$. This is not true in more general parameterisations, like the Euler angle parameterisation.

## 3. Optimal precision and conditions for optimal states

In this section we demonstrate the necessary and sufficient conditions our photonic-probe states must satisfy to reach optimal precision and we find what this optimal precision is. Our results also allow us to readily identify these photonic states.

Reaching the optimal precision allowed by the quantum Cramér-Rao inequality consists of two parts: saturating the quantum Cramér-Rao inequality and attaining the smallest value of the trace of the inverse quantum Fisher information matrix. We examine these conditions and their implications in sections 3.2 and 3.3, respectively.

We proceed by considering protocols for spin systems and deriving optimality results along similar lines to [24] for ancilla-based protocols. Then utilising our mapping introduced in section 2.1, we translate these results in terms of photonic states and processes. We find they lead to novel and important consequences for optical metrology which we explore.

We also show how the spin-system analogue to our photonic unitary estimation protocol allows for a much wider class of optimal states than has been identified [25].

### 3.1. Quantum Fisher information matrix for unitary estimation

The spin-system analogue to our unitary estimation protocol is represented in figure 3. It begins with a $n$-particle input state $\rho_0 = \bigotimes_{k=x,y,z}\rho_k$, where $\rho_k$ are $N$-particle input states. For a single run of the protocol, $I_{\alpha\beta}$ is a sum

of the quantum Fisher information matrices with respect to each sub-protocol. Thus,

$$I_{\alpha\beta}(\rho_0) = \sum_{k=x,y,z} I_{\alpha\beta}(\rho_k). \tag{12}$$

This additivity condition follows from equation (9), the purity of $\rho_0$ and $\tilde{U} = U_s^{\otimes 3N}$.

The state $\rho_0$ is in general not a symmetric state, while each $\rho_k$ is symmetric (see section 2.1). This means all partial traces of $\rho_k$ depend only on the number of subsystems traced out and not on *which* subsystem is traced out. Therefore all one-particle reduced states of $\rho_k$ are identical and we can define $\rho_k^{[1]} \equiv \mathrm{tr}_{[i]}(\rho_k)$ for any $i$. All two-particle reduced states are also identical and we define $\rho_k^{[2]} \equiv \mathrm{tr}_{[i,j]}(\rho_k)$ for any $i$, $j$ where $i \neq j$. Applying our locally independent parameterisation $t_\alpha \equiv \sigma_\alpha / \sqrt{2}$ in equations (11) and (12), the quantum Fisher information matrix for our protocol simplifies to the form

$$I_{\alpha\beta}(\rho_0) = \mathrm{Re}\{2N\,[\mathrm{tr}(\rho_{\mathrm{tot}}^{[1]}\sigma_\alpha\sigma_\beta)]\} + 2N(N-1)\mathrm{tr}(\rho_{\mathrm{tot}}^{[2]}(\sigma_\alpha \otimes \sigma_\beta)) - 2N^2 \sum_{k=x,y,z} \mathrm{tr}(\rho_k^{[1]}\sigma_\alpha)\mathrm{tr}(\rho_k^{[1]}\sigma_\beta), \tag{13}$$

where $\rho_{\mathrm{tot}}^{[1]} \equiv \sum_{k=x,y,z} \rho_k^{[1]}$ and $\rho_{\mathrm{tot}}^{[2]} \equiv \sum_{k=x,y,z} \rho_k^{[2]}$.

### 3.2. Saturating the quantum Cramér-Rao inequality

Saturation of the quantum Cramér-Rao inequality is a non-trivial constraint in the multi-parameter setting. For pure states, it can be shown that a necessary and sufficient condition to attain this saturation is to satisfy $\mathrm{Im}[\mathrm{tr}(\rho_\theta \lambda_\alpha \lambda_\beta)] = 0$ [27]. Since $\lambda_\alpha$ are hermitian, this condition is equivalent to

$$\mathrm{tr}(\rho_\theta [\lambda_\alpha, \lambda_\beta]) = 0. \tag{14}$$

Following similar arguments as in section 2.3.2, for arbitrary $\tilde{U}$ and pure states $\rho_\theta = \tilde{U}\rho_0 \tilde{U}^\dagger$, equation (14) is equivalent to

$$\mathrm{tr}(\rho_0 [\tau_\alpha, \tau_\beta]) = 0. \tag{15}$$

For the scenario pictured in figure 3 with $\tilde{U} = U^{\otimes 3N}$ we now obtain $\mathrm{tr}(\rho_0 \tau_\alpha \tau_\beta) = \sum_{i=1}^{3N} \mathrm{tr}[\mathrm{tr}_{[i]}(\rho_0)(t_\alpha t_\beta)] + \sum_{i \neq j=1}^{3N} \mathrm{tr}[\mathrm{tr}_{[i,j]}(\rho_0)(t_\alpha \otimes t_\beta)]$. Using $\rho_0 = \bigotimes_{k=x,y,z} \rho_k$, equation (15) thus reduces to

$$\mathrm{tr}(\rho_{\mathrm{tot}}^{[1]}[t_\alpha, t_\beta]) = 0. \tag{16}$$

For the locally independent parameterisation this becomes $\mathrm{tr}(\rho_{\mathrm{tot}}^{[1]}[\sigma_\alpha, \sigma_\beta]) = 0$, which is uniquely satisfied when $\rho_{\mathrm{tot}}^{[1]} \propto \mathbf{1}$. We can write $\rho_z^{[1]} = \mathbf{1}/2 + \sum_{k=x,y,z} b_k \sigma_k$ where $b_k$ are constants. Then using $\rho_x^{[1]} = h\rho_z^{[1]}h^\dagger$ and $\rho_y^{[1]} = h_c \rho_z^{[1]} h_c^\dagger$, where $h = (1/\sqrt{2})\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ and $h_c = (1/\sqrt{2})\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$, we find $\rho_{\mathrm{tot}}^{[1]} = (3/2)\mathbf{1} + (b_x + b_y + b_z)\sigma_x + b_z \sigma_y + (2b_x + b_z)\sigma_z$. Thus $\rho_{\mathrm{tot}}^{[1]} = (3/2)\mathbf{1}$ exactly when $b_k = 0$ for all $k = x$, $y$, $z$. Therefore a sufficient and necessary condition for the quantum Cramér-Rao inequality to be saturated is

$$\rho_z^{[1]} = \frac{\mathbf{1}}{2}. \tag{17}$$

We can translate equation (17) into an equivalent condition on photonic states. The most general pure $N$-particle two-mode bosonic state in the Fock basis is

$$|\Psi\rangle = \sum_{M=0}^{N} c_M |M, N-M\rangle. \tag{18}$$

Suppose the state in equation (18) is in the $HV$ basis. We can map it into its spin state counterpart using equation (1):

$$|\Psi\rangle \longleftrightarrow |\xi_0\rangle_{\mathrm{spin}} = \sum_{M=0}^{N} c_M \frac{1}{\sqrt{\binom{N}{M}}} \sum_j \Pi_j(|\uparrow\rangle^{\otimes M} \otimes |\downarrow\rangle^{\otimes(N-M)}). \tag{19}$$

We identify $\rho_z^{[1]}$ with the one-particle reduced state of $|\xi_0\rangle_{\mathrm{spin}}$. Again using equation (1), we can map $\rho_z^{[1]}$ to its photonic counterpart $\rho^{[1]}$ and find

$$\rho^{[1]} = \frac{1}{N} \times \left[ \sum_{M=0}^{N} |c_M|^2 M |1, 0\rangle\langle 1, 0| + \sum_{M=0}^{N} |c_M|^2 (N-M) |0, 1\rangle\langle 0, 1| \right.$$
$$\left. + \sum_{M=0}^{N-1} c_M^* c_{M+1} \sqrt{(N-M)(M+1)} |1, 0\rangle\langle 0, 1| + \sum_{M=0}^{N-1} c_M c_{M+1}^* \sqrt{(N-M)(M+1)} |0, 1\rangle\langle 1, 0| \right]. \tag{20}$$

Since $\langle\Psi|a^\dagger a|\Psi\rangle = \sum_{M=0}^{N}|c_M|^2 M = N - \langle\Psi|b^\dagger b|\Psi\rangle$ and $\langle\Psi|ab^\dagger|\Psi\rangle = \sum_{M=0}^{N-1} c_M^* c_{M+1}\sqrt{(N-M)(M+1)}$, equation (20) can be simplified to

$$\rho^{[1]} = \frac{1}{N}(\langle\Psi|a^\dagger a|\Psi\rangle|1,0\rangle\langle 1,0| + \langle\Psi|b^\dagger b|\Psi\rangle|0,1\rangle\langle 0,1|)$$
$$+ \frac{1}{N}(\langle\Psi|ab^\dagger|\Psi\rangle|1,0\rangle\langle 0,1| + \langle\Psi|a^\dagger b|\Psi\rangle|0,1\rangle\langle 1,0|). \tag{21}$$

This means the photonic equivalent to equation (17) is

$$\rho^{[1]} = \frac{1}{2}|0,1\rangle\langle 0,1| + \frac{1}{2}|1,0\rangle\langle 1,0|. \tag{22}$$

Thus, the conditions for saturating the quantum Cramér-Rao inequality (i.e. maximally-mixed one-particle density matrix) rewritten in the photonic form are

$$\langle\Psi|a^\dagger a|\Psi\rangle = \langle\Psi|b^\dagger b|\Psi\rangle = \frac{N}{2},$$
$$\langle\Psi|a^\dagger b|\Psi\rangle = 0 = \langle\Psi|ab^\dagger|\Psi\rangle. \tag{23}$$

The condition $\langle\Psi|a^\dagger b|\Psi\rangle = 0$ corresponds to an absence of first-order coherence for the state [43]. An interpretation for an arbitrary state satisfying all the conditions of equation (23) is that the intensities in each mode remain unaltered by any two-mode interferometer.

The conditions in equation (23) show that NOON states (except $N = 1$) and Holland-Burnett states all saturate the quantum Cramér-Rao inequality. In fact, all states of the form $|M, N-M\rangle + |N-M, M\rangle$ saturate the quantum Cramér-Rao inequality, *except* when $M = (N \pm 1)/2$ (Yurke states for odd $N$ [36]). The Yurke states for even $N$ are $|\Psi\rangle = (1/\sqrt{2})(|N/2, N/2\rangle + |N/2 + 1, N/2 - 1\rangle)$ [37], which also do not saturate the quantum Cramér-Rao inequality.

The Fock states $|N, 0\rangle$ and $|0, N\rangle$ clearly violate the first equality in equation (23) and thus also do not saturate the quantum Cramér-Rao inequality. Using equation (13) for these states, we find $\mathrm{tr}(I^{-1}) = 3/(4N)$, which is shot-noise limited precision and agrees with known results in single-parameter estimation.

### 3.3. Optimal limit for the quantum Cramér-Rao inequality

To obtain the optimal precision in our protocol, we must find the lowest attainable value of $\mathrm{tr}(I^{-1})$. For this, we need first an argument from [24] for the general form of $I$ for optimal states. We begin with the Cauchy-Schwarz inequality, which implies $9 = \mathrm{tr}(\mathbf{1})^2 = [\mathrm{tr}(I^{-\frac{1}{2}}I^{\frac{1}{2}})]^2 \leqslant \mathrm{tr}(I)\mathrm{tr}(I^{-1})$. The minimum value of $\mathrm{tr}(I^{-1})|_{\min}$ occurs when the Cauchy-Schwarz inequality is saturated, or

$$\mathrm{tr}(I^{-1}) = \mathrm{tr}(I^{-1})|_{\min} = \frac{9}{\mathrm{tr}(I)}. \tag{24}$$

This requires $I_{\alpha\beta} \propto \delta_{\alpha\beta}$. Next we note that the saturation condition $\rho_z^{[1]} = \mathbf{1}/2$ must also be satisfied for optimal $\rho_0$, and the general form of $I_{\alpha\beta}(\rho_0)$ from equation (13) then becomes

$$I_{\alpha\beta}(\rho_0) = 2N[3\delta_{\alpha\beta} + (N-1)\mathrm{tr}[\rho_{\mathrm{tot}}^{[2]}(\sigma_\alpha \otimes \sigma_\beta)]]. \tag{25}$$

Inserting condition $I_{\alpha\beta} \propto \delta_{\alpha\beta}$ into equation (25) yields

$$\mathrm{tr}[\rho_{\mathrm{tot}}^{[2]}(\sigma_\alpha \otimes \sigma_\beta)] \propto \delta_{\alpha\beta}. \tag{26}$$

We identify the class of states $\rho_{\mathrm{tot}}^{[2]}$ satisfying equation (26) as follows. For the general form of $\rho_{\mathrm{tot}}^{[2]}$ we write $\rho_{\mathrm{tot}}^{[2]} = \rho_z^{[2]} + (h \otimes h)\rho_z^{[2]}(h \otimes h) + (h_c \otimes h_c)\rho_z^{[2]}(h_c^\dagger \otimes h_c^\dagger)$. We can expand $\rho_z^{[2]}$ in the basis of Pauli operators as $\rho_z^{[2]} = (\mathbf{1} \otimes \mathbf{1})/4 + \sum_{j,k=x,y,z} c_{jk}(\sigma_j \otimes \sigma_k)$, where $c_{jk} = c_{kj}$ by the symmetry of $\rho_z^{[2]}$. We also note that the requirement for $\rho_z^{[1]} = \mathbf{1}/2$ eliminates contributions from terms in $\sigma_j \otimes \mathbf{1}$ and $\mathbf{1} \otimes \sigma_k$.

The operators $h \otimes h$ and $h_c \otimes h_c$ act to permute off-diagonal terms of $\rho_z^{[2]}$ ($j \neq k$ or change sign for $c_{jk}$). From equation (26) it can be readily verified that since the off-diagonal contributions to $\rho_{\mathrm{tot}}^{[2]}$ must be zero, the off-diagonal contributions to $\rho_z^{[2]}$ must also disappear, i.e., $c_{jk} = 0$ for $j \neq k$. From equation (26) it also follows that $c_{xx} = c_{yy}$. Thus for the diagonal terms of $\rho_{\mathrm{tot}}^{[2]}$ we find

$$\rho_{\mathrm{tot}}^{[2]} = (3/4)\mathbf{1} \otimes \mathbf{1} + K \sum_{k=x,y,z}(\sigma_k \otimes \sigma_k), \tag{27}$$

where $K = (2c_{xx} + c_{zz})$. Inserting equation (27) into equation (25), we find $I_{\alpha\beta} = 2N(3 + 4K(N-1))\delta_{\alpha\beta}$. From the optimality condition in equation (24) this means optimal states should be those that maximise $K$. Since the eigenvalues of $\rho_{\mathrm{tot}}^{[2]}$ are $3/4 + K$ and $3/4 - 3K$, the maximum value of $K$ for maintaining physical states (i.e., eigenvalues range from 0 to 3) is $K = 1/4$. Optimal states are therefore those satisfying

$$\rho_z^{[2]} = \frac{\mathbf{1} \otimes \mathbf{1}}{4} + \frac{1}{2}\left(\frac{1}{4} - c_{zz}\right)(\sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y) + c_{zz}\sigma_z \otimes \sigma_z. \tag{28}$$

For these optimal states, we have from equation (25) that $I_{\alpha\beta} = 2N(N+2)\delta_{\alpha\beta}$, which implies

$$\mathrm{tr}(I_{\alpha\beta}^{-1})|_{\min} = \frac{3}{2N(N+2)}. \tag{29}$$

We remark that this displays Heisenberg scaling $\mathrm{tr}(I_{\alpha\beta}^{-1}) \sim \mathcal{O}(N^{-2})$, signalling a quantum advantage in multi-parameter estimation.

Following a similar derivation as in section 3.2, we can convert equation (28) to its photonic counterpart using equation (1)

$$\rho^{[2]} = \left(\frac{1}{4} + c_{zz}\right)(|2, 0\rangle\langle 2, 0| + |0, 2\rangle\langle 0, 2|) + 2\left(\frac{1}{4} - c_{zz}\right)|1, 1\rangle\langle 1, 1|. \tag{30}$$

For the general photonic state $|\Psi\rangle$ in equation (18), the equivalent photonic state to $\rho_z^{[2]}$ is

$$\begin{aligned}
\rho^{[2]} = \frac{1}{N(N-1)} \times \big[ &\langle\Psi|(a^\dagger a)(a^\dagger a - 1)|\Psi\rangle |2, 0\rangle\langle 2, 0| \\
&+ \langle\Psi|(b^\dagger b)(b^\dagger b - 1)|\Psi\rangle |0, 2\rangle\langle 0, 2| + 2\langle\Psi|(a^\dagger a)(b^\dagger b)|\Psi\rangle |1, 1\rangle\langle 1, 1| \\
&+ \langle\Psi|a^{\dagger 2} b^2|\Psi\rangle |0, 2\rangle\langle 2, 0| + \langle\Psi|(a^{\dagger 2} b^2)^\dagger|\Psi\rangle |2, 0\rangle\langle 0, 2| \\
&+ \sqrt{2}\,\langle\Psi|(a^\dagger b)(b^\dagger b - 1)|\Psi\rangle |0, 2\rangle\langle 1, 1| + \sqrt{2}\,\langle\Psi|((a^\dagger b)(b^\dagger b - 1))^\dagger|\Psi\rangle |1, 1\rangle\langle 0, 2| \\
&+ \sqrt{2}\,\langle\Psi|(a^\dagger b)(a^\dagger a)|\Psi\rangle |1, 1\rangle\langle 2, 0| + \sqrt{2}\,\langle\Psi|((a^\dagger b)(a^\dagger a))^\dagger|\Psi\rangle |2, 0\rangle\langle 1, 1| \big]. \tag{31}
\end{aligned}$$

Comparing equations (30) and (31), the necessary and sufficient conditions for a state to be optimal (i.e., achieving the optimal quantum Cramér-Rao inequality) can be written succinctly as

$$\begin{aligned}
\langle\Psi|a^\dagger a|\Psi\rangle &= \langle\Psi|b^\dagger b|\Psi\rangle = \frac{N}{2}, \\
\langle\Psi|a^\dagger a a^\dagger a|\Psi\rangle &= \langle\Psi|b^\dagger b b^\dagger b|\Psi\rangle, \\
\langle\Psi|a^\dagger b|\Psi\rangle &= 0 = \langle\Psi|a^\dagger a^\dagger a b|\Psi\rangle = \langle\Psi|a^\dagger a^\dagger b b|\Psi\rangle = \langle\Psi|a^\dagger b^\dagger b b|\Psi\rangle. \tag{32}
\end{aligned}$$

We observe that while the saturation of the Cramér-Rao inequality depends only on the first-order correlations, the optimality conditions also depend on the second-order correlations $\langle\Psi|a^\dagger a^\dagger ab|\Psi\rangle$, $\langle\Psi|a^\dagger a^\dagger bb|\Psi\rangle$ and $\langle\Psi|a^\dagger b^\dagger bb|\Psi\rangle$. This is related to the observation that in the corresponding spin system, the optimality condition is also a contraint on the two-particle reduced state, not only the one-particle reduced state.

Using equation (32), it can easily be shown that optimal states include all Holland-Burnett states and states of the form $|M, N - M\rangle + |N - M, M\rangle$ (including NOON states) *except* when $M = (N \pm 1)/2$ (e.g. Yurke states, which do not saturate the quantum Cramér-Rao inequality) *and* when $M = N/2 \pm 1$ (e.g. NOON state when $N = 2$ is sub-optimal).

### 3.4. Comparison to alternative protocols

To compare the performance of our protocol with related proposals in the literature, we now consider the general scenario in which an arbitrary probe state $\rho_0$ with $N$ particles can be repeatedly prepared, and used to obtain an estimate to an unknown unitary $U_s$. For this we adapt the arguments in sections 3.2 and 3.3 to compute the optimal precision achievable using different choices for $\rho_0$, which includes the possibility of a collective measurement on all $N$ particles of the state $U_S^{\otimes N} \rho_0 U_S^{\dagger \otimes N}$ for each run of the protocol. We note that that the ability to prepare arbitrary $\rho_0$ here is comparable to the ability in our protocol to prepare each of optimal $\rho_x$, $\rho_y$ and $\rho_z$ with $N$ particles ($\rho_x$, $\rho_y$ and $\rho_z$ are used for a third of the total measurements each). The spin versions of optimal $\rho_x$, $\rho_y$ and $\rho_z$ we have discussed (e.g., spin analogues of Holland-Burnett and NOON states under local rotations) are all entangled. This is similar to the situation in single-parameter estimation using entangled states of spins or qubits, for which it is well known that entanglement in the probe state plays a critical role in achieving supersensitivity [44].

To consider the case of arbitrary probe states for unitary estimation, we must now modify our earlier definitions of the one and two-particle reduced states to account for states $\rho_0$ which are not symmetric. We define the 'averaged' one-particle reduced state of $\rho_0$ as $\tilde{\rho}_0^{[1]} = (1/N)\sum_{i=1}^N \mathrm{tr}_{[i]}(\rho_0)$ and the 'averaged' two-particle reduced state of $\rho_0$ as $\tilde{\rho}_0^{[2]} = [1/(N(N-1))] \sum_{i>j}^N [\mathrm{tr}_{[i,j]}(\rho_0) + \mathrm{tr}_{[j,i]}(\rho_0)]$. Saturation of the quantum Cramér-Rao bound in equation (15) then reduces to $\mathrm{tr}(\rho_0[\tau_\alpha, \tau_\beta]) = (N/2)\mathrm{tr}(\tilde{\rho}_0^{[1]}[\sigma_\alpha, \sigma_\beta]) = 0$. By the same argument in section 3.2, this implies that $\tilde{\rho}_0^{[1]} = \mathbf{1}/2$. Now the quantum Fisher information matrix for a general $N$-particle probe state satisfying this condition takes the form $I_{\alpha\beta}(\rho_0) = 2N[\delta_{\alpha\beta} + (N-1)]\mathrm{tr}[\tilde{\rho}_0^{[2]}(\sigma_\alpha \otimes \sigma_\beta)]$, using equation (11). Optimality requires $I_{\alpha\beta} \propto \delta_{\alpha\beta}$ and this implies the constraint $\mathrm{tr}[\tilde{\rho}_0^{[2]}(\sigma_\alpha \otimes \sigma_\beta)] \propto \delta_{\alpha\beta}$, similar to equation (26). Since $\tilde{\rho}_0^{[1]} = \mathbf{1}/2$ and $\tilde{\rho}_0^{[2]}$ is symmetrised by definition, the optimality constraint restricts

$$\tilde{\rho}_0^{[2]} = \frac{\mathbf{1} \otimes \mathbf{1}}{4} + K \sum_{\alpha=x,y,z} \sigma_\alpha \otimes \sigma_\alpha. \tag{33}$$

Therefore $I_{\alpha\beta} = 2N\left[1 + 4K(N-1)\right]\delta_{\alpha\beta}$. Maximum precision comes from maximising $K$ for physical states (i.e. the eigenvalues for $\tilde{\rho}_0^{[2]}$ lie between 0 and 1). Since the eigenvalues of $\tilde{\rho}_0^{[2]}$ are $1/4 + K$ and $1/4 - 3K$, the maximum is at $K = 1/12$. Thus the optimal precision is $\mathrm{tr}(I_{\alpha\beta}^{-1})|_{\min} = 9/[2N(N+2)]$. Note that if we used a separable state of $3N$ particles instead of a single $N$-particle state, this reduces precision to $\mathrm{tr}(I_{\alpha\beta}^{-1})|_{\min} = 3/[2N(N+2)]$. It is also clear from the dependence of $I_{\alpha\beta}(\rho_0)$ on the symmetrised quantities $\tilde{\rho}_0^{[1]}$ and $\tilde{\rho}_0^{[2]}$ that there is no advantage to choices for $\rho_0$ which are not symmetric. Since the symmetric state probes are those that can correspond to photonic states, this implies that spin protocols do not have an advantage over photonic protocols in terms of achieving optimal precision.

We can also compare one round of our protocol with the use of a single entangled probe state $\rho_0$ of $3N$ spin particles. We can use the same argument in section 3.2 to derive the optimal precision in the latter case. Now the quantum Fisher information matrix takes the form $I_{\alpha\beta}(\rho_0) = 2N\left[3\delta_{\alpha\beta} + 3(3N-1)\mathrm{tr}[\rho_0^{[2]}(\sigma_\alpha \otimes \sigma_\beta)]\right]$, using equation (11), which contains extra contributions to the $\mathrm{tr}[\rho_0^{[2]}(\sigma_\alpha \otimes \sigma_\beta)]$ term compared to equation (25). These come from correlations in $\rho_0$ that do not exist when $\rho_0 = \rho_x \otimes \rho_y \otimes \rho_z$. The precision thus achievable would be $\mathrm{tr}(I_{\alpha\beta}^{-1})|_{\min} = 3/[2N(3N+2)]$. This represents better precision but requires additional preparation resources. On the other hand, when the restriction is made that probe states can only be prepared with correlations on up to $N$ spin particles, the optimal precision using our protocol versus repeated use of a single type of probe state is the same, with $\mathrm{tr}(I_{\alpha\beta}^{-1})|_{\min} = 3/[2N(N+2)]$ when $3N$ spin particles are used. Note that whatever protocol is used, the Fisher information analysis here can reliably predict precision only when large number of measurements are used (roughly 100's of measurements would typically be used for each estimate of $U_s$). In this large number limit, the precision scaling for both are similar.

Now we turn to optimal probe states, assuming that they can only be prepared with correlations on up to $N$ spin particles. We have shown for our protocol that any probe state whose two-particle reduced state in the $z$-basis obeys equation (28) can achieve optimal precision. However, if an alternative protocol is used with only one symmetric input state, as proposed in the recent protocol in [25], the stricter equation (33) must be obeyed. An example of such a state given in [25] is $\sum_{i=x,y,z}(|+\sigma_i\rangle^{\otimes N} + |-\sigma_i\rangle^{\otimes N})$ where $|\pm\sigma_i\rangle$ are the $\pm$ eigenstates of the Pauli matrics, in the large $N$ limit or when $N$ is a multiple of 8. For the analogous photonic protocol, these probe states are superpositions of NOON states with respect to the polarisation bases HV, DA and RL, respectively.

Another way of satisfying equation (33) is given in [24] and uses protocols which allow for additional correlations between the probe state $\rho_0$ with $N$ spin particles and an ancilla state $|i\rangle_{\mathrm{anc}}$ (which does not interact with the unknown unitary). The optimal precision $\mathrm{tr}(I_{\alpha\beta}^{-1})|_{\min} = 3/[2N(N+2)]$ is the same as above with $N$-spin probe states. In [24], it is argued that all states proportional to $\sum_{i=0}^{B-1} |s_i\rangle^{\otimes N} \otimes |i\rangle_{\mathrm{anc}}$ satisfy equation (33) whenever the states $\{|s_0\rangle, \cdots, |s_{B-1}\rangle\}$ for the probe satisfies $(1/B)\sum_{i=0}^{B-1}|s_i\rangle^{\otimes 2}\langle s_i|^{\otimes 2} = (\mathbf{1} \otimes \mathbf{1})/4 + (1/12)\sum_{\alpha=x,y,z}\sigma_\alpha \otimes \sigma_\alpha$ (also known as a 2-design) and the states $\{|0\rangle_{\mathrm{anc}}, \cdots, |B-1\rangle_{\mathrm{anc}}\}$ for the ancilla are orthonormal. This is true for the Pauli basis where $|s_0\rangle = |\sigma_x\rangle$, $|s_1\rangle = |-\sigma_x\rangle$, $|s_2\rangle = |\sigma_y\rangle$, $|s_3\rangle = |-\sigma_y\rangle$, $|s_4\rangle = |\sigma_z\rangle$ and $|s_5\rangle = |-\sigma_z\rangle$. Another example is given by the tetrahedral basis, such as $|s_0\rangle = |0\rangle$, $|s_1\rangle = (1/\sqrt{3})|0\rangle - (1/\sqrt{2} + i/\sqrt{6})|1\rangle$, $|s_2\rangle = (1/\sqrt{3})|0\rangle + (1/\sqrt{2} - i/\sqrt{6})|1\rangle$, and $|s_3\rangle = (1/\sqrt{3})|0\rangle + i\sqrt{2/3}\,|1\rangle$. In addition, [24] presents generalisations to higher-spin particles. Our photon-spin mapping defined in section 2.1 also generalises, and the results in [24] for higher-spin particles can therefore be directly translated to systems which use photonic states to probe linear-optical unitaries on $>2$ modes.

Finally we note that for any optimal choices for $\rho_x$, $\rho_y$ and $\rho_z$ in our protocol, which must satisfy equation (28), the state $(1/3)(\rho_x \otimes |0\rangle\langle 0|_{\mathrm{anc}} + \rho_y \otimes |1\rangle\langle 1|_{\mathrm{anc}} + \rho_z \otimes |2\rangle\langle 2|_{\mathrm{anc}})$ for the ancilla-assisted protocol automatically satisfies equation (33) and achieves the same precision. In particular this shows that collective measurements extending over $3N$ particles in our protocol (i.e. on $U_S^{\otimes N}\rho_x U_S^{\dagger\otimes N}$, $U_S^{\otimes N}\rho_y U_S^{\dagger\otimes N}$ and $U_S^{\otimes N}\rho_z U_S^{\dagger\otimes N}$ together) cannot further improve the achievable precision.

## 4. Linear-optical protocols using photon counting

Our previous analysis identifies the input states necessary for optimality. However, it does not indicate which projective measurements can be used to achieve this optimality. Here we focus on measurements that can be implemented in linear-optical experiments. In single-parameter estimation, Holland-Burnett and NOON states display Heisenberg scaling using photon-number-counting measurements. In this section, we find the precision achieved by general product states (which include Holland-Burnett states) and $N = 2, 3$ NOON states using photon-number-counting measurements.

### 4.1. Fisher information for photon-number-counting measurements

Consider the situation wherein the photonic state $|\Psi\rangle = \sum_{M=0}^{N} c_M|M, N - M\rangle$ in the Fock basis is used to probe an unknown linear-optical unitary U, with photon-number-counting measurements on each mode at the output. The probability of detecting the final state in the number state $|M_d, N - M_d\rangle$ is $|\langle M_d, N - M_d|U|\Psi\rangle|^2$, where $M_d$ is an integer $0 \leqslant M_d \leqslant N$. These probability distributions for $M_d = 0, \cdots, N$ uniquely determine the Fisher information matrix.

Explicit forms for the probability distributions can be obtained using the Schwinger representation. This representation identifies $J_x = (a^\dagger b + ab^\dagger)/2$, $J_y = (a^\dagger b - ab^\dagger)/(2i)$ and $J_z = (a^\dagger a - b^\dagger b)/2$, where $J_x, J_y, J_z$ are the angular momentum operators acting on spin-$j$ states along the $x$, $y$, $z$ bases. The Schwinger representation therefore maps the photonic state $|M, N - M\rangle$ to a spin-$j$ state with quantum number $m$, represented by[6] $|j, m\rangle$, where $j = N/2$ and $m = M - N/2$. For convenience, we write U in terms of the Euler angle decomposition $U = \exp(i\psi_1 J_z)\exp(i\psi_2 J_y)\exp(i\psi_3 J_z)$, where $\psi_1$, $\psi_2$ and $\psi_3$ are Euler angles. Thus, given a known output state $|j, m_d\rangle$, where $m_d = M_d - N/2$, the probability distribution $|\langle M_d, N - M_d|U|\Psi\rangle|^2$ can be expressed as a function of the three Euler angles. It becomes $P(m_d, \psi_1, \psi_2, \psi_3) = |\sum_{m=-j}^{j} c_m e^{i(\psi_3 m + \psi_1 m_d)} d_{m_d,m}^j(\psi_2)|^2$. Here $d_{m_d,m}^j(\psi_2) \equiv \langle j, m_d|\exp(i\psi_2 J_y)|j, m\rangle$ are the Wigner d-matrices (for a derivation of these see [45], noting a different convention).

In our protocol, there are three types of input states and measurements corresponding to different polarisation bases *HV, DA* and *RL*. The corresponding Fisher information matrices are denoted $F^{HV}$, $F^{DA}$ and $F^{RL}$, respectively. Due to the additivity condition in [33], the total Fisher information matrix is a sum of all three contributions $F^{HV}, F^{DA}$ and $F^{RL}$.

We can relate $F^{DA}$ and $F^{RL}$ to $F^{HV}$ by observing how the probability distribution with respect to *DA* (denoted $P_{DA}$) and *RL* (denoted $P_{RL}$) can be transformed into the probability distribution with respect to *HV* (denoted $P_{HV}$) by a change in Euler angles. Suppose $\{\psi_1, \psi_2, \psi_3\}$ are the Euler angles corresponding to the transformation with respect to the *HV* basis. Then we can write $P_{DA}(m_d, \psi_1, \psi_2, \psi_3) \equiv |\langle j, m_d|_{DA} U|\Psi\rangle_{DA}|^2 = |\langle j, m_d|_{HV} U'|\Psi\rangle_{HV}|^2 \equiv P_{HV}(m_d, \psi_1', \psi_2', \psi_3')$ where $U' = \exp(i\psi_1' J_z)\exp(i\psi_2' J_y)\exp(i\psi_3' J_z)$. Thus the Euler angles $\{\psi_1', \psi_2', \psi_3'\}$ are defined by $P_{HV}(m_d, \psi_1', \psi_2', \psi_3') = P_{DA}(m_d, \psi_1, \psi_2, \psi_3)$. Similarly, the Euler angles $\{\psi_1'', \psi_2'', \psi_3''\}$ are defined by $P_{HV}(m_d, \psi_1'', \psi_2'', \psi_3'') = P_{RL}(m_d, \psi_1, \psi_2, \psi_3)$.

We define matrices $W'$ and $W''$ by $W'_{\alpha\beta} \equiv \partial\psi_\alpha'/\partial\psi_\beta$ and $W''_{\alpha\beta} \equiv \partial\psi_\alpha''/\partial\psi_\beta$. Therefore,

$$F^{DA}(\psi_1, \psi_2, \psi_3) = W'^T F^{HV}(\psi_1', \psi_2', \psi_3') W'$$
$$F^{RL}(\psi_1, \psi_2, \psi_3) = W''^T F^{HV}(\psi_1'', \psi_2'', \psi_3'') W''. \tag{34}$$

For practical purposes including computing probability distributions and performing estimation from data, it is convenient to use the Euler angle parameters (or any other simple parameterisation). However, we have seen that locally independent parameters are better suited to quantifying precision.

We can transform the Fisher information matrix in the Euler parameterisation $F_{Euler}$ into the locally independent parameterisation (with parameters $\{\theta_1, \theta_2, \theta_3\}$) by $F = J^T F_{Euler} J$, where matrix $J$ is defined as $J_{\alpha\beta} \equiv \partial\psi_\alpha/\partial\theta_\beta$. Since the locally independent parameters are defined by $t_\alpha \equiv \sigma_m/\sqrt{2}$, we can find $J$ using

$$\frac{\sigma_\alpha}{\sqrt{2}} = t_\alpha = iU_s^\dagger \sum_k \frac{\partial U_s}{\partial\psi_k} J_{k\alpha}. \tag{35}$$

The last term in equation (35) can be computed by replacing $U_s$ with its matrix representation $\mathcal{M}$ (see appendix A), expressed in its Euler angle decomposition as $\mathcal{M} = \exp(i\psi_1\sigma_z/2)\exp(i\psi_2\sigma_y/2)\exp(i\psi_3\sigma_z/2)$. Since the optimal precision is characterised by $\mathrm{tr}(F^{-1})$, we find $\mathrm{tr}(F^{-1}) = \mathrm{tr}(VF_{Euler}^{-1})$, where

$$V = (J^{-1})^T J^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 0 & \cos(\psi_2) \\ 0 & 1 & 0 \\ \cos(\psi_2) & 0 & 1 \end{pmatrix}. \tag{36}$$

$V$ becomes a simple rescaling factor around $\psi_2 = \pm\pi/2$. Thus we expect the Euler parameters to behave similarly to locally independent parameters around $\{\psi_1, \psi_2 = \pm\pi/2, \psi_3\}$. Suppose we denote infinitesimal deviation from these Euler angles by $\{\delta\psi_1, \delta\psi_2, \delta\psi_3\}$. Then $U_s(\psi_1 + \delta\psi_1, \pm\pi/2 + \delta\psi_2, \psi_3 + \delta\psi_3) \simeq U_s(\psi_1, \psi_2 = \pm\pi/2, \psi_3)R_z^\dagger(1 \pm i\delta\psi_1\sigma_x/2 + i\delta\psi_2\sigma_y/2 + i\delta\psi_3\sigma_z/2)R_z$ where $R_z = \exp(i\psi_3\sigma_z/2)$ is a rotation acting on the generators. Since the sign changes and fixed rotation $R_z$ preserve the essential features of the locally independent parameterisation, this confirms that Euler parameterisation is equivalent to locally independent parameterisation up to a scaling factor.

We also note that $\det(V) = \sin(\psi_2)^2/8$, which goes to zero at points $\psi_2 = 0$ and $\psi_2 = \pm\pi$. Thus, at these points, the inverse Fisher information with respect to the Euler parameterisation diverges. This can happen when one cannot gather any information about any one (or more) parameter(s). For instance, when $\psi_2 = 0$,

---

[6] See [45] for a derivation.

$U_s = \exp(i(\psi_1 + \psi_3)\sigma_z/2)$. Here one is estimating only two parameters $\psi_1 + \psi_3$ and $\psi_2 = 0$ instead of three parameters. Similarly when $\psi = \pm\pi$, one is estimating two parameters $\psi_3 - \psi_1$ and $\psi_2 = \pm\pi$ instead of three parameters, since $U_s = \exp(\pm i(\pi/2)\sigma_y)\exp(i(\psi_3 - \psi_1)\sigma_z/2)$. However, this singularity is an artificial product of the Euler parameterisation and not essential to the estimation protocol. Other parameterisations will lead to singularities at different points. To distinguish these artificial singularities from genuine singularities that might arise from actual limitations of the estimation scheme, we can use alternative choices of parameterisation.

The tools in this section are sufficient to find the optimal precision possible in our protocol when using photon-number-counting measurements and $N$-photon probe states.

## 4.2. Product states

We now examine the precision achievable by product states $|M, N - M\rangle$ using photon-number-counting measurements. In single-parameter estimation, the Fisher information is $F_{single} = N + 2M(N - M)$, corresponding to $U_s = \exp(i\psi_2 J_y)$ for fixed $\psi_1 = 0 = \psi_3$ [10]. This is maximised for Holland-Burnett states (i.e., $M = N/2$), where $F_{single} = N(N + 2)/2$. For multi-parameter estimation using product states, the Fisher information matrix $F^{HV}(\psi_1, \psi_2, \psi_3)$ has only one non-zero element, $F_{22}^{HV}(\psi_2)$, since the probability distributions using product states depend on $\psi_2$, but not $\psi_1$ or $\psi_3$. Hence, the scaling of $\mathrm{tr}(F^{-1})$ can be shown, using equation (34), to be equivalent to that of single-parameter estimation, where

$$\mathrm{tr}(F^{-1}) = \frac{1}{F_{single}}\mathrm{tr}(V\tilde{W}^{-1}). \tag{37}$$

Here the matrix $\tilde{W}$ is defined by $\tilde{W}_{\alpha\beta} = \delta_{2\alpha}\delta_{2\beta} + W'_{2\alpha}W'_{2\beta} + W''_{2\alpha}W''_{2\beta}$. It captures the effect of the three sets of polarisation bases used and is independent of the input state. The matrix $\tilde{W}$ can be easily computed using $\cos^2(\psi'_2/2) = \cos^2(\psi_2/2)\cos^2((\psi_1 + \psi_3)/2) + \sin^2(\psi_2/2)\sin^2((\psi_1 - \psi_3)/2)$ and $\cos^2(\psi''_2/2) = \cos^2(\psi_2/2)\cos^2((\psi_1 + \psi_3)/2) + \sin^2(\psi_2/2)\cos^2((\psi_1 - \psi_3)/2)$.

Our numerical simulations show the minimum value of $\mathrm{tr}(V\tilde{W}^{-1})$ to be $3/2$. This indicates that Holland-Burnett states are the only product states under photon-number-counting measurements that achieve an identical scaling in $N$ with the theoretical optimal in equation (29), with the minimum $\mathrm{tr}(F^{-1}) = 3/(N(N + 2))$. However, this is a factor of two higher than the predicted value for optimal measurements. It is interesting to compare to single-parameter estimation, where the optimal precision for Holland-Burnett states do not in fact share exactly the same scaling in $N$ as the theoretical optimal scaling of $1/N^2$.
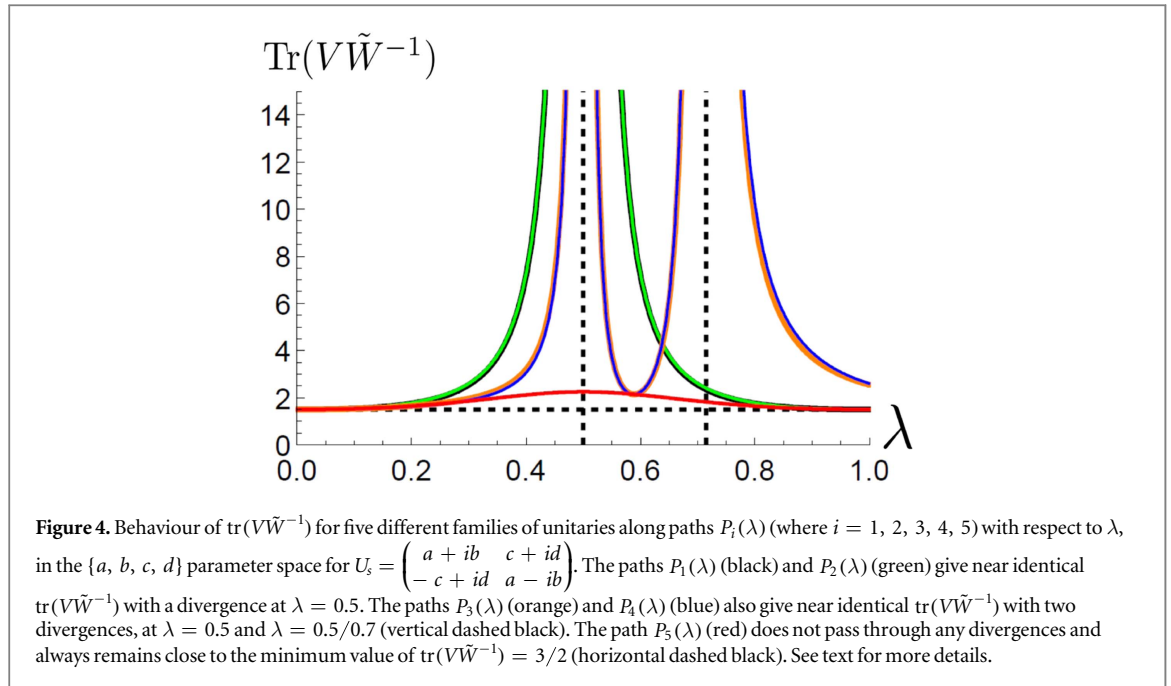
Holland-Burnett states are thus strong candidates for practical implementation of our protocol. As well as providing near-optimal precision using photon-number-counting measurements, they are more experimentally accessible than NOON states for $N > 2$ [46] and also perform better in the presence of photon loss.

We also observe that the precision in multi-parameter estimation $\mathrm{tr}(F^{-1})$ is *dependent* on the parameters of the unknown unitary. This is different to single-parameter estimation for Holland-Burnett states, where the Fisher information is independent of the unknown parameter. We can see this dependence by first noting the positions of the minima, which occur at $\{\psi_1, \psi_2, \psi_3\} = \{0, \pm\pi/2, \pm\pi/2\}$, $\{\pm\pi/2, \pm\pi/2, 0\}$, $\{\pm\pi, \pm\pi/2, \pi/2\}$, $\{\pm3\pi/2, \pm\pi/2, 0\}$. These minima occur at $\psi_2 = \pm\pi/2$, where the Euclidean parameterisation coincides with the locally independent parameterisation. If we take $U_s = \begin{pmatrix} a + ib & c + id \\ -c + id & a - ib \end{pmatrix}$, where $a, b, c, d \in \mathfrak{R}$ and $a^2 + b^2 + c^2 + d^2 = 1$, then the minima occur when $|a| = |b| = |c| = |d| = 1/2$. This coincides with the results found in [11] that use process fidelity instead of Fisher information.

We can examine the dependence of $\mathrm{tr}(F^{-1})$ on $U_s = \begin{pmatrix} a + ib & c + id \\ -c + id & a - ib \end{pmatrix}$ by showing how $\mathrm{tr}(V\tilde{W}^{-1})$ changes with respect to unitaries that are near the minimal unitary $u_{min} = \{a, b, c, d\} = \{1/2, 1/2, 1/2, 1/2\}$ (i.e., the unitary where $\mathrm{tr}(F^{-1})$ attains its lowest value). Unitaries near $u_{min}$ can be accessed from $u_{min}$ through different paths in the $\{a, b, c, d\}$ parameter space. Each point along a path represents a different unitary we want to estimate and we can study the behaviour of $\mathrm{tr}(V\tilde{W}^{-1})$ at each unitary along each path. We can define such five paths $P_i(\lambda) \equiv p_i(\lambda)/\|p_i(\lambda)\|$ where $i = 1, 2, 3, 4, 5$ and each path is parameterised by $\lambda \in [0, 1]$ in the following way: $p_1 = u_{min} - \{\lambda, 0, 0, 0\}$, $p_2 = u_{min} - \{\lambda, 0, \lambda, 0\}$, $p_3 = u_{min} - \{\lambda, 0.7\lambda, \lambda, 0\}$, $p_4 = u_{min} - \{0.7\lambda, \lambda, 0.7\lambda, 0\}$ and $p_5 = u_{min} - \{\lambda, \lambda, \lambda, 0\}$. See figure 4 for a plot showing how $\mathrm{tr}(V\tilde{W}^{-1})$ changes along $p_i(\lambda)$. There are two divergences that occur at $\lambda = 0.5$ and $\lambda = 0.5/0.7$, which correspond to at least one of $a, b, c, d$ having value zero.

Dependence of the Fisher information on the unitary matrix itself is in fact common in single-parameter estimation when not dealing with so-called path-symmetric states (which include Holland-Burnett and NOON states [47]) or when the effects of experimental imperfections are accounted for. Adaptive schemes such as in [48, 49] can be used to optimise precision given these dependencies.

**Figure 4.** Behaviour of $\mathrm{tr}(V\tilde{W}^{-1})$ for five different families of unitaries along paths $P_i(\lambda)$ (where $i = 1, 2, 3, 4, 5$) with respect to $\lambda$, in the $\{a, b, c, d\}$ parameter space for $U_s = \begin{pmatrix} a + ib & c + id \\ -c + id & a - ib \end{pmatrix}$. The paths $P_1(\lambda)$ (black) and $P_2(\lambda)$ (green) give near identical $\mathrm{tr}(V\tilde{W}^{-1})$ with a divergence at $\lambda = 0.5$. The paths $P_3(\lambda)$ (orange) and $P_4(\lambda)$ (blue) also give near identical $\mathrm{tr}(V\tilde{W}^{-1})$ with two divergences, at $\lambda = 0.5$ and $\lambda = 0.5/0.7$ (vertical dashed black). The path $P_5(\lambda)$ (red) does not pass through any divergences and always remains close to the minimum value of $\mathrm{tr}(V\tilde{W}^{-1}) = 3/2$ (horizontal dashed black). See text for more details.

### 4.3. NOON states

We now examine the precision achieved by $N = 2, 3$ NOON states under photon-number-counting measurements. We restrict our attention to small-$N$ NOON states since proposed schemes for efficient generation of large-$N$ NOON states achieve fidelity considerably less than 1 [50][7], or require feedforward [51, 52] which is technically-challenging. Although all large-$N$ quantum states with high phase sensitivity are very sensitive to photon losses, the problem is particularly acute for NOON states, for which the loss of a single photon due to losses acting independently on both modes causes a complete loss of phase sensitivity.

We find $\mathrm{tr}(F^{-1})$ for $N = 2$ and $N = 3$ NOON states by directly computing $F^{HV}$, $F^{DA}$ and $F^{RL}$ from the corresponding probability distributions. Plots for $\mathrm{tr}(F^{-1})$ along the paths $P_i(\lambda)$ for the $N = 2$ NOON state are given in figure 5 and for the $N = 3$ NOON state in figure 6.

For the case of $N = 2$, we note that although a unitary $U_s = h_c$ can be used to convert the Holland-Burnett state into the NOON state, this transformation does not commute with $h$, used for the $DA$ polarisation-basis measurement. Hence, the general behaviour of the precision computed for the two states should be expected to be quite different.

Numerical investigation based on a random search over 1000 Haar-random matrices (that represent unitary transformations we want to estimate) yields a minimum $\mathrm{tr}(F^{-1}) = 0.377$ for $N = 2$ NOON states. This is very close to the $N = 2$ Holland-Burnett state with $\mathrm{tr}(F^{-1}) = 0.375$. Neither of these states are optimal which requires $\mathrm{tr}(F^{-1}) = 0.1875$.

A similar search for $N = 3$ NOON states yields a minimum $\mathrm{tr}(F^{-1}) = 0.167$. This is slightly better precision than the value 0.2, which is given for $N = 3$ by the formula $3/N(N + 2)$ (which applies for Holland-Burnett states with even $N$). Note that the optimal value $\mathrm{tr}(F^{-1}) = 0.1$ is not achieved.
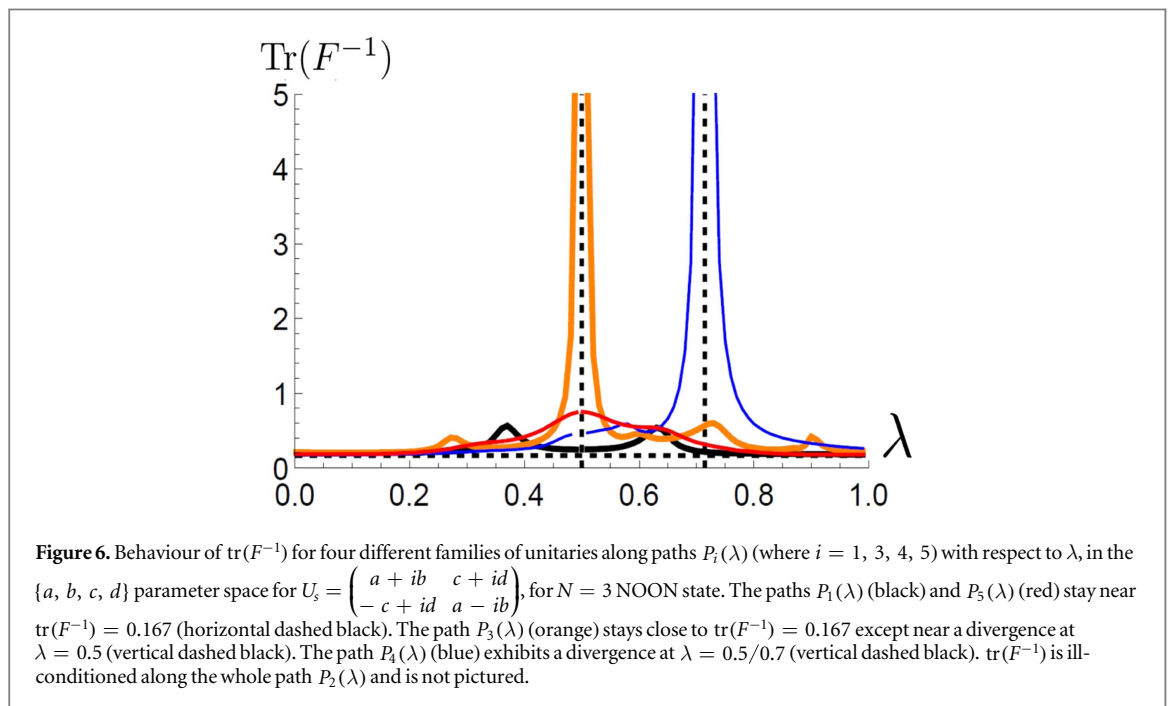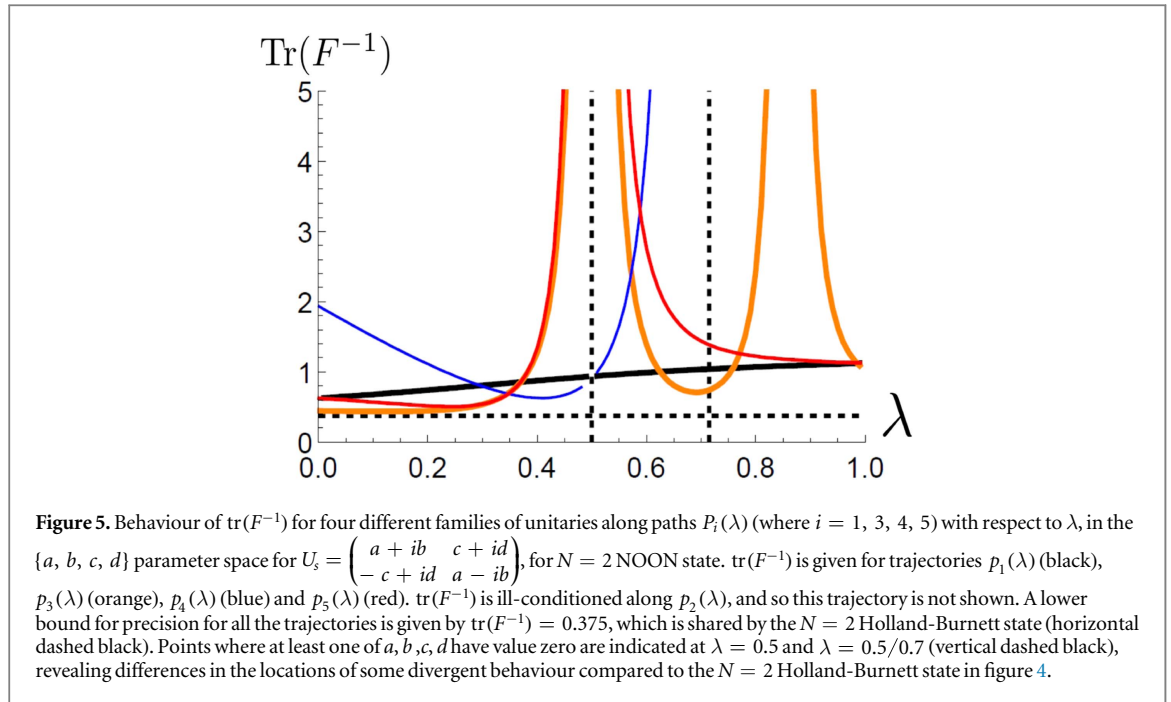
Thus the optimality of $N = 2$ and $N = 3$ NOON states together with photon-number-counting measurements for single-parameter estimation no longer holds true in the multi-parameter estimation protocol that we consider. Furthermore, the best achievable precision by these NOON states is similar to that achievable using Holland-Burnett states.

### 4.4. $N = 2, 3$ states and photon-number-counting measurements

It is interesting to consider if there exist $N = 2, 3$ states which can reach the optimal bound on precision in equation (29) using photon-number-counting measurements. For $N = 2$, the Holland-Burnett state is the only state whose precision saturates the optimal bound in equation (29) for general measurements and for $N = 3$, the NOON state is the only such optimal state (see appendix B). Due to the results in 4.2 and 4.3, this means that there is no $N = 2, 3$ state that can saturate the optimal precision using photon-number-counting measurements.

However, there may still exist states that can perform better than Holland-Burnett and NOON states under photon-number-counting measurements only. It is interesting to explore the precision attained by these states,

---

[7] We note that the fidelities achieved in the scheme in [50] are more than 90% for NOON components with high photon number.

**Figure 5.** Behaviour of $\mathrm{tr}(F^{-1})$ for four different families of unitaries along paths $P_i(\lambda)$ (where $i = 1, 3, 4, 5$) with respect to $\lambda$, in the $\{a, b, c, d\}$ parameter space for $U_s = \begin{pmatrix} a + ib & c + id \\ -c + id & a - ib \end{pmatrix}$, for $N = 2$ NOON state. $\mathrm{tr}(F^{-1})$ is given for trajectories $p_1(\lambda)$ (black), $p_3(\lambda)$ (orange), $p_4(\lambda)$ (blue) and $p_5(\lambda)$ (red). $\mathrm{tr}(F^{-1})$ is ill-conditioned along $p_2(\lambda)$, and so this trajectory is not shown. A lower bound for precision for all the trajectories is given by $\mathrm{tr}(F^{-1}) = 0.375$, which is shared by the $N = 2$ Holland-Burnett state (horizontal dashed black). Points where at least one of $a, b, c, d$ have value zero are indicated at $\lambda = 0.5$ and $\lambda = 0.5/0.7$ (vertical dashed black), revealing differences in the locations of some divergent behaviour compared to the $N = 2$ Holland-Burnett state in figure 4.



**Figure 6.** Behaviour of $\mathrm{tr}(F^{-1})$ for four different families of unitaries along paths $P_i(\lambda)$ (where $i = 1, 3, 4, 5$) with respect to $\lambda$, in the $\{a, b, c, d\}$ parameter space for $U_s = \begin{pmatrix} a + ib & c + id \\ -c + id & a - ib \end{pmatrix}$, for $N = 3$ NOON state. The paths $P_1(\lambda)$ (black) and $P_5(\lambda)$ (red) stay near $\mathrm{tr}(F^{-1}) = 0.167$ (horizontal dashed black). The path $P_3(\lambda)$ (orange) stays close to $\mathrm{tr}(F^{-1}) = 0.167$ except near a divergence at $\lambda = 0.5$ (vertical dashed black). The path $P_4(\lambda)$ (blue) exhibits a divergence at $\lambda = 0.5/0.7$ (vertical dashed black). $\mathrm{tr}(F^{-1})$ is ill-conditioned along the whole path $P_2(\lambda)$ and is not pictured.

but finding these states and simulating their precision is difficult. It requires the computation of 18 independent terms for the Fisher information matrix (each of which is dependent on multi-photon statistics), that correspond to three independent measurements. From the examples of Holland-Burnett and NOON states, we have seen that there is also a high sensitivity of the precision on the unitary and the task for optimising over random unitaries is already complicated even for any given $N = 2, 3$ state. Furthermore, there are no known methods for generating arbitrary states with fixed total photon number in linear-optical measurements (especially for $N > 2$). If a $N = 3$ state with better precision under photon-number-counting measurements were found, it may be difficult to experimentally generate. This is similar to the case in single-parameter estimation, where in general there are only limited cases when an experimentally-motivated measurement does saturate the quantum Cramér-Rao bound.

We also note that, for arbitrary $N$, the precision $\mathrm{tr}(F^{-1})$ for Holland-Burnett states under photon-number-counting measurements only differs from the optimal precision by a factor of 2. Therefore, practical considerations may make Holland-Burnett states more desirable to use over other possible states, which in

principle provide marginally better performance but which would be more difficult to generate. While Holland-Burnett and NOON states are not optimal under photon-number-counting measurements, it is an interesting question for future investigation what kinds of measurements are optimal for these states. However, it is unclear if there are any 'simple' optimal measurements, as there are few experimentally available options other than photon-number-counting measurements.

## 5. Summary and outlook

In summary, we have developed a formalism to study quantum-enhanced $SU(2)$-estimation using $N$-particle photonic states. We have derived easy-to-use, necessary and sufficient conditions that these photonic states must satisfy to achieve the optimal precision in $SU(2)$-estimation. We also interpreted these results in terms of photon interferometry. Our results showed some key differences between multi-parameter and single-parameter estimation.

We found that, unlike single-parameter estimation (without loss and decoherence), Holland-Burnett states and NOON states provide a strongly unitary-dependent precision, making adaptive measurements essential. Holland-Burnett states are optimal, although they are only near optimal using photon-number-counting measurements. This makes Holland-Burnett states experimentally preferable to high-$N$ NOON states which are difficult to generate. In addition, $N = 2$ and $N = 3$ NOON states do not achieve optimal precision under photon-number-counting measurements.

Our results clearly show how multi-parameter estimation is not a simple generalisation of single-parameter estimation and thus invites further theoretical study.

By using a mapping between photonic and spin systems, we have argued how to relate multi-parameter estimation schemes for photonic and spin systems. As a first application, we have shown how the spin analogue to our photonic protocol allows more optimal states than an earlier proposal for spin systems. The mapping we have described will potentially inform improvements for future protocols for both systems.

Extending our results to more practically important schemes presents some exciting challenges. For example, the use of multi-mode squeezed states and homodyne measurements could potentially enable quantum enhancement for unitary estimation at high-$N$. Extensions of the general approach in this paper to $SU(d)$ estimation for $d > 2$ may have implications for quantum computing models like boson sampling. An important question for future study is how the potential for quantum-enhanced precision changes for non-unitary processes that include the effects of photon loss and decoherence. This question has been a focal point for work on single-parameter estimation in recent years. In particular, the Heisenberg limit has been largely superseded by revised scaling laws that account for imperfections and which reveal much-reduced potential for supersensitivity using non-classical probe states [53].

Furthermore, there are many alternative notions of quantum enhancement compared to the one used in this paper, that should be considered for different applications with specific restrictions on physical resources. In particular, recent work on single-parameter estimation shows the importance of critically comparing the performance of single and multi-pass protocols using non-classical and classical probe states where there is a requirement to maximise precision per absorbed photon, as is key when measuring fragile samples [54].

## Acknowledgements

## Appendix A. Photonic to spin system mapping

Using the mapping in equation (1) we can describe a transformation of the two-mode photonic state under unitary operator $U$ in terms of the evolution of $N$ spin-1/2 particles. Let each spin-1/2 particle transform under the $2 \times 2$ unitary matrix represented by $\mathcal{M}$. Then $\mathcal{M}$ and $U$ are related by

$$\begin{pmatrix} \tilde{a}^\dagger \\ \tilde{b}^\dagger \end{pmatrix} \equiv \begin{pmatrix} Ua^\dagger U^\dagger \\ Ub^\dagger U^\dagger \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} a^\dagger \\ b^\dagger \end{pmatrix} \equiv \mathcal{M}^T \begin{pmatrix} a^\dagger \\ b^\dagger \end{pmatrix}, \tag{A.1}$$

where $U$ is confined to a linear-optical process. Here $\mathcal{M}$ can be interpreted as acting on a single-photon state.

We begin with the correspondence between the creation operators of the photonic and spin states $a^\dagger \leftrightarrow a_\uparrow^\dagger$, $b^\dagger \leftrightarrow a_\downarrow^\dagger$ from our mapping. Thus, after unitary evolution, we have the correspondence $Ua^\dagger U^\dagger = \tilde{a}^\dagger \leftrightarrow U_s a_\uparrow^\dagger U_s^\dagger \equiv \tilde{a}_\uparrow^\dagger$ and $Ub^\dagger U^\dagger = \tilde{b}^\dagger \leftrightarrow U_s a_\downarrow^\dagger U_s^\dagger \equiv \tilde{a}_\downarrow^\dagger$, where $U_s$ is the unitary operator acting on the spin degrees of freedom. To find the correspondence between $U_s$ and $U$, we can see from equation (A.1) that $U|1, 0\rangle = Ua^\dagger U^\dagger|0, 0\rangle = \alpha a^\dagger|0, 0\rangle + \beta b^\dagger|0, 0\rangle \leftrightarrow \alpha|\uparrow\rangle + \beta|\downarrow\rangle = U_s a_\uparrow^\dagger U_s^\dagger |0\rangle = U_s a_\uparrow^\dagger|0\rangle = U_s|\uparrow\rangle$ and $U|0, 1\rangle = Ub^\dagger U^\dagger|0, 0\rangle = \gamma a^\dagger|0, 0\rangle + \delta b^\dagger|0, 0\rangle \leftrightarrow \gamma|\uparrow\rangle + \delta|\downarrow\rangle = U_s|\downarrow\rangle$. Therefore we can write

$$\begin{pmatrix} \tilde{a}^\dagger \\ \tilde{b}^\dagger \end{pmatrix} \equiv \begin{pmatrix} Ua^\dagger U^\dagger \\ Ub^\dagger U^\dagger \end{pmatrix} \longleftrightarrow \begin{pmatrix} U_s a_\uparrow^\dagger U_s^\dagger \\ U_s a_\downarrow^\dagger U_s^\dagger \end{pmatrix} = \mathcal{M}^T \begin{pmatrix} a_\uparrow^\dagger \\ a_\downarrow^\dagger \end{pmatrix}. \tag{A.2}$$

We can find a matrix representation for $U_s$ by choosing a representation for the spin eigenstates

$$|\uparrow\rangle \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\downarrow\rangle \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{A.3}$$

Inserting this representation into the relations $U_s|\uparrow\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle$ and $U_s|\downarrow\rangle = \gamma|\uparrow\rangle + \delta|\downarrow\rangle$, we can see that a $2 \times 2$ matrix representation of $U_s$ when it acts on the spin states is equivalent to the matrix $\mathcal{M}$. We can now see the correspondence between the evolution of the photonic two-mode state and the evolution of the spin state

$$U|M, N - M\rangle \longleftrightarrow U_s^{\otimes N}|\xi_0\rangle_{\text{spin}}, \tag{A.4}$$

where a matrix representation of $U_s$ is equivalent to $\mathcal{M}$.

## Appendix B. Proof that the Holland-Burnett (NOON) state is the only optimal state when $N = 2$ ($N = 3$)

The optimality condition for photonic states can be written as a constraint on the two-particle reduced density matrix of its analogous spin state. This constraint can be expressed in the spin state form in equation (28) and in the photonic form in equation (30). For an $N = 2$ state, the two-particle reduced density matrix is simply the matrix itself. This means that a simple inspection of equation (30) reveals the Holland-Burnett state as the only pure state satisfying this constraint.

A general $N = 3$ photonic state takes the form $|\Psi\rangle = c_0|0, 3\rangle + c_1|1, 2\rangle + c_2|2, 1\rangle + c_3|3, 0\rangle$. This means that its equivalent spin state can be written as $|\xi_0\rangle_{\text{spin}} = c_0|\downarrow\downarrow\downarrow\rangle + (c_1/\sqrt{3})(|\uparrow\downarrow\downarrow\rangle + |\downarrow\uparrow\downarrow\rangle + |\downarrow\downarrow\uparrow\rangle) + (c_2/\sqrt{3})(|\uparrow\uparrow\downarrow\rangle + + |\uparrow\downarrow\uparrow\rangle + |\downarrow\uparrow\uparrow\rangle) + c_3|\uparrow\uparrow\uparrow\rangle$, which must satisfy equation (28) for optimal states. We can rewrite equation (28) in matrix form as

$$\rho_z^{[2]} = \begin{pmatrix} \frac{1}{4} + c_{zz} & 0 & 0 & 0 \\ 0 & \frac{1}{4} - c_{zz} & \frac{1}{4} - c_{zz} & 0 \\ 0 & \frac{1}{4} - c_{zz} & \frac{1}{4} - c_{zz} & 0 \\ 0 & 0 & 0 & \frac{1}{4} + c_{zz} \end{pmatrix}. \tag{B.1}$$

The two-particle reduced density matrix $\text{tr}_{[2]}(|\xi_0\rangle\langle\xi_0|_{\text{spin}})$ for $|\xi_0\rangle_{\text{spin}}$ can be shown to be

$$\text{tr}_{[2]}(|\xi_0\rangle\langle\xi_0|_{\text{spin}}) = \begin{pmatrix} |c_3|^2 + \frac{|c_2|^2}{3} & \frac{c_3 c_2^*}{\sqrt{3}} + \frac{c_1^* c_2}{3} & \frac{c_3 c_2^*}{\sqrt{3}} + \frac{c_1^* c_2}{3} & \frac{c_3 c_1^*}{\sqrt{3}} + \frac{c_2 c_0^*}{\sqrt{3}} \\ \frac{c_1 c_2^*}{3} + \frac{c_2 c_3^*}{\sqrt{3}} & \frac{|c_1|^2 + |c_2|^2}{3} & \frac{|c_1|^2 + |c_2|^2}{3} & \frac{c_1 c_0^*}{\sqrt{3}} + \frac{c_1^* c_2}{3} \\ \frac{c_1 c_2^*}{3} + \frac{c_2 c_3^*}{\sqrt{3}} & \frac{|c_1|^2 + |c_2|^2}{3} & \frac{|c_1|^2 + |c_2|^2}{3} & \frac{c_1 c_0^*}{\sqrt{3}} + \frac{c_1^* c_2}{3} \\ \frac{c_0 c_2^*}{\sqrt{3}} + \frac{c_1 c_3^*}{\sqrt{3}} & \frac{c_0 c_1^*}{\sqrt{3}} + \frac{c_1 c_2^*}{3} & \frac{c_0 c_1^*}{\sqrt{3}} + \frac{c_1 c_2^*}{3} & |c_0|^2 + \frac{|c_1|^2}{3} \end{pmatrix}. \tag{B.2}$$

It is then straightforward to show that $\text{tr}_{[2]}(|\xi_0\rangle\langle\xi_0|_{\text{spin}})$ only satisfies equation (B.1) when $c_1 = 0 = c_2$. Note that while the cases (i) $c_0 = 0 = c_2 = c_3$; (ii) $c_0 = 0 = c_1 = c_3$; (iii) $c_1 = 0 = c_2 = c_3$ in equation (B.2) obeys equation (B.1) for the zero matrix elements of $\rho_z^{[2]}$, they do not obey the diagonal terms of the matrix $\rho_z^{[2]}$ in equation (B.1). The $c_1 = 0 = c_2$ condition corresponds to the $N = 3$ NOON state $|\Psi\rangle = (1/\sqrt{2})(|3, 0\rangle + |0, 3\rangle)$.

# References

[1] Taylor M 2015 *Quantum Microscopy of Biological Systems* (Switzerland: Springer)
[2] Udem T, Holzwarth R and Hänsch T W 2002 *Nature* **416** 233–7
[3] Bloom B, Nicholson T, Williams J, Campbell S, Bishof M, Zhang X, Zhang W, Bromley S and Ye J 2014 *Nature* **506** 71–5
[4] O'Brien J L 2007 *Science* **318** 1567–70
[5] Toyli D M, Charles F, Christle D J, Dobrovitski V V and Awschalom D D 2013 *Proc. Natl Acad. Sci.* **110** 8417–21
[6] Abbott B *et al* 2016 *Phys. Rev. Lett.* **116** 061102
[7] Giovannetti V, Lloyd S and Maccone L 2011 *Nat. Photon.* **5** 222–9
[8] Nielsen M A and Chuang I L 2010 *Quantum Computation and Quantum Information* (New York: Cambridge University Press)
[9] Bollinger J, Itano W, Wineland D and Heinzen D 1996 *Phys. Rev. A* **54** R4649
[10] Durkin G A and Dowling J P 2007 *Phys. Rev. Lett.* **99** 070801
[11] Zhou X Q, Cable H, Whittaker R, Shadbolt P, O'Brien J L and Matthews J C 2015 *Optica* **2** 510–6
[12] Holland M and Burnett K 1993 *Phys. Rev. Lett.* **71** 1355
[13] Liero H and Zwanzig S 2016 *Introduction to the Theory of Statistical Inference* (Boca Raton, FL: CRC Press)
[14] Ciampini M A, Spagnolo N, Vitelli C, Pezzé L, Smerzi A and Sciarrino F 2016 *Sci. Rep.* **6** 28881
[15] Humphreys P C, Barbieri M, Datta A and Walmsley I A 2013 *Phys. Rev. Lett.* **111** 070403
[16] Yue J D, Zhang Y R and Fan H 2014 *Sci. Rep.* **4** 5933
[17] Vidrighin M D, Donati G, Genoni M G, Jin X M, Kolthammer W S, Kim M, Datta A, Barbieri M and Walmsley I A 2014 *Nat. Commun.* **5** 3532
[18] Crowley P J, Datta A, Barbieri M and Walmsley I A 2014 *Phys. Rev. A* **89** 023845
[19] Bagan E, Baig M and Munoz-Tapia R 2004 *Phys. Rev. A* **69** 050303
[20] Bagan E, Baig M and Munoz-Tapia R 2004 *Phys. Rev. A* **70** 030301
[21] Chiribella G, D'Ariano G, Perinotti P and Sacchi M 2004 *Phys. Rev. Lett.* **93** 180503
[22] Hayashi M 2006 *Phys. Lett. A* **354** 183–9
[23] Ballester M A 2004 *Phys. Rev. A* **69** 022303
[24] Ballester M A 2005 arXiv: quant-ph/0507073
[25] Baumgratz T and Datta A 2016 *Phys. Rev. Lett.* **116** 030801
[26] Fujiwara A 2001 *Phys. Rev. A* **65** 012316
[27] Matsumoto K 2002 *J. Phys. A* **35** 3111
[28] Holstein T and Primakoff H 1940 *Phys. Rev.* **58** 1098
[29] Biedenharn L C and Van Dam H 1965 *Quantum Theory of Angular Momentum: a Collection of Reprints and Original Papers* (Academic Press)
[30] Dicke R H 1954 *Phys. Rev.* **93** 99
[31] Greenberger D M, Horne M A and Zeilinger A 1989 *Going Beyond Bell's Theorem* (Amsterdam: Springer) 69–72
[32] Tóth G and Apellaniz I 2014 *J. Phys. A* **47** 424006
[33] Cramér H 1947 Mathematical methods of statistics (JSTOR)
[34] Braunstein S L and Caves C M 1994 *Phys. Rev. Lett.* **72** 3439
[35] Lee H, Kok P and Dowling J P 2002 *J. Mod. Opt.* **49** 2325–38
[36] Yurke B 1986 *Phys. Rev. Lett.* **56** 1515
[37] Yurke B, McCall S L and Klauder J R 1986 *Phys. Rev. A* **33** 4033
[38] Helstrom C 1967 *Phys. Lett. A* **25** 101–2
[39] Helstrom C 1968 *IEEE Trans. Inf. Theory* **14** 234–42
[40] Paris M G 2009 *Int. J. Quantum Inf.* **7** 125–37
[41] Ragy S, Jarzyna M and Demkowicz-Dobrzanski R 2016 *Phys. Rev. A* **94** 052108
[42] Szczykulska M, Baumgratz T and Datta A 2016 *Advances in Physics: X* **1** 621–39
[43] Loudon R 2000 *The Quantum Theory of Light* (Oxford: Oxford University Press)
[44] Giovannetti V, Lloyd S and Maccone L 2006 *Phys. Rev. Lett.* **96** 010401
[45] Sakurai J J and Napolitano J 2011 *Modern Quantum Mechanics* (Addison-Wesley)
[46] Datta A, Zhang L, Thomas-Peter N, Dorner U, Smith B J and Walmsley I A 2011 *Phys. Rev. A* **83** 063836
[47] Hofmann H F 2009 *Phys. Rev. A* **79** 033822
[48] Sugiyama T, Turner P S and Murao M 2012 *Phys. Rev. A* **85** 052107
[49] Mahler D, Rozema L, Darabi A, Ferrie C, Blume-Kohout R and Steinberg A 2013 *Phys. Rev. Lett.* **111** 183601
[50] Hofmann H F and Ono T 2007 *Phys. Rev. A* **76** 031806
[51] Cable H and Dowling J P 2007 *Phys. Rev. Lett.* **99** 163604
[52] Cable H, Laloë F and Mullin W 2011 *Phys. Rev. A* **83** 053626
[53] Escher B M, de Matos Filho R L and Davidovich L 2011 *Nature Phys.* **7** 406–11
[54] Birchall P M, O'Brien J L, Matthews J C and Cable H 2016 arXiv:1602.07561