



Heriot-Watt University
Research Gateway

Average number is an insufficient metric for interferometry

Citation for published version:

Branford, D & Rubio, J 2021, 'Average number is an insufficient metric for interferometry', *New Journal of Physics*, vol. 23, 123041. <https://doi.org/10.1088/1367-2630/ac3571>

Digital Object Identifier (DOI):

[10.1088/1367-2630/ac3571](https://doi.org/10.1088/1367-2630/ac3571)

Link:

[Link to publication record in Heriot-Watt Research Portal](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

New Journal of Physics

Publisher Rights Statement:

© 2021 The Author(s).

General rights

Copyright for the publications made accessible via Heriot-Watt Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

Heriot-Watt University has made every reasonable effort to ensure that the content in Heriot-Watt Research Portal complies with UK legislation. If you believe that the public display of this file breaches copyright please contact open.access@hw.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

PAPER • OPEN ACCESS

Average number is an insufficient metric for interferometry

To cite this article: Dominic Branford and Jesús Rubio 2021 *New J. Phys.* **23** 123041

View the [article online](#) for updates and enhancements.

You may also like

- [Peer review declaration](#)
- [Peer Review Declaration Form Guidance](#)
- [Peer review declaration](#)



PAPER

Average number is an insufficient metric for interferometry

OPEN ACCESS

RECEIVED
21 July 2021REVISED
18 October 2021ACCEPTED FOR PUBLICATION
2 November 2021PUBLISHED
22 December 2021

Original content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](#).

Any further distribution
of this work must
maintain attribution to
the author(s) and the
title of the work, journal
citation and DOI.

Dominic Branford^{1,*}  and Jesús Rubio² ¹ Institute of Photonics and Quantum Sciences, Heriot-Watt University, Edinburgh, EH14 4AS, United Kingdom² Department of Physics and Astronomy, University of Exeter, Stocker Road, Exeter, EX4 4QL, United Kingdom

* Author to whom any correspondence should be addressed.

E-mail: D.Branford@hw.ac.uk

Keywords: quantum metrology, optical interferometry, variable-number states

Abstract

We argue that analysing schemes for metrology solely in terms of the average particle number can obscure the number of particles effectively used in informative events. For a number of states we demonstrate that, in both frequentist and Bayesian frameworks, the average number of a state can essentially be decoupled from the aspects of the total number distribution associated with any metrological advantage.

1. Introduction

Quantum sensing techniques offer the possibility of improved precision per (average) particle [1]. In phase estimation, the precision $\Delta\varphi \gtrsim 1/\sqrt{\langle N \rangle}$ —the shot-noise limit (SNL)—available from repeated use of single photons or classical light can be surpassed, instead attaining $\Delta\varphi \gtrsim 1/\langle N \rangle$ —Heisenberg scaling—through quantum correlations such as entanglement and squeezing [1–5].

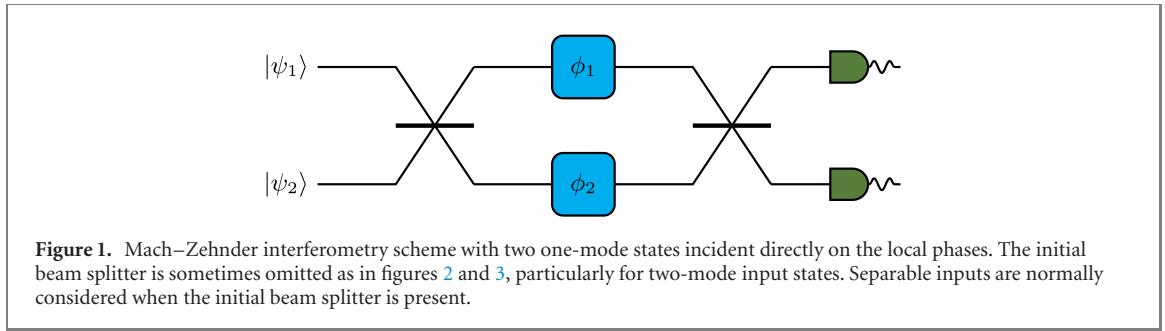
Only being a quadratic improvement, the tens of photons found in the quantum state-of-art [6, 7] cannot compare directly with the intensity of classical laser light. Indeed, only recently has the SNL been surpassed in absolute terms [8]. While latest generation gravitational wave detectors utilise squeezed light, the current frequency-independent squeezing does not decrease quantum noise in absolute terms [9], but can benefit certain gravitational wave frequencies, or reduce the required classical light and associated classical noise contributions.

Instead, the argument for more immediate applications of the quantum metrology toolbox have focussed on the idea of delicate samples [10–15]. While a piece of glass or a waveplate may be robust to strong laser light, biological systems may be expected to suffer photodamage [16]. Such biological limits give cause to want to constrain the amount of (damaging) light passing through the sample itself.

These constraints can entail limiting the total, average, or single-shot number of photons the sample is exposed to. While optimal states in a truncated Fock-space setting are known [2, 17], this strict setting necessitates excluding otherwise reasonable states—that can likely be generated in a more scalable fashion—including squeezed vacuum which have small but non-zero overlap with large Fock states. As a result, even when the delicate sample motivation is invoked, average number is often exclusively used as a resource.

For states without a fixed total number, the quadratic-in- $\langle N \rangle$ variance of Heisenberg scaling is not an absolute limit and a greater per-shot precision for fixed $\langle N \rangle$ has been predicted under local estimation [18, 19], although the regime where local estimation applies is not easily accessed in such settings [20–23]. A bound $\Delta\varphi \gtrsim 1/\sqrt{\langle N^2 \rangle}$ has been proposed as an alternate or complementary limit [24–26] to the $\Delta\varphi \gtrsim 1/\langle N \rangle$ Heisenberg scaling.

In this work we see that average number $\langle N \rangle$ is not only insufficient to construct an absolute bound to phase estimation but that it can bear no relevance to either the metrological performance of variable number states or the number of particles that may be passing through a sample on any given shot. This leads us to conclude that average number is not just an occasionally imperfect proxy for the cost of certain states, but a potentially dangerous misnomer as to the exposure a sample may receive.



2. Interferometry

Typically, analyses of interferometry, such as the full Mach–Zehnder scheme illustrated in figure 1, focus on maximising the precision for a given average number of photons [2, 5]. While there are families of fixed number states—particularly N00N [27, 28] and Holland–Burnett [2, 29]—states which always have exactly the same number of photons, in general variable number states are of interest [2, 4]. Among states without a fixed total number, Gaussian probes such as squeezed states can be considered to have relatively normal photon number distributions [4, 30]. However, states with more irregular photon number distributions have also been considered [18, 19, 31].

In this section we introduce the framework of local estimation based around the Cramér–Rao bound (CRB) [32] which we use in our first pass of the problem. We revisit this under a Bayesian framework in section 5 to allow a more general control of the initial level of ignorance [33].

2.1. Precision in quantum mechanics

For a given probe state ρ and measurement $\mathbf{\Pi}$ the precision attainable by any *unbiased* estimate of a parameter is bounded by the CRB. Measurement-independent lower bounds on the CRB then exist as quantum Cramér–Rao bound (QCRBs), which depend only on the quantum state. For single-parameter estimation the symmetric logarithmic derivative (SLD) QCRB is the relevant QCRB, as there exists a measurement which can satisfy the QCRB inequality. This forms a hierarchy of inequalities bounding the variance of the estimator [34, 35]

$$(\Delta\varphi)^2 \geq \frac{1}{\nu \mathcal{I}(\rho(\varphi); \mathbf{\Pi})} \geq \frac{1}{\nu \mathcal{J}(\rho(\varphi))}, \quad (1)$$

where ν is the number of repetitions, $\mathcal{I}(\rho(\varphi); \mathbf{\Pi})$ the (state- and measurement-dependent) classical Fisher information (CFI), and $\mathcal{J}(\rho(\varphi))$ the (state- but not measurement-dependent) SLD quantum Fisher information (QFI) (henceforth QFI).

2.2. Precision in interferometry

For pure states, the QFI for estimation of a phase shift $\varphi = \phi_1 - \phi_2$ is straightforwardly given by [4, 35, 36]

$$\mathcal{J}_\varphi(|\psi\rangle) = \langle (\hat{n}_1 - \hat{n}_2)^2 \rangle - \langle \hat{n}_1 - \hat{n}_2 \rangle^2 \quad (2)$$

from the generating Hamiltonian $K = (\hat{n}_1 - \hat{n}_2)/2$. The QCRB then gives a bound

$$(\Delta\varphi)^2 \geq \frac{1}{\nu [\langle (\hat{n}_1 - \hat{n}_2)^2 \rangle - \langle \hat{n}_1 - \hat{n}_2 \rangle^2]}. \quad (3)$$

For mixed states the required expressions can be far more involved [4, 35, 37] and the mixed state QFIs necessary for this work are derived in appendix A1.

Among fixed-number states the maximum QFI comes from the N00N state [1, 2], which is a balanced superposition of all photons in one mode or the other

$$|\psi_{\text{N00N}}\rangle = \frac{|N0\rangle + |0N\rangle}{\sqrt{2}}, \quad (4)$$

directly incident on the phase shifts as figure 2. Using equation (2), $\hat{n}_1|\psi_{\text{N00N}}\rangle = N|N0\rangle/\sqrt{2}$ and $\hat{n}_2|\psi_{\text{N00N}}\rangle = N|0N\rangle/\sqrt{2}$ give

$$\mathcal{J}(|\psi_{\text{N00N}}\rangle) = N^2, \quad (5)$$

giving a Heisenberg-scaling precision $(\Delta\varphi)^2 \geq 1/\nu N^2$.

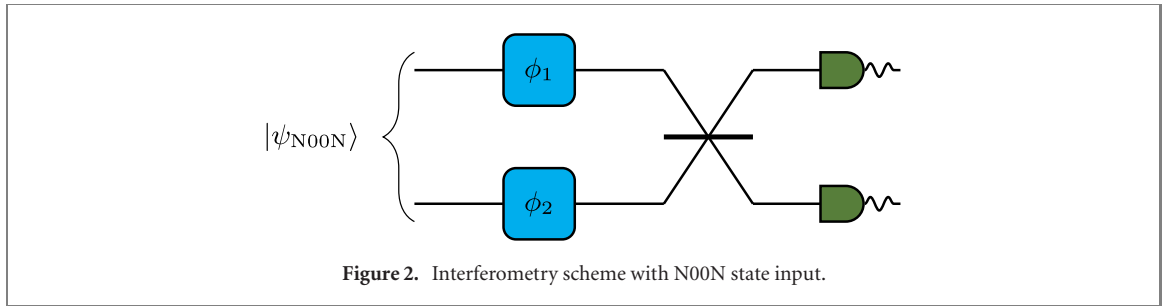


Figure 2. Interferometry scheme with N00N state input.

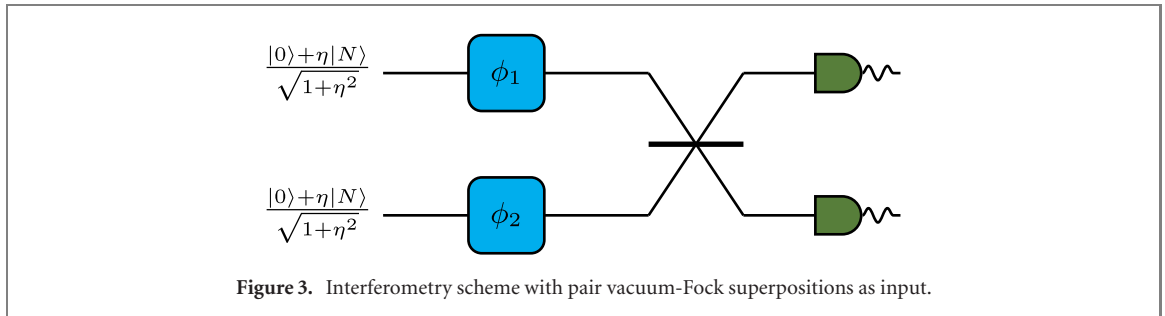


Figure 3. Interferometry scheme with pair vacuum-Fock superpositions as input.

2.3. Phase references

The conventional measurement in optical interferometry is photon counting which is a phase-insensitive technique, although it can be used with linear optics to resolve relative phases it is insensitive to a local phase. Where a state of indefinite photon number is used with phase-insensitive measurements the pure state (single-parameter) QCRB is not necessarily sufficient to properly model the problem [4, 38].

A single-mode state of indefinite number can—does, if pure—possess coherence between components of different total number which accumulate a local phase. This phase is not directly accessible through phase-insensitive measurements—the probabilities $\{|\langle\psi|n\rangle|^2\}$ depend only on the magnitude of the amplitude in the Fock basis—without introducing (an) additional state(s) with a well-defined relative phase. This is done both explicitly in the two-mode Mach–Zehnder or implicitly in the measurement as e.g. a local oscillator in homodyne and heterodyne detection [39].

While these phase-sensitive measurements are reasonable to consider in general, their availability fundamentally changes the formulation of the problem: it becomes capable to resolve a local phase shift with single-mode probe states [38, 40]. As such a Mach–Zehnder scheme with phase-sensitive measurements may contain unnecessary overhead and perform worse than a phase-sensitive measurement on a single-mode probe state under an equivalent resource.

When we wish to explicitly consider phase-insensitive measurements (such as photon counting in figure 1) the input probe state should be modified to erase coherences with respect to different degenerate eigenstate subspaces of the total number operators as [38]

$$\rho_\psi = \frac{1}{2\pi} \int_0^{2\pi} d\vartheta e^{i\vartheta(\hat{n}_1 + \hat{n}_2)} |\psi\rangle\langle\psi| e^{-i\vartheta(\hat{n}_1 + \hat{n}_2)}, \quad (6)$$

whence it follows $e^{i\zeta(\hat{n}_1 + \hat{n}_2)} \rho_\psi e^{-i\zeta(\hat{n}_1 + \hat{n}_2)} = \rho_\psi$. Unless $|\psi\rangle$ is a state of fixed total number then ρ_ψ is necessarily a mixed state and $\mathcal{J}(\rho_\psi)$ cannot be evaluated by equation (2), however this does not preclude $\mathcal{J}(\rho_\psi) = \mathcal{J}(|\psi\rangle)$ which is in fact true for all the states we consider in the following section.

3. Obfuscation through average number

We first consider the vacuum-Fock superposition states

$$|\psi_{0N}^{\otimes 2}\rangle = \frac{(|0\rangle + \eta|N\rangle)^{\otimes 2}}{1 + \eta^2} = \frac{|0\rangle + \eta(|N0\rangle + |0N\rangle) + \eta^2|NN\rangle}{1 + \eta^2} \quad (7)$$

states which are incident directly on a relative phase as in figure 3. These can be considered a Fock state analog of the Rivas-Luis state [18] to which we would expect the main ideas of this section to generalise.

These $|\psi_{0N}\rangle$ states are a simple example to illustrate why we should be concerned with average number representing the entire description, ‘danger’, and power of a variable number state. We do not propose

them as viable probe states: the generation of such a state $|\psi_{0N}\rangle$ can be expected to be demanding and difficult to scale [41], before worrying about the relative magnitudes of the amplitudes or the simultaneous generation of two copies with fixed and known relative phase. Nor as worthwhile probe states: as well as numerous information-theoretic concerns made previously [20, 22], we will demonstrate these in fact offer no advantage over an *equivalent* N00N state.

3.1. Arbitrary precision from fixed average number states

On average $|\psi_{0N}^{\otimes 2}\rangle$ has

$$\langle N \rangle = \langle \psi_{0N}^{\otimes 2} | \hat{n}_{\text{Tot}} | \psi_{0N}^{\otimes 2} \rangle = \frac{\eta^2 N}{2(1 + \eta^2)} \quad (8)$$

photons. The QFI for $\varphi = \phi_1 - \phi_2$ can be evaluated from equation (2) to

$$\mathcal{J}_\varphi(|\psi_{0N}^{\otimes 2}\rangle) = \frac{2\eta^2 N^2}{(1 + \eta^2)^2}. \quad (9)$$

Equation (9) is quadratic in N while the average photon number is linear in N and so we attain Heisenberg scaling for a fixed η .

However, rewriting equation (9) in terms of N and $\langle N \rangle$ —by solving equation (8) for η , namely $\eta = \sqrt{\frac{\langle N \rangle}{2N - \langle N \rangle}}$ which is valid for $N > \langle N \rangle / 2$ —we have

$$\mathcal{J}_\varphi(|\psi_{0N}^{\otimes 2}\rangle) = \langle N \rangle \left(N - \frac{\langle N \rangle}{2} \right), \quad (10)$$

which can be made arbitrarily large for any fixed $\langle N \rangle$ by making N arbitrarily large and adjusting η accordingly.

Observations of such large QFIs have been made before [18, 19]. One problem that should be anticipated from similar states is in finite trials due to issues with the necessary repetitions [20, 21, 42, 43] or necessary prior knowledge [44–46]. While these limitations amply motivate alternative probe states, here we will show that the apparent surpassing of Heisenberg scaling in $\langle N \rangle$ can be explained also within local estimation itself.

With a N00N state the same precision can be reached with $\sqrt{2}\eta N / (1 + \eta^2)$ photons (or at least approximated with the nearest integers). For $\eta \geq 1/\sqrt{2}$ this requires fewer photons on average, however when $\eta \ll 1$ (which is the case when equation (10) is made to substantially exceed Heisenberg scaling) the N00N state requires far more photons than the average number in the vacuum-Fock superpositions for the same precision. This suggests the vacuum-Fock superposition (with $\eta \ll 1$) should be preferred over a N00N state with equivalent QFI due to the lower average number. We will argue that, even if both states can be easily generated, such a predilection is ill-founded and may be deeply misguided.

3.2. Excessive damage through detection events

We can divide the potential measurement outcomes into three groups: (a) neither detector clicks, (b) both detectors click a total of N times, (c) both detectors click a total of $2N$ times. The latter two events subdivide into a combinatoric wealth of different distributions between the two detectors. Already we can see that when we register photons as having passed through the sample, the number thereof can be entirely detached from $\langle N \rangle$, given in equation (8), which vanishes in the $\eta \ll 1$ limit even though N remains comparatively large.

In each case we can identify distinct amounts of information which will be returned:

- (a) No photons passed through the relative phase, no clicks, no information;
- (b) N photons passed through the relative phase shift in superposition and their resulting distribution between the two detectors is parameter-dependent;
- (c) N photons passed through each arm ($2N$ total) and so no relative phase was accumulated.

In the first case one may be disappointed as we learn precisely nothing, though any sample is unscathed. In the second, N photons pass through the sample and the QFI and their counts display parameter-dependence. In the last though, $2N$ photons passed through the sample, yet the counting statistics are entirely uninformative.

Where photon numbers in the single-shot are of utmost concern the existence of such an event could destroy the sample or necessitate the experiment be run with lower average numbers at the expense of precision. Yet this is not apparent if we rely on the average number to be a sufficient approximation for the number of photons in the single-shot. Nor were we expecting to only gain information when at least N —rather than $\langle N \rangle$ —photons pass through the sample.

3.3. The role of coherences across total number

We can consider the projection into the subspaces spanned by $\{|00\rangle\}$, $\{|N0\rangle, |0N\rangle\}$, and $\{|NN\rangle\}$ —the effective state to phase-insensitive measurements—of $|\psi_{0N}^{\otimes 2}\rangle$ according to equation (6), namely $\rho_{\psi_{0N}^{\otimes 2}}$. This gives us the state

$$\rho_{\psi_{0N}^{\otimes 2}} = \frac{|00\rangle\langle 00| + 2\eta^2|\psi_{N00N}\rangle\langle\psi_{N00N}| + \eta^4|NN\rangle\langle NN|}{(1 + \eta^2)^2}. \tag{11}$$

The QFI of $\rho_{\psi_{0N}^{\otimes 2}}$, evaluated in appendix A.1.1 using the SLD, is

$$\mathcal{J}(\rho_{\psi_{0N}^{\otimes 2}}) = \frac{2\eta^2 N^2}{(1 + \eta^2)^2}, \tag{12}$$

matching equation (9), whence we see that the optimal measurement for the probe state $|\psi_{0N}^{\otimes 2}\rangle$ is phase-insensitive. Moreover, as only the $|\psi_{N00N}\rangle$ component in both $|\psi_{0N}^{\otimes 2}\rangle$ and $\rho_{\psi_{0N}^{\otimes 2}}$ acquires a φ dependence ($e^{i\varphi(\hat{n}_1 - \hat{n}_2)/2}|MM\rangle = |MM\rangle$) we can identify that a sufficient measurement to attain the QCRB is one which both attains the QCRB for $|\psi_{N00N}\rangle$ and distinguishes between components of different total photon number—the photon counting illustrated in figure 3 is one such measurement.

This probabilistic mixture of the definite total-number components of $|\psi_{0N}^{\otimes 2}\rangle$ is no longer mode-separable. Now, the only coherence is found in the (mode-entangled) N00N state term, which was implicitly present in the expansion of the tensor product in equation (7). The appearance of this N00N state component is not coincidental but is responsible—alongside its vanishing weighting for $\eta \ll 1$ —for the arbitrarily large QFI of the $|\psi_{0N}^{\otimes 2}\rangle$ state.

3.4. A probabilistic alternative (interpretation)

While we introduced $\rho_{\psi_{0N}^{\otimes 2}}$ in the previous section as the effective state seen by phase-insensitive measurements it is itself a valid probe state—one which is just as capable as $|\psi_{0N}^{\otimes 2}\rangle$. Indeed whether created independently or by a suitable collapse on $|\psi_{0N}^{\otimes 2}\rangle$ it is informative to understanding the underpinnings of the better-than-Heisenberg behavior displayed by both $|\psi_{0N}^{\otimes 2}\rangle$ and $\rho_{\psi_{0N}^{\otimes 2}}$.

Convexity of the QFI upperbounds the QFI of any mixed state as [47–49]

$$\mathcal{J}\left(\sum_k p_k |\psi_k\rangle\langle\psi_k|\right) \leq \mathcal{I}(\{p_k\}) + \sum_k p_k \mathcal{J}(|\psi_k\rangle\langle\psi_k|), \tag{13}$$

where the two terms are the CFI of the probabilities attached to the pure state decomposition and the average of the QFIs of the associated pure states.

For $\rho_{\psi_{0N}^{\otimes 2}}$ using the decomposition given in equation (11), the convexity inequality is satisfied with $\mathcal{J}(\psi_{N00N}) = N^2$ and $p_{\psi_{N00N}} = 2\eta^2/(1 + \eta^2)^2$, and $\mathcal{J}(|00\rangle) = \mathcal{J}(|NN\rangle) = \mathcal{I}(\{\frac{1}{(1+\eta^2)^2}, \frac{2\eta^2}{(1+\eta^2)^2}, \frac{\eta^4}{(1+\eta^2)^2}\}) = 0$. Thus using $\rho_{\psi_{0N}^{\otimes 2}}$ ($|\psi_{0N}^{\otimes 2}\rangle$) as probe state is metrologically equivalent to probabilistically using the equivalent N00N state with the same N and equivalent probability $p_{\psi_{N00N}} = \langle\psi_{N00N}|\rho_{\psi_{0N}^{\otimes 2}}|\psi_{N00N}\rangle = |\langle\psi_{N00N}|\psi_{0N}^{\otimes 2}\rangle|^2$.

This is not to say that $\rho_{\psi_{0N}^{\otimes 2}}$ and $|\psi_{0N}^{\otimes 2}\rangle$ are entirely equivalent. If we prepare $\rho_{\psi_{0N}^{\otimes 2}}$ by probabilistically generating $|00\rangle$, $|\psi_{N00N}\rangle$, or $|NN\rangle$ or measuring total number of $|\psi_{0N}^{\otimes 2}\rangle$ we can technically improve on the strategy by not preparing or discarding the $|NN\rangle$ component. Without material change to the interferometry scheme using instead

$$\rho_{\psi_{0NN}} = \frac{1 + \eta^4}{(1 + \eta^2)^2}|00\rangle\langle 00| + \frac{2\eta^2}{(1 + \eta^2)^2}|\psi_{N00N}\rangle\langle\psi_{N00N}|, \tag{14}$$

has a negligible effect on $\langle N \rangle$ with $2\eta^2 N/(1 + \eta^2)^2$ dominating $2\eta^4 N/(1 + \eta^2)^2$ when $\eta \ll 1$, but still has the same QFI $2\eta^2 N/(1 + \eta^2)^2$. More importantly though it disposes of the potentially damaging but metrologically useless $2N$ events, making a significant difference to the worst-case in spite of a vanishing change in $\langle N \rangle$.

As $|\psi_{0N}^{\otimes 2}\rangle$ is to $\rho_{\psi_{0N}^{\otimes 2}}$, a coherent form of $\rho_{\psi_{0NN}}$ can also be defined as

$$|\psi_{0NN}\rangle = \frac{\sqrt{1 + \eta^4}}{1 + \eta^2}|0\rangle + \frac{\sqrt{2\eta}}{1 + \eta^2} \frac{|N0\rangle + |0N\rangle}{\sqrt{2}}, \tag{15}$$

which shares average number and QFI with $\rho_{\psi_{0NN}}$. However, as with $|\psi_{0N}^{\otimes 2}\rangle$, this can be thought of as nothing more than a coherent method to probabilistically apply a N00N state.

Table 1. States, select properties of their total number distributions, and QFI for the discussed pure states.

ρ	$\langle N \rangle$	$\langle 0 \rho 0 \rangle$	$\mathcal{J}(\rho)$	$\frac{\mathcal{J}(\rho)}{1-\langle 0 \rho 0 \rangle}$
$ \psi_{N00N}\rangle$	N	0	N^2	N^2
$ \psi_{0N}^{\otimes 2}\rangle, \rho_{\psi_{0N}^{\otimes 2}}$	$\frac{2\eta^2 N}{1+\eta^2}$	$\frac{1}{(1+\eta^2)^2}$	$\frac{2\eta^2 N^2}{(1+\eta^2)^2}$	$\frac{N^2}{1+\frac{\eta^2}{2}}$
$ \psi_{0NN}\rangle, \rho_{\psi_{0NN}}$	$\frac{2\eta^2 N}{(1+\eta^2)^2}$	$\frac{1+\eta^4}{(1+\eta^2)^2}$	$\frac{2\eta^2 N^2}{(1+\eta^2)^2}$	N^2

3.5. Implications of satisfying the convexity bound

At this point we can be convinced that the metrological value of $\rho_{\psi_{0N}^{\otimes 2}}$ or $\rho_{\psi_{0NN}}$ is the (effective) probabilistic application of a N00N state. Considering either the expansion of the tensor product in equation (7) or the (equal QFI) effective state for phase-insensitive measurements in equation (11), the same appears to be true of $|\psi_{0N}^{\otimes 2}\rangle$. Here again we find ourselves dissatisfied with average number: we have associated the precision of our probe state directly with a single component of fixed number N which is detached from $\langle N \rangle$ (with arbitrary η).

This is not to say that the probabilistic application of the N00N state is a valuable or even viable strategy compared to deterministic application. When vacuum is applied nothing is done and the measurement results are of no value—non-zero contributions to both average number and QFI come solely from the N00N state component. It is simply the interplay between the QFI displaying a quadratic scaling in N while average number is only linear in N which—as noted previously in coherent [18] and incoherent [4, 26] mixtures—gives rise to the arbitrarily large QFI discussed in section 3.1.

Indeed, if probabilistically preparing vacuum were a mechanism to genuinely increase precision while reducing sample exposure we could note that we can apply the vacuum as probe state and discard the (deterministic) measurement results without reducing the quality of the estimate. However, if we are not using the measurement results and consider the vacuum state to have no effect on the sample [50] no practical distinction can be drawn between performing or simulating these vacuum applications. This would open the farcical situation that any Heisenberg-scaling scheme can be ‘improved’ with respect to the average number by *simulating* a classical mixture of the probe state and a vacuum probe. Such improvement is patently false as the same useful applications of a Heisenberg-scaling probe state have simply been diluted with additional zero-cost, zero-information probe states. Instead, to avoid this loophole of guaranteed yet free arbitrary precision from any better than shot-noise probe(s), we should conclude that the cost of the classical mixture cannot be well-represented by the average number $\langle N \rangle$.

3.6. A more appropriate resource

We have considered a number of different states: $|\psi_{N00N}\rangle$, $|\psi_{0N}^{\otimes 2}\rangle$, $\rho_{\psi_{0N}^{\otimes 2}}$, $\rho_{\psi_{0NN}}$, and $|\psi_{0NN}\rangle$, possessing various total number distributions, average number, and QFI. Yet in each instance the QFI is that of the N -particle N00N state scaled by the probability of collapsing each state into that N -particle N00N state. The states also share an identical SLD operator as appendix A.1.1 shows.

It is our view that these—and the associated mixed states generated by projecting in total number—all derive their metrological power solely from their possession of a N00N state component. This makes the other—where $\eta \ll 1$, largely vacuum—components only relevant to the average number (for a given amplitude of the N00N state). If one were to condition on events where a photon is detected (a necessary but not sufficient condition for such an event to be informative as to the phase shift) then, as table 1 shows, the states with a vacuum-component appear no-better, if not worse.

Rather than η being a parameter which can be tuned to increase the QFI, it should be considered a parameter to reduce $\langle N \rangle$ by obfuscating the N00N component solely responsible for the metrological power of all these states. As such N , not $\langle N \rangle$, is a more representative resource for these states in terms of both the photons passing through the sample and their metrological performance.

4. Probabilistic sensing

The states of the previous example are particularly convenient to analyse, however, the fundamental conceit of doing nothing—by coherent or incoherent means—to artificially deflate average number can be applied more generally. In both cases the average number of the resultant state becomes detached from the original state responsible for all phase-sensitivity giving a false impression of improved sensitivity per photon.

4.1. Probabilistically sensing

A two-mode state $|\phi_M\rangle$ with average number M which is either used (with probability p) or replaced with vacuum (with probability $1 - p$)—without adapting the interferometer or measurement—as input to the interferometer is described by the state

$$\rho = (1 - p)|0\rangle\langle 0| + p|\phi_M\rangle\langle\phi_M|, \quad (16)$$

which—by definition of $|\phi_M\rangle\langle\phi_M|$ —contains on average pM photons.

The QFI is (see appendix A.1.2)

$$\mathcal{J}(\rho) = 4p \left(\langle\phi'|\phi'\rangle - |\langle\phi'|\phi\rangle|^2 \frac{1 - p|\langle\phi|0\rangle|^2}{1 - |\langle\phi|0\rangle|^2} \right) = p\mathcal{J}(|\phi\rangle) - \frac{p(1 - p)\langle\hat{n}_1 - \hat{n}_2\rangle^2 |\langle\phi|0\rangle|^2}{1 - |\langle\phi|0\rangle|^2}, \quad (17)$$

which is upperbounded as $\mathcal{J}(\rho) \leq p\mathcal{J}(|\phi\rangle)$, with the equality holding for any $|\phi\rangle$ either orthogonal to the vacuum state ($\langle 0|\phi\rangle = 0$), or symmetric in the two modes ($\langle\hat{n}_1 - \hat{n}_2\rangle = 0$). Adapting the probability with which $|\phi_M\rangle$ is used to fix p to be $\langle N\rangle/M$ —for any $\langle N\rangle \leq M$ —means the QFI is at most

$$\mathcal{J}(\rho) \leq \langle N\rangle \frac{\mathcal{J}(|\phi_M\rangle)}{M}. \quad (18)$$

For states $|\phi_M^S\rangle$ which only reach a shot-noise scaling $\mathcal{J}(|\phi_M^S\rangle) \sim M$, equation (18) gives that $\mathcal{J}(\rho^S) \lesssim \langle N\rangle$ —classical states are constrained to only give a classical scaling when applied in such a probabilistic manner. By comparison a Heisenberg scaling state possessing $\mathcal{J}(|\phi_M^H\rangle) \sim M^2$, gives a QFI $\mathcal{J}(\rho_M^H) \lesssim \langle N\rangle M$, which—at least when $\langle\phi|0\rangle \sim 0$ or $\langle\hat{n}_1 - \hat{n}_2\rangle \sim 0$ —can be taken to be arbitrarily large with suitably large M , while—for any $M > \langle N\rangle$ — p can be adjusted to achieve the fixed average number $\langle N\rangle$.

Here it may also make sense to consider M as the more appropriate resource constraint for the states defined in equation (16). With respect to M , $\mathcal{J}(\rho_M^H)$ does not display any better-than-Heisenberg scaling and p acts again to reduce the average number without positively contributing to the metrological performance. However, this relies on $|\phi_M\rangle$ itself being well-represented by its average number; this fails (as it should) if $|\phi_M\rangle = |\psi_{0N}^{\otimes 2}\rangle$.

4.2. Coherent probabilistic applications

Now we turn to a coherent form of the probabilistic scheme where

$$|\psi\rangle = \sqrt{1 - p}|0\rangle + \sqrt{p}|\phi_M\rangle. \quad (19)$$

Limiting to $\langle 0|\phi_M\rangle = 0$ such that $|\psi\rangle$ is normalised, we have $\hat{n}_j|\psi\rangle = \sqrt{p}\hat{n}_j|\phi_M\rangle$ and so the QFI is

$$\mathcal{J}(|\psi\rangle) = p\langle(\hat{n}_1 - \hat{n}_2)^2\rangle_{\phi_M} - p^2\langle\hat{n}_1 - \hat{n}_2\rangle_{\phi_M}^2. \quad (20)$$

This coherent probabilistic scheme can technically improve on the probabilistic scheme outlined in section 4.1 by the relatively simple comparison of equations (17) and (20), noting equation (20) is greater than $p\mathcal{J}(|\phi_M\rangle)$ due to the additional factor p reducing the $\langle\hat{n}_1 - \hat{n}_2\rangle_{\phi_M}^2$ term. Again, with a Heisenberg scaling $|\phi_M\rangle$, $|\psi\rangle$ can achieve a better-than-Heisenberg scaling with respect to $\langle N\rangle$, but not M .

4.3. Average number

In both the incoherent and coherent settings discussed, we can perform the same conditioning on events where a non-zero number of photons are detected (which is necessary but not sufficient for the event to be informative). Considering $\langle 0|\phi_M\rangle = 0$, photons are detected with probability p and precision can only increase when this happens. Conditioning on such detection events—retaining $p = \langle N\rangle/M$ —we find for ρ

$$\frac{\mathcal{J}(\rho)}{p} = \mathcal{J}(|\phi_M\rangle) \quad (21)$$

and for $|\psi\rangle$

$$\frac{\mathcal{J}(|\psi\rangle)}{p} = \mathcal{J}(|\phi_M\rangle) + (1 - p)\langle\hat{n}_1 - \hat{n}_2\rangle_{\phi_M}^2 \quad (22)$$

where the latter term can be upper-bounded by M^2 (as $1 - p \leq 1$ and $\langle\hat{n}_1 - \hat{n}_2\rangle^2 \leq \langle\hat{n}_1 + \hat{n}_2\rangle^2$) which means for Heisenberg-scaling $|\phi_M\rangle$ with $\mathcal{J}(|\phi_M\rangle) = M^2$, $\frac{\mathcal{J}(|\psi\rangle)}{p} \leq (2 - p)\mathcal{J}(|\phi_M\rangle)$.

5. Bayesian analysis

In section 3 we focused exclusively on the CRB and QFI as a tool to argue that certain variable-number states derive their metrological power from a common source, and that this fact can be obscured when the average number is employed as the resource. We now consider a Bayesian single-shot analysis to show that these observations transcend the local nature of the CRB.

5.1. Metrological power in the presence of prior information

Instead of a hierarchy of inequalities (e.g. equation (1)) which must be saturated sequentially in order to describe the attainable precision, the Bayesian framework enables a flexible formulation of optimisation problems whose solution, if available, readily provides the *best* possible measurement and phase estimate [51], together with the associated uncertainty, for an appropriate criterion of performance [52].

Under the mean square error criterion [17, 52], the optimal single-shot measurement is given by the eigenvectors of an operator S , solution to the equation $S\rho + \rho S = 2\rho'$, where $\rho = \int_a^b d\varphi z(\varphi)\rho(\varphi)$, $\rho' = \int_a^b d\varphi z(\varphi)\rho(\varphi)\varphi$, and $z(\varphi)$ is a prior probability with support bounded by a and b [43, 53]. In this work, $\rho(\varphi) = \exp(-i\varphi K)\rho(0)\exp(i\varphi K)$, with $K = (\hat{n}_1 - \hat{n}_2)/2$ and $\rho(0)$ the initial state fed to the interferometer. Furthermore, the j th eigenvalue of S gives the optimal estimate for the phase shift φ when the j th projector is measured, and the optimal uncertainty is [43, 53], for a single shot [54],

$$\langle(\Delta\varphi)^2\rangle_{\text{opt}} = \int_a^b d\varphi z(\varphi)\varphi^2 - \text{Tr}(\rho S^2), \quad (23)$$

where $\langle(\Delta\varphi)^2\rangle$ denotes the prior-averaged mean square error [4, 33, 55]. The optimal phase estimates associated with equation (23) are generally biased, for, in Bayesian estimation, limiting the problem to unbiased estimators is not needed, nor necessarily beneficial. This is in contrast to earlier sections where $(\Delta\varphi)^2$ refers to the variance of an unbiased estimator, with unbiasedness being required there as biased estimators can display low variance despite, due to large bias, high error.

Next, it is noted that the term $\text{Tr}(\rho S^2)$ is suggestive of the mixed-state definition for the QFI (cf appendix A.1). However, since the prior $z(\varphi)$ enters equation (23) non-trivially, a more careful examination is needed if we are to find a sensible quantifier of metrological power in Bayesian interferometry.

If the prior is sufficiently narrow, equation (23) may reasonably be approximated as $\langle(\Delta\varphi)^2\rangle_{\text{opt}} \approx \sigma_0^2 [1 - \sigma_0^2 \mathcal{J}(\rho)]$ [43, 56, 57], where $\mathcal{J}(\rho)$ is the QFI employed so far and $\sigma_0^2 = \int_a^b d\varphi z(\varphi)\varphi^2 - [\int_a^b d\varphi z(\varphi)\varphi]^2$ is the initial uncertainty. This motivates rewriting equation (23) in a similar form, but now without the aforementioned approximation, as

$$\langle(\Delta\varphi)^2\rangle_{\text{opt}} = \sigma_0^2 [1 - \sigma_0^2 \mathcal{P}(\rho; z)], \quad (24)$$

where we have defined the quantity

$$\mathcal{P}(\rho; z) := \frac{\text{Tr}(\rho S^2) - \text{Tr}(\rho S)^2}{\sigma_0^4}, \quad (25)$$

and $\text{Tr}(\rho S) = \int_a^b d\varphi z(\varphi)\varphi$, with ρ and ρ' denoting the initial and the prior-averaged parameter-encoded states. Note that equation (25) recovers the QFI locally, i.e. $\mathcal{P}(\rho; z) \rightarrow \mathcal{J}(\rho)$ in the limit of a narrow prior. We now argue that $\mathcal{P}(\rho; z)$ is a valid and more general quantifier of metrological power.

By construction, $0 \leq \mathcal{P}(\rho; z) \leq 1/\sigma_0^2$. If $\mathcal{P}(\rho; z) = 0$, then $\langle(\Delta\varphi)^2\rangle_{\text{opt}} = \sigma_0^2$, and so no information is gained by the application of the scheme. On the contrary, if $\mathcal{P}(\rho; z) = 1/\sigma_0^2$, then $\langle(\Delta\varphi)^2\rangle_{\text{opt}} = 0$, which would imply that the relative phase is perfectly resolved. Enhancing the precision is thus equivalent to reducing the prior error by making $\mathcal{P}(\rho; z)$ larger. Unlike the QFI, $\mathcal{P}(\rho; z)$ can, regardless of ρ , only be unbounded when $\sigma_0 = 0$, where even an infinite amount of information would not improve the estimate since there is no uncertainty to be reduced.

The quantifier $\mathcal{P}(\rho; z)$ thus possesses the properties that a good measure of metrological power should have, while generalising the QFI. Specifically, in the limit where $\mathcal{P}(\rho; z) \rightarrow \mathcal{J}(\rho)$, equation (25) leads to the *same* hierarchy of probes predicted by local estimation; otherwise, the Bayesian quantifier will generally reveal more accurate information about metrological enhancements.

Note that, in general, one should use a criterion of performance that respects the periodicity of phase shifts [4, 58], and equation (23) does not. Nevertheless, if the prior is effectively concentrated within an interval of length $W = (b - a) \lesssim 2$, then the square error giving rise to equation (23) approximates a sine square error, which is a valid cost function for periodic parameters [4, 43, 58]. Admittedly, this restricts the quantifier $\mathcal{P}(\rho; z)$ to an intermediate regime between local estimation ($W \ll 1$) and a fully global

estimation framework ($W \leq 2\pi$). Yet, such a regime has been shown to be sufficiently non-local (i.e. to allow for sufficiently wide priors) to reveal non-trivial effects beyond QFI analyses in interferometry [43], and so it suffices for our purpose.

5.2. Average number in Bayesian interferometry

We now demonstrate that the decoupling of average number and the origin of the metrological power is not limited to local estimation. To ensure equivalent prior knowledge in each case, we omit a direct analysis of $\psi_{0N}^{\otimes 2}$ and ψ_{0NN} , and instead use their effective phase-insensitive measurement forms $\rho_{\psi_{0N}^{\otimes 2}}$ and $\rho_{\psi_{0NN}}$, which are also of independent interest viz section 3. Appendix A.2 provides the details of these calculations.

Consider a N00N state and prior $z(\varphi) = 1/W$, with support $\varphi \in [-W/2, W/2]$; the Bayesian quantifier (now omitting the prior dependency) from equation (25) reads

$$\mathcal{P}(|\psi_{N00N}\rangle) = \kappa(NW/2)N^2, \quad (26)$$

where $\kappa(x)$ is function with range $[0, 1]$ and defined as $\kappa(x) = 9(x \cos x - \sin x)^2/x^6$. The limit $\kappa \rightarrow 1$, which is realised when local prior information is available (i.e. $W \rightarrow 0$), leads to $\mathcal{P}(|\psi_{N00N}\rangle) = N^2$; this not only is consistent with the QFI, but also bounds $\mathcal{P}(|\psi_{N00N}\rangle)$ for any finite W . By fixing $NW = \zeta$, with finite W , we further recover the well-known result that N00N states achieve a Heisenberg scaling

$$\mathcal{P}(|\psi_{N00N}\rangle) = \kappa(\zeta/2)N^2 \quad (27)$$

subject to the prior being localised to an interval $W = \zeta/N$. Moreover, for $\zeta \lesssim 1$ we have $\kappa \sim 1$, implying that $\mathcal{P}(|\psi_{N00N}\rangle) \sim N^2$ when the phase is initially localised to an interval $W \sim 1/N$ [22, 44].

Now, compare this scheme with the variable-number state in equation (11), $\rho_{\psi_{0N}^{\otimes 2}}$, which, assuming the same flat prior, leads to

$$\mathcal{P}(\rho_{\psi_{0N}^{\otimes 2}}) = \frac{2\eta^2 \mathcal{P}(|\psi_{N00N}\rangle)}{(1+\eta^2)^2} = \frac{2\eta^2 N^2 \kappa(\zeta/2)}{(1+\eta^2)^2}. \quad (28)$$

Recalling that, for this scheme, $\langle N \rangle = 2\eta^2 N/(1+\eta^2)$, equation (28) can be written as

$$\mathcal{P}(\rho_{\psi_{0N}^{\otimes 2}}) = \kappa(\zeta/2)\langle N \rangle \left(N - \frac{\langle N \rangle}{2} \right) \approx \kappa(\zeta/2)\langle N \rangle N, \quad (29)$$

where the approximation assumes that, for a fixed (and finite) $\langle N \rangle$, $N \gg \langle N \rangle$. In other words, $\rho_{\psi_{0N}^{\otimes 2}}$ cannot provide a precision scaling better than $\mathcal{P}(\rho_{\psi_{0N}^{\otimes 2}}) \sim \langle N \rangle N$, where N , and not $\langle N \rangle$, gauges the precision. Moreover, the Heisenberg scaling is not violated with respect to N .

As with the QFI, here too we are led to conclude that N may be a more representative resource constraint. This is further reinforced by the existence of a key relationship between N and the prior width W —namely, $W \lesssim 1/N$ —which is needed for $\mathcal{P}(\rho_{\psi_{0N}^{\otimes 2}}) \sim \langle N \rangle N$ to hold with a coefficient of order unity. Given that the average number $\langle N \rangle$ plays no role in such relationship, using $\langle N \rangle$ as a proxy without examining the role of N could lead to underestimating the prior information needed to exploit the metrological power of a state. For instance, we could be led to rely on $W \sim 1/\langle N \rangle$ which, should it be an insufficient amount of prior information, risks giving rise to exposure to a high-power probe state that does not extract information beyond the initial knowledge; such a zero-gain in spite of a prior $W \sim 1/\langle N \rangle$ has been noted previously [46].

Still, there is no doubt that $\rho_{\psi_{0N}^{\otimes 2}}$ has *some* metrological value—Luis [42] also showed in the context of Bayesian estimation that the single-mode $|\psi_{0N}^{\otimes 2}\rangle$ has some value—however, as in the QFI discussion, its source may be linked, tentatively, to the N00N-state component. First, note that the constraint $W \lesssim 1/N$ is a consequence of the form of the Bayesian precision for N00N states, regardless of the variable-number state to which the N00N-state component may belong to.

Secondly, if we generalise the notion of resource introduced in section 3.6 from $\mathcal{J}(\rho)/p$ to $\mathcal{P}(\rho)/p$, where p was the probability of detecting photons at the output ports, we find

$$\frac{\mathcal{P}(\rho_{\psi_{0N}^{\otimes 2}})}{p} = \frac{\mathcal{P}(|\psi_{N00N}\rangle)}{1 + \frac{\eta^2}{2}} \leq \mathcal{P}(|\psi_{N00N}\rangle), \quad (30)$$

where we have used that, for $\rho_{\psi_{0N}^{\otimes 2}}$, $p = 1 - 1/(1 + \eta^2)^2$. Equation (30) indicates that the sensitivity achieved by $\rho_{\psi_{0N}^{\otimes 2}}$ can never surpass that of a probabilistic application of its N00N-state component. If we instead examine the mixed-N00N state $\rho_{\psi_{0NN}}$ in equation (14), which removes the possibility of

non-informative and potentially damaging events (i.e. those associated with $|\text{NN}\rangle$), then we have the exact identity

$$\frac{\mathcal{P}(\rho_{\psi_{0\text{NN}}})}{p} = \mathcal{P}(|\psi_{\text{N00N}}\rangle), \quad (31)$$

where now $p = 2\eta^2/(1 + \eta^2)^2$.

Finally, in appendix A.2 we show that—as with the SLD—the quantum estimator S is that for a N00N state even when $\rho_{\psi_{0\text{N}}^{\otimes 2}}$ or $\rho_{\psi_{0\text{NN}}}$ are employed. Given the fundamental role played by S in Bayesian estimation—it contains the optimal phase estimates and their measurement scheme—this is perhaps the most convincing argument.

We conclude that the unsuitability of the average photon number to capture the true source of metrological advantage, as well as the real damage done to the sample, is not a byproduct of the local nature of the QFI, but a more general feature of quantum metrology. More practically, equations (27), (30) and (31) extend the validity of the findings in table 1, based on the QFI, to a non-local regime with moderate prior knowledge (i.e. $W \lesssim 2$).

6. Discussion

6.1. Understanding Heisenberg scaling with only an average number constraint

We have seen that the close-to-vacuum states $(|0\rangle + \eta|\phi\rangle)/\sqrt{1 + \eta^2}$ appear to owe their metrological-advantage to being effective probabilistic application of Heisenberg-scaling states. We would expect these observations to carry over to the likes of the Rivas-Luis states; section 4.2 does in principle cover them, but only through a state of the form $|\phi\rangle = |0, \xi\rangle + |\xi, 0\rangle + \alpha|\xi, \xi\rangle$ which we do not believe has garnered any direct analysis. This already raises major questions as to whether the apparent Heisenberg-violating scalings are either: actually Heisenberg-violating, or quantum in origin.

An explicitly better-than-Heisenberg scaling requires N to be explicitly parameterised in terms of η and $\langle N \rangle$, otherwise we simply have a highly favourable pre-factor to a QFI linear in $\langle N \rangle$ as equation (10), or an unfavourable pre-factor to a QFI quadratic in N as equation (9). Moreover, if we take N as our resource for these states, the parameterised form (e.g. equation (10)) is bound by a Heisenberg-scaling in N —both in isolation and conditioning on photon-involving events (table 1).

Arbitrarily large precision gains are, in any case, explicitly forbidden when working with finite prior information (i.e. σ_0 is finite), since the Bayesian quantifier $\mathcal{P}(\rho; z)$ is upper-bounded by $1/\sigma_0^2$. Even this finite improvement can be hard to achieve with better-than-Heisenberg strategies requiring either a very large number of repetitions [20, 59] or a very large amount of prior knowledge [21, 46]. Moreover, there is an extensive body of evidence supporting that a Heisenberg-scaling in the total number used across all repetitions cannot be surpassed [20, 22, 26, 60, 61].

As to the quantum origin, coherence between different total number subspaces does not contribute to the precision and performance is no different to spending a fraction of trials using N00N states and a fraction doing nothing. One can thus argue that the Heisenberg-scaling derives from the quintessentially quantum N00N state but (in the cases discussed herewanyhe apparent better-than-Heisenberg precision on top of this comes from the wholly classical ability to sometimes do nothing whether exactly or—due to coherence—only effectively.

6.2. Noise

Optical loss is known to be detrimental to precision in interferometry [62]. Of particular relevance to this work is the inevitable reduction to a shot-noise scaling in sufficiently high loss [63–65]. This gives rise to a bound on the QFI of a variable number state ρ with average number $\langle N \rangle$ of [66]

$$\mathcal{J}(\rho) \leq \frac{\langle N \rangle \gamma}{1 - \gamma}, \quad (32)$$

for fractional loss rate γ . In low noise, equation (32) has a substantial pre-factor on $\langle N \rangle$ which does not preclude Heisenberg-scaling ($\gamma/(1 - \gamma)$ can exceed $\langle N \rangle$), but for sufficiently high loss any scheme reduces to a shot-noise scaling. In such a regime $\langle N \rangle$ may suffice, though photodamage may remain a concern if the loss involves absorption or does not precede the sample.

In discussing the potential photodamage we limited our model to the number of photons passing through the probe. In practice, a more intricate model explicitly factoring in some absorption (damage) dynamic may be desirable both in quantifying a state's metrological performance and the potential damage it may cause by its application. Whether a given level of exposure, or absorption or damage is considered can give rise to different precisions [10, 67].

6.3. Beyond an average number constraint

Although Heisenberg-scaling is ordinarily identified as a quadratic scaling in $\langle N \rangle$ [2, 5, 68, 69], previous works have identified a linear scaling in $\langle N^2 \rangle$ (quadratic in $\sqrt{\langle N^2 \rangle}$) as the correct fundamental limit when variable number states are used in interferometry [24–26, 70]. For $|\psi_{0N}^{\otimes 2}\rangle$, which has $\langle N^2 \rangle = 2\eta^2(1 + 2\eta^2)N^2/(1 + \eta^2)^2$, the QFI indeed scales linearly in $\langle N^2 \rangle$ with $\mathcal{J}(|\psi_{0N}^{\otimes 2}\rangle)/\langle N^2 \rangle = 1/(1 + 2\eta^2)$. The coefficient $\mathcal{J}(|\psi_{0N}^{\otimes 2}\rangle)/\langle N^2 \rangle$ is of order unity in the Heisenberg-violating $\eta \ll 1$ regime and vanishing—as the QFI itself vanishes—in the $\eta \gg 1$ regime where the input approximates phase-insensitive Fock states incident directly on the phase shifts. Similarly, one may use equations (26)–(28) to show that, in the Bayesian framework, $\mathcal{P}(\rho_{\psi_{0N}^{\otimes 2}})/\langle N^2 \rangle \sim 1/(1 + 2\eta^2)$.

One could consider using $\sqrt{\langle N^2 \rangle}$ rather than or in addition to $\langle N \rangle$ as a constraint; this gives a much larger cost for $|\psi_{0N}^{\otimes 2}\rangle$ that better matches the average photon number when conditioned on events where a non-zero number of photons is detected. For fixed number and coherent states $\sqrt{\langle N^2 \rangle} = \langle N \rangle$, while $\sqrt{\langle N^2 \rangle} \approx \langle N \rangle$ for squeezed vacuum states [2, 71]. This allows the existing shot-noise and Heisenberg scalings to carry over for those states under a $\langle N^2 \rangle$ constraint, while marking out the vacuum-Fock or Rivas-Luis states as much more expensive than average number alone suggests.

Working to an additional constraint of $\langle N^2 \rangle$ may make more sense to understand the potential exposure to a sample, particularly in the Heisenberg-violating regimes we explored earlier, however it is not sufficient to alleviate worst-case concerns. Consider a superposition of primarily squeezed vacuum with a vanishing component of a much higher energy Fock state, in some sense an inverse situation to $|\psi_{0N}^{\otimes 2}\rangle$; we would expect to find an apparently reasonable sensitivity: quadratic in $\langle N \rangle$ and $\sqrt{\langle N^2 \rangle}$ due to the squeezed vacuum contribution dominating, but still possessing a non-zero probability of a potentially damaging event. While this situation would be better for bounding $\langle N \rangle^2$ and $\langle N^2 \rangle$, we can still have a rare, damaging, yet uninformative event—whether in terms of the number of photons passing through the sample or explicitly modeling absorption. Simply constraining the maximum number—essentially truncating the Fock space by total photon number—goes a long way to alleviate this, but such a truncation may be stricter than need be, forbidding for example the conventional Gaussian states.

Instead, there is still some cause to desire a more intricately weighted measure of information per cost—such as an extension of the quantities $\mathcal{J}(|\psi\rangle)/(1 - |\langle 0|\psi\rangle|^2)$ and $\mathcal{P}(|\psi\rangle; z)/(1 - |\langle 0|\psi\rangle|^2)$ employed in sections 3.6 and 5.2, respectively—capable of ensuring good sensing value per-photon in every aspect of the probe state. More general measures of fluctuations in the total number distribution may be sufficient for this purpose. Even simple heuristics may be of use: the total number distribution of Gaussian or fixed-number states are singly-peaked (around $\langle N \rangle$) while the vacuum-Fock and Rivas-Luis states are multiply-peaked (around 0, x , and $2x$, with x being the N of the Fock state and the average number of the squeezed vacuum state).

7. Conclusions

We have argued that average number can fail to capture properties of the full total number distribution of probe states relevant to metrological power, photodamage, and necessary prior information. While average number has been known not to properly constrain the theoretical precision of a probe state [24, 26], we argue this should be attributed to an over-reliance on average number to capture a potentially complex, multi-faceted number distribution rather than the existence of certain anomalous states.

Finally, we emphasise that while the examples considered here are informative as to the role average number should or should not be allowed to play in evaluating and comparing interferometry schemes, this does not negate the practical inaccessibility of supposed better-than-Heisenberg scaling states [20–23, 33, 72], nor can it circumvent the value of states and measurements which are practical to prepare [73] and noise-resilient [74].

Acknowledgments

We are grateful to Francesco Albarelli for valuable discussions. We thank Erik Gauger and Alfredo Luis for comments. DB acknowledges support from UK EPSRC Grant EP/R030413/1. JR acknowledges support from UK EPSRC Grant No. EP/T002875/1.

Data availability statement

No new data were created or analysed in this study.

Appendix A Precision calculations

A.1. Mixed state QCRB

For a mixed state ρ the SLD QFI is not given directly by equation (2) but can be found as [34, 35]

$$\mathcal{J}_\varphi(\rho) = \text{Tr}(\rho L_\varphi^2), \quad (\text{A1})$$

where L_φ is the the SLD operator defined by

$$\frac{\partial \rho}{\partial \varphi} = \frac{L_\varphi \rho + \rho L_\varphi}{2}. \quad (\text{A2})$$

A.1.1. Mixed state relatives of the N00N and vacuum-Fock states

In section 3 we have parameter-dependent mixed states of the form

$$\rho(\varphi) = \alpha|00\rangle\langle 00| + \beta|\psi_{\text{N00N}}(\varphi)\rangle\langle \psi_{\text{N00N}}(\varphi)| + (1 - \alpha - \beta)|\text{NN}\rangle\langle \text{NN}|, \quad (\text{A3})$$

whose derivative is

$$\partial_\varphi \rho = i\frac{\beta N}{2}e^{iN\varphi}|N0\rangle\langle 0N| - i\frac{\beta N}{2}e^{-iN\varphi}|0N\rangle\langle N0|, \quad (\text{A4})$$

and so the SLD is

$$L_\varphi = iN(e^{iN\varphi}|N0\rangle\langle 0N| - e^{-iN\varphi}|0N\rangle\langle N0|), \quad (\text{A5})$$

which is also the SLD operator for any pure state where $|\psi_{\text{N00N}}\rangle$ is the only phase-sensitive component as $L_\varphi|\psi_{\text{N00N}}(\varphi)\rangle = \partial_\varphi|\psi_{\text{N00N}}(\varphi)\rangle$. The QFI for equation (A3) is then

$$\mathcal{J}(\rho) = \beta N^2, \quad (\text{A6})$$

which generalises the specific mixed state QFIs given in section 3.

A.1.2. Probabilistic sensing

In section 4 we consider the state equation (16), whose QFI is most conveniently derived using a non-orthogonal basis [37]. With the non-orthogonal basis $\{|0\rangle, |\phi_M\rangle, |\phi'_M\rangle\}$, where $|\phi'_M\rangle = \partial_\varphi|\phi_M\rangle$, the SLD solution is

$$L_\varphi = \frac{2}{p(1 - |\langle 0|\phi_M\rangle|^2)} [p\langle \phi'_M|\phi_M\rangle\langle 0|\phi_M\rangle\langle 0| + p\langle \phi_M|\phi'_M\rangle\langle \phi_M|0\rangle\langle \phi_M|0\rangle - \langle 0|\phi_M\rangle\langle 0|\phi'_M| - \langle \phi_M|0\rangle\langle \phi'_M|0\rangle + |\phi'_M\rangle\langle \phi_M| + |\phi_M\rangle\langle \phi'_M|], \quad (\text{A7})$$

recalling $\langle \phi_M|\phi'_M\rangle + \langle \phi'_M|\phi_M\rangle = 0$ [35] and $\langle 0|\phi'_M\rangle = 0$ as $\langle 0|\phi_M(\varphi)\rangle = \langle 0|e^{i\varphi(\hat{n}_1 - \hat{n}_2)}|\phi_M(0)\rangle = \langle 0|\phi_M(0)\rangle$. The QFI is then

$$\mathcal{J}(\rho) = 4p(\langle \phi'_M|\phi'_M\rangle + \langle \phi_M|\phi'_M\rangle^2) + 4p(1 - p)\frac{|\langle \phi_M|0\rangle|^2\langle \phi'_M|\phi_M\rangle^2}{1 - |\langle \phi_M|0\rangle|^2}, \quad (\text{A8})$$

which is upperbounded by $4p(\langle \phi'_M|\phi'_M\rangle + \langle \phi_M|\phi'_M\rangle^2)$ as $\langle \phi_M|\phi'_M\rangle + \langle \phi'_M|\phi_M\rangle = \langle \phi_M|\phi'_M\rangle + \langle \phi_M|\phi'_M\rangle^* = 0$ entails $\langle \phi_M|\phi'_M\rangle^2 \leq 0$. Finally, $|\phi'_M\rangle = i\frac{\hat{n}_1 - \hat{n}_2}{2}|\phi_M\rangle$ gives equation (17).

A.2. Bayesian quantifier

Section 5.1 introduces a measure of metrological power, $\mathcal{P}(\rho; z) = [\text{Tr}(\varrho S^2) - \text{Tr}(\varrho S)^2]/\sigma_0^4$, capable of accommodating a finite and moderate amount of prior information. Given the flat prior in the main text, i.e. $z(\varphi) = 1/W$, when $\varphi \in [-W/2, W/2]$, such quantifier becomes

$$\mathcal{P}(\rho) = \frac{144 \text{Tr}(\varrho S^2)}{W^4}, \quad (\text{A9})$$

where we have used that $\sigma_0^2 = W^2/12$ and

$$\text{Tr}(\varrho S) = \int_a^b d\varphi z(\varphi) \varphi \propto \int_{-W/2}^{W/2} d\varphi \varphi = 0. \quad (\text{A10})$$

The optimal quantum estimator S is given by $S\varrho + \varrho S = 2\varrho'$, where, in this case,

$$\varrho = \frac{1}{W} \int_{-W/2}^{W/2} d\varphi \rho(\varphi), \quad \varrho' = \frac{1}{W} \int_{-W/2}^{W/2} d\varphi \rho(\varphi) \varphi. \quad (\text{A11})$$

As with the SLD in equation (A2), S is determined by a Lyapunov equation. However, the similarity ends here: unlike in the frequentist approach, in the Bayesian framework we integrate the phase-dependent state with respect to the prior, making the final solution parameter-independent, and the derivative $\partial_\varphi \rho$ no longer plays a role (although it *does* reemerge in the limit of a narrow prior, as one would expect from a local approach [43, 56]).

Consider now the family of parameter-dependent states in equation (A3). The operators in equation (A11) read

$$\varrho = \alpha|00\rangle\langle 00| + (1 - \alpha - \beta)|NN\rangle\langle NN| + \frac{\beta s_{NW}}{NW}(|0N\rangle\langle N0| + |N0\rangle\langle 0N|) + \frac{\beta}{2}(|0N\rangle\langle 0N| + |N0\rangle\langle N0|) \quad (\text{A12})$$

and

$$\varrho' = i \frac{\beta (NW c_{NW} - 2 s_{NW})}{2N^2 W} (|0N\rangle\langle N0| - |N0\rangle\langle 0N|), \quad (\text{A13})$$

where $s_{NW} = \sin(NW/2)$ and $c_{NW} = \cos(NW/2)$. The solution to $S\varrho + \varrho S = 2\varrho'$ is then

$$S = i \frac{NW c_{NW} - 2 s_{NW}}{N^2 W} (|0N\rangle\langle N0| - |N0\rangle\langle 0N|). \quad (\text{A14})$$

Introducing equations (A12) and (A14) in equation (A9), we finally arrive at

$$\mathcal{P}(\rho) = \frac{144\beta(NW c_{NW} - 2 s_{NW})^2}{N^4 W^6} = \kappa(NW/2)\beta N^2, \quad (\text{A15})$$

where we have defined the function $\kappa(x) = 9(x \cos x - \sin x)^2/x^6$. The results discussed in section 5.2 can be then recovered from equation (A15) as follows: $\beta = 1$ for equation (26) (N00N state), and $\beta = 2\eta^2/(1 + \eta^2)^2$ for both equation (28) (state $\rho_{\psi_{0N}^{\otimes 2}}$) and $\mathcal{P}(\rho_{\psi_{0NN}})$ in equation (31).

We further note that the optimal quantum estimator S in equation (A14) is independent of α and β , and so identical for all the states represented by equation (A3), which include the N00N state. This leads to the third argument given in section 5.2 to support the idea that the metrological advantage of states such as $\rho_{\psi_{0N}^{\otimes 2}}$ and $\rho_{\psi_{0NN}}$, both included in equation (A3), originates in their N00N-state components.

ORCID iDs

Dominic Branford  <https://orcid.org/0000-0001-8989-7464>

Jesús Rubio  <https://orcid.org/0000-0002-8193-8273>

References

- [1] Giovannetti V, Lloyd S and Maccone L 2006 *Phys. Rev. Lett.* **96** 010401
- [2] Lang M D and Caves C M 2014 *Phys. Rev. A* **90** 025802
- [3] Tóth G and Apellaniz I 2014 *J. Phys. A: Math. Theor.* **47** 424006
- [4] Demkowicz-Dobrzański R, Jarzyna M and Kołodyński J 2015 *Progress in Optics* vol 60 ed E Wolf (Amsterdam: Elsevier) pp 345–435
- [5] Sahota J and Quesada N 2015 *Phys. Rev. A* **91** 013808
- [6] Vahlbruch H, Mehmet M, Danzmann K and Schnabel R 2016 *Phys. Rev. Lett.* **117** 110801
- [7] Wang H *et al* 2019 *Phys. Rev. Lett.* **123** 250503
- [8] Slussarenko S, Weston M M, Chrzanowski H M, Shalm L K, Verma V B, Nam S W and Pryde G J 2017 *Nat. Photon.* **11** 700
- [9] Radiation-pressure effects in these detectors introduce a trade-off where a maximum sensitivity is reached beyond which the negative contribution from back-action outweighs the reduction in shot-noise without frequency-dependent measurements and/or squeezing [75].
- [10] Wolfgramm F, Vitelli C, Beduini F A, Godbout N and Mitchell M W 2013 *Nat. Photon.* **7** 28
- [11] Taylor M A, Janousek J, Daria V, Knittel J, Hage B, Bachor H.-A and Bowen W P 2013 *Nat. Photon.* **7** 229
- [12] Taylor M A and Bowen W P 2016 *Phys. Rep.* **615** 1
- [13] Triginer Garces G, Chrzanowski H M, Daryanoosh S, Thiel V, Marchant A L, Patel R B, Humphreys P C, Datta A and Walmsley I A 2020 *Appl. Phys. Lett.* **117** 024002
- [14] Casacio C A, Madsen L S, Terrasson A, Waleed M, Barnscheidt K, Hage B, Taylor M A and Bowen W P 2021 *Nature* **594** 201
- [15] Xavier J, Yu D, Jones C, Zossimova E and Vollmer F 2021 *Nanophotonics* **10** 1387
- [16] Cole R 2014 *Cell Adhesion Migration* **8** 452
- [17] Demkowicz-Dobrzański R 2011 *Phys. Rev. A* **83** 061802
- [18] Rivas Á and Luis A 2012 *New J. Phys.* **14** 093052
- [19] Zhang Y R, Jin G R, Cao J P, Liu W M and Fan H 2012 *J. Phys. A: Math. Theor.* **46** 035302
- [20] Tsang M 2012 *Phys. Rev. Lett.* **108** 230401
- [21] Giovannetti V and Maccone L 2012 *Phys. Rev. Lett.* **108** 210404
- [22] Berry D W, Hall M J W, Zwierz M and Wiseman H M 2012 *Phys. Rev. A* **86** 053813
- [23] Pezzé L 2013 *Phys. Rev. A* **88** 060101

- [24] Hofmann H F 2009 *Phys. Rev. A* **79** 033822
- [25] Hyllus P, Pezzè L and Smerzi A 2010 *Phys. Rev. Lett.* **105** 120501
- [26] Pezzè L, Hyllus P and Smerzi A 2015 *Phys. Rev. A* **91** 032103
- [27] Mitchell M W, Lundeen J S and Steinberg A M 2004 *Nature* **429** 161
- [28] Giovannetti V, Lloyd S and Maccone L 2011 *Nat. Photon.* **5** 222
- [29] Holland M J and Burnett K 1993 *Phys. Rev. Lett.* **71** 1355
- [30] Lvovsky A I 2015 *Photonics* ed D L Andrews (New York: Wiley) pp 121–63
- [31] Lee C, Oh C, Jeong H, Rockstuhl C and Lee S.-Y 2019 *J. Phys. Commun.* **3** 115008
- [32] Kay S M 1998 *Fundamentals of Statistical Signal Processing: Estimation Theory* (Englewood Cliffs, NJ: Prentice-Hall)
- [33] Jarzyna M and Demkowicz-Dobrzański R 2015 *New J. Phys.* **17** 013010
- [34] Braunstein S L and Caves C M 1994 *Phys. Rev. Lett.* **72** 3439
- [35] Paris M G A 2009 *Int. J. Quantum Inf.* **07** 125
- [36] Lang M D and Caves C M 2013 *Phys. Rev. Lett.* **111** 173601
- [37] Genoni M G and Tufarelli T 2019 *J. Phys. A: Math. Theor.* **52** 434002
- [38] Jarzyna M and Demkowicz-Dobrzański R 2012 *Phys. Rev. A* **85** 011801
- [39] Leonhardt U and Paul H 1995 *Prog. Quantum Electron.* **19** 89
- [40] Monras A 2006 *Phys. Rev. A* **73** 033821
- [41] Loredò J C *et al* 2019 *Nat. Photon.* **13** 803
- [42] Luis A 2017 *Phys. Rev. A* **95** 032113
- [43] Rubio J and Dunningham J 2019 *New J. Phys.* **21** 043037
- [44] Hall M J W and Wiseman H M 2012 *Phys. Rev. X* **2** 041006
- [45] Hayashi M, Vinjanampathy S and Kwek L C 2018 *J. Phys. B: At. Mol. Opt. Phys.* **52** 015503
- [46] Rubio J and Dunningham J 2020 *Phys. Rev. A* **101** 032114
- [47] Tóth G and Petz D 2013 *Phys. Rev. A* **87** 032324
- [48] Alipour S and Reza khani A T 2015 *Phys. Rev. A* **91** 042104
- [49] Ng S, Ang S Z, Wheatley T A, Yonezawa H, Furusawa A, Huntington E H and Tsang M 2016 *Phys. Rev. A* **93** 042121
- [50] It is possible that these vacuum runs could in practice be beneficial, offering time for the sample to recover and avoid photobleaching effects, or actively detrimental, exacerbating drift effects or extending exposure to detrimental environmental or background noise processes. Ignoring any such noise processes and considering photons passing through the sample (rather than applying some model of absorption [67]) the vacuum state can be considered to have no effect on any sample.
- [51] While quantum metrology typically focuses on measurement strategies, quantum estimation à la Bayes can provide not only the optimal measurement and estimate for a given probe, as we do here, but also inform which initial probe is globally optimal [76].
- [52] Helstrom C W 1976 *Quantum Detection and Estimation Theory* (New York: Academic)
- [53] Personick S 1971 *IEEE Trans. Inf. Theory* **17** 240
- [54] Note that the form of equation (23) is still valid in multi-shot scenarios with collective measurements [43].
- [55] Li Y, Pezzè L, Gessner M, Ren Z, Li W and Smerzi A 2018 *Entropy* **20** 628
- [56] This follows from a linear expansion of $\rho(\varphi)$ around the prior average $\bar{\varphi} = \int_a^b d\varphi z(\varphi)\varphi$ as $\rho(\varphi) \approx \rho(\bar{\varphi}) + \frac{\partial \rho}{\partial \varphi} \Big|_{\bar{\varphi}} (\varphi - \bar{\varphi})$ which gives $\varrho \approx \rho(\bar{\varphi})$ and $\varrho' \approx \bar{\varphi} \rho'(\bar{\varphi}) + \sigma_0^2 \frac{\partial \rho}{\partial \varphi} \Big|_{\bar{\varphi}}$. Under this expansion we find $S = \bar{\varphi} + \sigma_0^2 L_{\varphi} \Big|_{\bar{\varphi}}$ with L_{φ} being the SLD, with which equation (23) becomes $\langle (\Delta\varphi)^2 \rangle_{\text{opt}} \simeq \sigma_0^2 [1 - \sigma_0^2 \mathcal{J}(\rho)]$ [43, appendix D].
- [57] Similar results have reported an exact equality for a Gaussian prior [76, 77].
- [58] Holevo A 2011 *Probabilistic and Statistical Aspects of Quantum Theory* (Berlin: Springer)
- [59] Rubio J, Knott P and Dunningham J 2018 *J. Phys. Commun.* **2** 015027
- [60] Braunstein S L, Lane A S and Caves C M 1992 *Phys. Rev. Lett.* **69** 2153
- [61] Górecki W, Demkowicz-Dobrzański R, Wiseman H M and Berry D W 2020 *Phys. Rev. Lett.* **124** 030501
- [62] Demkowicz-Dobrzański R, Dorner U, Smith B J, Lundeen J S, Wasilewski W, Banaszek K and Walmsley I A 2009 *Phys. Rev. A* **80** 013825
- [63] Kołodyński J and Demkowicz-Dobrzański R 2010 *Phys. Rev. A* **82** 053804
- [64] Escher B M, de Matos Filho R L and Davidovich L 2011 *Nat. Phys.* **7** 406
- [65] Demkowicz-Dobrzański R, Kołodyński J and Guţă M 2012 *Nat. Commun.* **3** 1063
- [66] Demkowicz-Dobrzański R, Banaszek K and Schnabel R 2013 *Phys. Rev. A* **88** 041802
- [67] Perarnau-Llobet M, Malz D and Cirac J I 2021 *Quantum* **5** 446
- [68] Ou Z Y 1997 *Phys. Rev. A* **55** 2598
- [69] Luis A 2004 *Phys. Lett. A* **329** 8
- [70] Gessner M, Pezzè L and Smerzi A 2018 *Phys. Rev. Lett.* **121** 130503
- [71] Loudon R and Knight P L 1987 *J. Mod. Opt.* **34** 709
- [72] Hall M J W, Berry D W, Zwiernik M and Wiseman H M 2012 *Phys. Rev. A* **85** 041802
- [73] Thomas-Peter N *et al* 2011 *New J. Phys.* **13** 055024
- [74] Datta A, Zhang L, Thomas-Peter N, Dorner U, Smith B J and Walmsley I A 2011 *Phys. Rev. A* **83** 063836
- [75] Kimble H J, Levin Y, Matsko A B, Thorne K S and Vyatchanin S P 2001 *Phys. Rev. D* **65** 022002
- [76] Macieszczak K, Fraas M and Demkowicz-Dobrzański R 2014 *New J. Phys.* **16** 113002
- [77] Demkowicz-Dobrzański R, Górecki W and Guţă M 2020 *J. Phys. A: Math. Theor.* **53** 363001