Random Bosonic States for Robust Quantum Metrology

M. Oszmaniec,[1] R. Augusiak,[1,2] C. Gogolin,[1,3] J. Kołodyński,[1] A. Acín,[1,4] and M. Lewenstein[1,4]

[1]*ICFO-Institut de Ciencies Fotoniques, The Barcelona Institute of Science and Technology, 08860 Castelldefels (Barcelona), Spain*[*]

[2]*Center for Theoretical Physics, Polish Academy of Sciences, Al. Lotników 32/46, 02-668 Warsaw, Poland*

[3]*Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Straße 1, 85748 Garching, Germany*

[4]*ICREA-Institució Catalana de Recerca i Estudis Avançats, Lluis Companys 23, 08010 Barcelona, Spain*

(Received 20 March 2016; revised manuscript received 17 August 2016; published 2 December 2016)

We study how useful random states are for quantum metrology, i.e., whether they surpass the classical limits imposed on precision in the canonical phase sensing scenario. First, we prove that random pure states drawn from the Hilbert space of distinguishable particles typically do not lead to superclassical scaling of precision even when allowing for local unitary optimization. Conversely, we show that random pure states from the symmetric subspace typically achieve the optimal Heisenberg scaling without the need for local unitary optimization. Surprisingly, the Heisenberg scaling is observed for random isospectral states of arbitrarily low purity and preserved under loss of a fixed number of particles. Moreover, we prove that for pure states, a standard photon-counting interferometric measurement suffices to typically achieve resolution following the Heisenberg scaling for all values of the phase at the same time. Finally, we demonstrate that metrologically useful states can be prepared with short random optical circuits generated from three types of beam splitters and a single nonlinear (Kerr-like) transformation.

DOI: 10.1103/PhysRevX.6.041044

Subject Areas: Interdisciplinary Physics, Quantum Physics, Quantum Information

I. INTRODUCTION

Quantum metrology opens the possibility to exploit quantum features to measure unknown physical quantities with accuracy surpassing the constraints dictated by classical physics [1–4]. Classically, by employing N probes to independently sense a parameter, the mean-squared error of estimation scales, at best, as $1/N$. This resolution is also known as the standard quantum limit (SQL) [5]. Quantum mechanics, however, allows one to engineer entangled states of N particles which, when used as probes, can lead to resolutions beyond the SQL. Crucially, in the canonical phase-sensing scenario, a precision scaling like $1/N^2$, known as the Heisenberg limit (HL) [6], can be attained. In practice, the destructive impact of noise must also be taken into account [7–9], but quantum-enhanced resolutions have been successfully observed in optical interferometry [10,11] (including gravitational-wave detection [12]), ultracold ion spectroscopy [13,14], atomic magnetometry [15,16], and in entanglement-assisted atomic clocks [17,18].

A fundamental question is to understand which quantum states offer an advantage for quantum metrology. Quantum-enhanced parameter sensitivity can only be observed with systems exhibiting interparticle entanglement [19]; thus, such enhanced sensitivity can be used to detect multipartite entanglement [20–23] and lower bound the number of particles being entangled [24,25]. However, the precise connection between entanglement and a quantum metrological advantage is not fully understood so far.

It is known that states achieving the optimal sensitivity must have entanglement between all their particles [25], like the Greenberger-Horne-Zeilinger (GHZ) state (equivalent to the N00N state in optical interferometry), yet there also exist classes of such states that are useless from the metrological perspective [26]. The optimal states, however, belong to the symmetric (bosonic) subspace, from which many states have been recognized to offer a significant advantage in quantum metrology [26–28]. On the other hand, a very weak form of entanglement—so-called undistillable entanglement—may lead to Heisenberg scaling [29], while any superclassical scaling arbitrarily close to the HL ($1/N^{2-\epsilon}$ with $\epsilon > 0$) can be achieved with states whose geometric measure of entanglement vanishes in the limit $N \to \infty$ [30].

Here, we go beyond merely presenting examples of states leading to quantum-enhanced precision. Instead, we conduct a systematic study by analyzing typical properties of the quantum and classical Fisher information on various ensembles of quantum states. First, we show that states of distinguishable particles typically are *not* useful for

[*]michal.oszmaniec@icfo.es

metrology, despite having a large amount of entanglement as measured by the entanglement entropy [31–34] and various other measures [35–38]. On the contrary, we show that most pure random states from the symmetric (bosonic) subspace of any local dimension achieve resolutions at the HL. Moreover, we prove that the usefulness of random symmetric states is robust against the loss of a finite number of particles, and it also holds for mixed states with fixed spectra (as long as the distance from the maximally mixed state in the symmetric subspace is sufficiently large). This result remain in stark contrast to the case of GHZ states, which completely lose their (otherwise ideal) phase sensitivity upon loss of just a single particle. Third, we show that, even for a fixed measurement, random pure bosonic states typically allow one to sense the phase at the HL. Concretely, this holds in the natural quantum-optics setting of photon-number detection in output modes of a balanced beam splitter [39] and independently of the value of the parameter. Finally, we demonstrate that states generated using random circuits with gates from a universal gate set on the symmetric subspace consisting only of beam splitters and a single nonlinear Kerr-like transformation also typically achieve Heisenberg scaling—again, even for a fixed measurement. Since all of our findings also equally apply to standard atomic interferometry [40–43], we hope that our work not only shows that metrological usefulness is a more generic feature than previously thought, but also opens up new possibilities for quantum-enhanced metrology based on random states.

Lastly, let us note that, as metrological usefulness of quantum states is tantamount to the notion of state macroscopicity [44], our results directly apply in this context [45,46]. Moreover, since states attaining HL can be used to approximately clone N quantum gates into as many as N^2 gates as $N \to \infty$, one can immediately use our findings to also infer that typical symmetric states provide a resource allowing for optimal asymptotic replication of unitary gates [47,48].

Our results are based on leveraging recent insights concerning the continuity of quantum Fisher information [30], measure concentration techniques [35,49–51], recently proven results about the spectral gap in the special unitary group [52], as well as the theory of approximate t-designs [53–55].

Our work sheds new light on the role of symmetric states in quantum metrology [26–28,46,56]. In particular, it clarifies the usefulness of symmetric states from the typicality perspective [26], but it also analytically confirms the findings about their typical properties previously suggested by numerical computations [46,56].

The remainder of the paper is organized as follows. In Sec. II, we introduce the setting of quantum parameter estimation, including the classical and quantum Fisher information and their operational interpretation. In Sec. III, we familiarize the reader with the technique of measure concentration and introduce some additional notation. In Sec. IV, we present our results on the lack of usefulness of random states of distinguishable particles for quantum metrology. Subsequently, in Sec. V, we show that states from the symmetric (bosonic) subspace typically attain the HL. Next, in Sec. VI, we analyze the robustness of metrological usefulness of such states under noise and loss of particles. In Sec. VII, we turn our attention to the classical Fisher information in a concrete measurement setup. We prove that with random pure symmetric states, Heisenberg scaling can typically be achieved with a physically accessible measurement setup—essentially, a Mach-Zehnder interferometer with particle-number detectors. Finally, in Sec. VIII, we demonstrate that symmetric states whose metrological properties effectively mimic those of random symmetric states can be prepared by short random circuits generated from a universal gate set in the symmetric subspace. We conclude our work in Sec. IX.

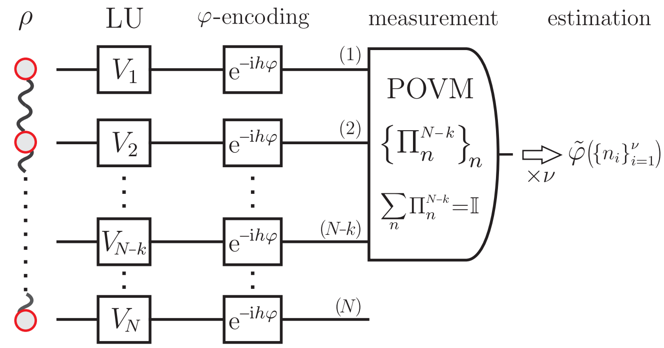

FIG. 1. Local-unitary (LU) optimized lossy quantum metrology protocol: A given state ρ is adjusted with LU operations $V_1, \ldots, V_N$ to probe, as precisely as possible, small fluctuations of the parameter φ, which is independently and unitarily encoded into each of the N constituent particles. Finally, only $N-k$ particles are measured, reflecting the possibility of losing k of them. The most general measurement is then described by a positive-operator-valued measure (POVM), i.e., a collection of positive semidefinite operators $\{\Pi_n^{N-k}\}_n$ that act on the remaining $N-k$ particles and sum up to the identity. The process is repeated ν times, in order to construct the most sensitive estimate of the parameter, $\tilde{\varphi}$, based on the measurement outcomes $\{n_i\}_{i=1}^{\nu}$.

II. QUANTUM METROLOGY

We consider the canonical phase-sensing scenario of quantum metrology [6], in which one is given N devices (black boxes) that encode a phaselike parameter φ. The task is to determine the optimal strategy of preparing quantum probes so that, after interaction with the devices, the probes can be measured in a way that reveals the highest sensitivity to fluctuations of the parameter φ. Crucially, in quantum mechanics, one has the freedom to apply the devices “in parallel” to a (possibly entangled) state ρ of N particles—see Fig. 1. The whole process is assumed to be repeated many

$\nu \gg 1$ times so that sufficient statical data are always guaranteed. Note that parameter sensing is the most "optimistic" metrology setting [57], in the sense that any type of general phase-estimation problem that also accounts for the nonperfect prior knowledge about the value of φ or finite size of the measurement data is bound to be more difficult [58].

Let $p_{n|\varphi}$ be the probability that (in any given round) the measurement will result in the outcome n given that the initial state is ρ and the parameter is φ. Then, the mean-squared error $\Delta^2\tilde{\varphi}$ of any unbiased and consistent estimator $\tilde{\varphi}$ of φ is lower bounded by the Cramér-Rao bound (CRB) [57]:

$$\Delta^2\tilde{\varphi} \geq \frac{1}{\nu F_{\mathrm{cl}}(\{p_{n|\varphi}\})}, \tag{1}$$

where

$$F_{\mathrm{cl}}(\{p_{n|\varphi}\}) \coloneqq \sum_n \frac{1}{p_{n|\varphi}} \left(\frac{\mathrm{d}p_{n|\varphi}}{\mathrm{d}\varphi}\right)^2 \tag{2}$$

is the classical Fisher information (FI). Importantly, in the phase-sensing scenario, the CRB (1) is guaranteed to be tight in the limit of large ν [57]. The classical Fisher information hence quantifies, in an operationally meaningful way, what the ultimate resolution to the fluctuations of φ is with a given measurement and state.

The quantum device is taken to act on a single particle (photon, atom, etc.) with Hilbert space $\mathcal{H}_{\mathrm{loc}}$ of finite dimension $d \coloneqq |\mathcal{H}_{\mathrm{loc}}|$ via a Hamiltonian h, which, without loss of generality, we assume to be traceless (any nonzero trace of h can always be incorporated into an irrelevant global phase factor). The device unitarily encodes the unknown parameter φ by performing the transformation $\mathrm{e}^{-\mathrm{i}h\varphi}$. The multiparticle state $\rho \in \mathcal{H}_N \coloneqq (\mathcal{H}_{\mathrm{loc}})^{\otimes N}$ moves along the trajectory

$$\varphi \mapsto (\mathrm{e}^{-\mathrm{i}h\varphi})^{\otimes N} \rho (\mathrm{e}^{\mathrm{i}h\varphi})^{\otimes N}, \tag{3}$$

corresponding to the unitary evolution with the global Hamiltonian

$$H \coloneqq H_N \coloneqq \sum_{j=1}^{N} h^{(j)}. \tag{4}$$

The measurement is defined by a POVM, $\{\Pi_n^N\}_n$, acting on the whole system (or, while accounting for particle losses, only on the remaining $N-k$ particles—see Fig. 1) and satisfying $\sum_n \Pi_n^N = \mathbb{1}$. It yields outcome n with probability

$$p_{n|\varphi} = \mathrm{tr}(\Pi_n^N \mathrm{e}^{-\mathrm{i}H\varphi} \rho \mathrm{e}^{\mathrm{i}H\varphi}). \tag{5}$$

In the seminal work, Braunstein and Caves [59] showed how to quantify the maximal usefulness of a state ρ in the above scenario by maximizing the classical Fisher information (2) over all possible POVMs. The resulting quantity is called quantum Fisher information (QFI). It depends solely on the quantum state ρ being considered and the Hamiltonian H responsible for the parameter encoding, and we denote it by $F(\rho, H)$. A QFI scaling faster than linear with N (for fixed local Hamiltonian h) ultimately leads to superclassical resolutions by virtue of the CRB (1). Since the Hamiltonian (4) is local and parameter independent, the ultimate HL is unambiguously given by $F_{\mathrm{cl}} \propto N^2$, with superclassical scaling being possible solely due to the entanglement properties of ρ and not due to a nonlocal or nonlinear parameter dependence [60–62]. Importantly, thanks to the unitary character of the parameter encoding (3), resolutions that scale beyond SQL can indeed be attained in metrology [8,9].

Although in the phase-sensing scenario the optimal measurement can be designed for a particular value of φ, it is often enough to know, prior to the estimation, that the parameter lies within a sufficiently narrow window of its potential values, as then there exist a sequence of measurements that eventually—in the limit of many protocol repetitions ($\nu \to \infty$ in Fig. 1) with measurements adaptively adjusted—yields a classical Fisher information that still achieves the QFI [63]. Crucially, the optimal scaling of the QFI achievable in the phase-sensing scenario is proportional to N^2, which proves the Heisenberg scaling to indeed be the ultimate one.

For the sake of having concise terminology, we call a family of states for increasing N "useful" for quantum sensing if there exists a Hamiltonian h in Eq. (3) such that the corresponding QFI scales faster than N [i.e., $F(\rho, H) \notin \mathcal{O}(N)$] in the limit of large N. In contrast, we say that the family of states is "not useful" for quantum sensing (and hence also for all less "optimistic" metrological scenarios) if its QFI scales asymptotically, at most, like N [i.e., $F(\rho, H) \in \mathcal{O}(N)$]. We adopt the above nomenclature for the sake of brevity and concreteness. However, let us stress that states reaching beyond SQL, despite not preserving superclassical precision scaling, may also yield dramatic precision enhancement (e.g., squeezed states in gravitational detectors [12]), which, in fact, then guarantees their rich entanglement structure [20–25]. Nevertheless, it is the superclassical precision scaling that manifests the necessary entanglement properties to be maintained with the system size. In particular, its protection at the level of QFI has recently allowed scientists to design novel noise-robust metrology protocols [64,65].

In the remainder of this section, we give a mathematical definition of the QFI and discuss some of its properties. The QFI has an elegant geometric interpretation [59] as the "square of the speed" along the trajectory (3) measured with respect to the Bures distance $d_B(\rho, \sigma) \coloneqq \sqrt{2[1 - \mathcal{F}(\rho, \sigma)]}$, where $\mathcal{F}(\rho, \sigma) \coloneqq \mathrm{tr}\sqrt{\rho^{1/2}\sigma\rho^{1/2}}$ is the fidelity. This allows one to define the QFI geometrically [66]:

$$F(\rho, H) := [\lim_{\varphi\to 0} 2 d_B(\rho, e^{-iH\varphi}\rho e^{iH\varphi})/|\varphi|]^2. \quad (6)$$

This geometric interpretation of QFI is key for the derivation of the following results. Using the spectral decomposition of a quantum state $\rho = \sum_{j=1}^{d^N} p_j |e_j\rangle\langle e_j|$, with $p_j \geq 0$ denoting its eigenvalues, we can write the QFI more explicitly as [2,3]

$$F(\rho, H) = 2 \sum_{j,k:\, p_j+p_k\neq 0} \frac{(p_j - p_k)^2}{p_j + p_k} |\langle e_j|H|e_k\rangle|^2, \quad (7)$$

which for pure states, $\rho = \psi$, simplifies further to the variance $\Delta^2 H|_\psi$ of H, that is,

$$F(\psi, H) = 4[\mathrm{tr}(\psi H^2) - \mathrm{tr}(\psi H)^2] =: 4\Delta^2 H|_\psi. \quad (8)$$

Let us also recall that the QFI is a convex function on the space of quantum states, which shows that mixing states can never increase their parameter sensitivity above their average sensitivity. This, together with the fact that the QFI is also additive on product states [2,3], directly proves that only entangled states can lead to resolutions beyond the SQL.

III. CONCENTRATION OF MEASURE PHENOMENON

Before presenting our main results, we briefly discuss a key ingredient for their proofs—the concentration of measure phenomenon [49–51]. For a more detailed discussion, we point the reader to Appendix A. For any finite-dimensional Hilbert space, we denote by $\mu(\mathcal{H})$ the Haar measure on the special unitary group $\mathrm{SU}(\mathcal{H})$. The Haar measure can be thought of as the uniform distribution over unitary transformations. We denote by $\mathrm{Pr}_{U\sim\mu(\mathcal{H})}(A(U))$ the probability that a statement A holds for unitaries U drawn from the measure $\mu(\mathcal{H})$ and by

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{H})} f(U) := \int_{\mathrm{SU}(\mathcal{H})} d\mu(U) f(U) \quad (9)$$

the expectation value of a function $f\colon \mathrm{SU}(\mathcal{H}) \to \mathbb{R}$. Our findings concern the typical value of such functions. For example, $f(U)$ could be the QFI of some family of so-called isospectral states, i.e., states of the form $U\rho U^\dagger$, with ρ some fixed state on $\mathcal{H}$ and $U \sim \mu(\mathcal{H})$ a unitary drawn from the Haar measure on $\mathrm{SU}(\mathcal{H})$ [37]. Note that as $F(U\rho U^\dagger, H) = F(\rho, U^\dagger H U)$ [this follows directly from Eq. (7)], all our results can also be interpreted as being about random Hamiltonians instead of random states.

To show that for almost all U the value of such a function is close to the typical value and that this typical value is close to the average, we repeatedly employ the following concentration of measure inequality [50]:

$$\mathop{\mathrm{Pr}}_{U\sim\mu(\mathcal{H})}\left(\left|f(U) - \mathop{\mathbb{E}}_{U\sim\mu(\mathcal{H})} f\right| \geq \epsilon\right) \leq 2\exp\left(-\frac{|\mathcal{H}|\epsilon^2}{4L^2}\right). \quad (10)$$

It holds for every $\epsilon \geq 0$ and every function $f\colon \mathrm{SU}(\mathcal{H}) \to \mathbb{R}$ that is Lipschitz continuous [with respect to the geodesic distance on $\mathrm{SU}(\mathcal{H})$] and thus possesses its corresponding Lipschitz constant L. Recall that the Lipschitz constant gives the bound on how fast the value of a function can change under a change of its argument. For a formal definition of L, see Eq. (A5) in Appendix A, where we explicitly prove bounds on Lipschitz constants of all the functions relevant for our considerations. Here, let us only note that as both the FI (2) and the QFI (6) are nontrivial (in particular, nonlinear) functions of quantum states, we need to resort to advanced techniques of differential geometry.

Before we move on to our results, we introduce a minimal amount of additional notation: Given two functions f, g, we write $f(N) \in \Theta(g(N))$ if both functions have the same behavior in the limit of large N (up to a positive multiplicative constant) and write $f(N) \in \mathcal{O}(g(N))$ if there exists a constant C such that $f(N) \leq C g(N)$ in the limit of large N. Slightly abusing notation, we sometimes also use the symbols $\Theta(f(N))$ and $\mathcal{O}(f(N))$ to denote an arbitrary function with the same asymptotic behavior as f. For any operator X, we denote its operator norm by

$$\|X\| := \sup_{|\psi\rangle\neq 0} \frac{\|X|\psi\rangle\|}{\||\psi\rangle\|}, \quad (11)$$

where $\||\psi\rangle\| = \sqrt{\langle\psi|\psi\rangle}$ stands for the standard vector norm. Then, the trace and the Hilbert-Schmidt norms of X are defined as $\|X\|_1 := \mathrm{tr}\sqrt{X^\dagger X}$ and $\|X\|_{\mathrm{HS}} := \sqrt{\mathrm{tr}(X^\dagger X)}$, respectively. These generally obey the relation $\|X\| \leq \|X\|_{\mathrm{HS}} \leq \|X\|_1$.

IV. FUTILITY OF GENERAL RANDOM STATES

First, we show that Haar-random isospectral states of distinguishable particles are typically not useful for quantum metrology even if they are pure and hence typically highly entangled [31–33,35–38]. This remains true even if one allows for LU optimization before the parameter is encoded (see Fig. 1). Note that in the special case $d = 2$, the LU optimization of the input state is equivalent to an optimization over all unitary parameter encodings. The maximal achievable QFI with LU optimization is given by

$$F^{\mathrm{LU}}(\rho, H) := \sup_{V\in\mathrm{LU}} F(V\rho V^\dagger, H), \quad (12)$$

where LU denotes the local unitary group, i.e., a group consisting of unitaries of the form $V = V_1 \otimes V_2 \otimes \ldots \otimes V_N$, with V_j acting on the jth particle of the system (see Fig. 1). Despite the fact that for other states this sometimes boosts their QFI [26], we have the following no-go theorem for random states from the full Hilbert space $\mathcal{H}_N$ of distinguishable particles:

Theorem 1. (Most random states of distinguishable particles are not useful for metrology, even after LU optimization.) Fix a local dimension d, single-particle Hamiltonian h, and a pure state ψ_N on $\mathcal{H}_N$. Let $F^{\mathrm{LU}}(U) \coloneqq F^{\mathrm{LU}}(U\psi_N U^\dagger, H)$; then,

$$\Pr_{U\sim\mu(\mathcal{H}_N)}(F^{\mathrm{LU}}(U) \notin \Theta(N)) \le \exp\left(-\Theta\left(\frac{d^N}{N^2}\right)\right). \quad (13)$$

Proof sketch.—From Eq. (8), we have that $F(\psi, H) \le 4\mathrm{tr}(\psi H)$, which implies

$$\begin{aligned} &F^{\mathrm{LU}}(U\psi_N U^\dagger, H) \\ &\quad \le 4 \sup_{V\in\mathrm{LU}} \mathrm{tr}(U\psi_N U^\dagger V H^2 V^\dagger) \\ &\quad = 4 \sup_{V_1,\ldots,V_N} \sum_{j,k} \mathrm{tr}(U\psi_N U^\dagger V_j h^{(j)} V_j^\dagger V_k h^{(k)} V_k^\dagger). \quad (14)\end{aligned}$$

The terms with $j = k$ give a contribution of at most $4N\|h\|^2$. In the remaining terms, however, the operator $V_j h^{(j)} V_j^\dagger V_k h^{(k)} V_k^\dagger$ is traceless, so

$$\begin{aligned} &\sup_{V_1,\ldots,V_N} \sum_{j\ne k} \mathrm{tr}(U\psi_N U^\dagger V_j h^{(j)} V_j^\dagger V_k h^{(k)} V_k^\dagger) \\ &\qquad \le \|h\|^2 \sum_{j\ne k} \|\mathrm{tr}_{\neg j,k}(U\psi_N U^\dagger) - \mathbb{1}/d^2\|_1. \quad (15)\end{aligned}$$

But, the average of $\|\mathrm{tr}_{\neg j,k}(U\psi_N U^\dagger) - \mathbb{1}/d^2\|_1$ can be upper bounded exponentially [67] as

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{H}_N)} \|\mathrm{tr}_{\neg j,k}(U\psi_N U^\dagger) - \mathbb{1}/d^2\|_1 \le \frac{d^2}{\sqrt{d^N}}, \quad (16)$$

so

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{H}_N)} F^{\mathrm{LU}}(U) \le 4N\|h\|^2\left[1 + \frac{(N-1)d^2}{\sqrt{d^N}}\right]. \quad (17)$$

Conversely, a direct computation of the average non-LU-optimized QFI yields a lower bound of order $N\|h\|^2$. Application of a concentration inequality of the type given in Eq. (10) yields the claimed result (see Appendix D for details). ■

Because of the convexity of QFI, the typical behavior of the QFI on any unitary-invariant ensembles [66] of mixed density matrices in $\mathcal{H}_N$ can only be worse than that of pure states predicted by the above theorem. Furthermore, a bound similar to Eq. (17) can also be derived for Hamiltonians H with few-body terms, for example, those with finite or short-range interactions on regular lattices or those considered in Ref. [60]. Lastly, let us remark that, as we consider the most optimistic phase-sensing protocol, Theorem 1 also disproves the possibility of superclassical scalings of precision when considering random states in any general phase-estimation protocol [58], e.g., the Bayesian single-shot scenarios [68–70].

The above proof relies on the fact that most random states on $\mathcal{H}_N$ have very mixed two-particle marginals. Thus, high entanglement entropy is enough to make random pure states on $\mathcal{H}_N$ useless for quantum metrology. Complementing this, in Ref. [30], it has been proven that a nonvanishing geometric measure of entanglement $E_g(\psi) \in \Theta(1)$ is, at the same time, necessary for Heisenberg scaling [recall that the geometric measure of entanglement for a pure state ψ is defined as $E_g(\psi) \coloneqq 1 - \max_{\sigma\in\mathrm{SEP}} \mathrm{tr}(\psi\sigma)$, where SEP denotes the set of separable states in $\mathcal{D}(\mathcal{H}_N)$]. Interestingly, pure random states of N particles do typically satisfy $E_g(\psi) \approx 1$ [36]. This result, together with Theorem 1, shows that, contrary to a recent conjecture of Ref. [71], a high geometric measure of entanglement is not sufficient for states to exhibit superclassical precision scaling in quantum metrology. However, let us note that this is consistent with the numerical findings of Ref. [46]. The idea that the presence or absence of superclassical scaling of the QFI arises solely from the two-particle marginals has also recently been noted in Ref. [72].

V. USEFULNESS OF RANDOM SYMMETRIC STATES

We now turn to the study of random states from the symmetric (bosonic) subspace of N qudits, $\mathcal{S}_N \coloneqq \mathrm{span}\{|\psi\rangle^{\otimes N} : |\psi\rangle \in \mathcal{H}_{\mathrm{loc}}\}$, which is of dimension $|\mathcal{S}_N| = \binom{N+d-1}{N} \in \Theta(N^{d-1})$. This subspace of states contains metrologically useful states such as the GHZ state or the Dicke state [27] and naturally appears in experimental setups employing photons and bosonic atoms [3]. For the special case $d = 2$, it was proven in [26] that, with LU optimization, almost all pure symmetric states exhibit $F^{\mathrm{LU}} > 4N\|h\|^2$.

In what follows, we consider random isospectral symmetric states, i.e., states of the form $U\sigma_N U^\dagger$, with σ_N being a fixed state on $\mathcal{S}_N$ and $U \sim \mu(\mathcal{S}_N)$. By $\sigma_{\mathrm{mix}} = \mathbb{1}_{\mathcal{S}_N}/|\mathcal{S}_N|$, we denote the maximally mixed state in $\mathcal{S}_N$. In particular, we prove that such symmetric states typically achieve a Heisenberg-like scaling, provided that the spectrum of σ_N differs sufficiently from the spectrum of σ_{mix}. Interestingly, this holds even without LU optimization:

Theorem 2. (Most random isospectral symmetric states are useful for quantum sensing.) Fix a single-particle Hamiltonian h, local dimension d, and a state σ_N from the symmetric subspace $\mathcal{S}_N$ with eigenvalues $\{p_j\}_j$. Let $F(U) \coloneqq F(U\sigma_N U^\dagger, H)$; then,

$$\begin{aligned} &\Pr_{U\sim\mu(\mathcal{S}_N)}(F(U) < d_B(\sigma_N, \sigma_{\mathrm{mix}})^2\Theta(N^2)) \\ &\qquad \le \exp\left(-d_B(\sigma_N, \sigma_{\mathrm{mix}})^3\Theta(N^{d-1})\right). \quad (18)\end{aligned}$$

Proof sketch.—We use the standard integration techniques (see Appendix C for details) on the unitary group to show that

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)} F(U) = \mathrm{tr}(h^2)\mathcal{G}(N,d)\Lambda(\{p_j\}_j), \quad (19)$$

where

$$\mathcal{G}(N,d) \coloneqq \frac{4N(N+d)}{d(d+1)}\frac{|\mathcal{S}_N|}{|\mathcal{S}_N|+1}, \quad (20)$$

and $\Lambda(\{p_j\}_j)$ is a function of the eigenvalues $\{p_j\}_j$ which for pure states attains $\Lambda = 1$. Therefore, for the case of pure states, we have

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)} F(U\psi_N U^\dagger, H) = \mathrm{tr}(h^2)\mathcal{G}(N,d), \quad (21)$$

where ψ_N is an arbitrary pure state on $\mathcal{S}_N$. From this result, it clearly follows that the average QFI of random pure symmetric states exhibits Heisenberg scaling in the limit $N\to\infty$. Moreover, it turns out that $\Lambda(\{p_j\}_j)$ satisfies the inequality

$$\Lambda(\{p_j\}_j) \geq \frac{|\mathcal{S}_N|}{|\mathcal{S}_N|-1}\frac{d_B(\sigma_N,\sigma_{\mathrm{mix}})^2}{2}. \quad (22)$$

The inequality (18) now follows from concentration of measure inequalities on $\mathrm{SU}(\mathcal{S}_N)$ (see Appendix D for a detailed proof). ■

As the Bures distance to the maximally mixed state $d_B(\sigma_N,\sigma_{\mathrm{mix}})$ enters the theorem in a nontrivial way, we illustrate the power of the result by showing that even states that asymptotically move arbitrarily close to σ_{mix} still typically achieve a superclassical scaling:

Corollary 1. (Sufficient condition for usefulness of random isospectral symmetric states) Let $U\sigma_N U^\dagger$ be an ensemble of isospectral states from the symmetric subspace $\mathcal{S}_N$ with eigenvalues $\{p_j\}_j$. Theorem 2 implies that random states drawn from such an ensemble are typically useful for sensing, as for any $\alpha < \min\{1/2,(d-1)/3\}$, they yield a precision scaling $1/N^{2(1-\alpha)}$ provided that $d_B(\sigma_N,\sigma_{\mathrm{mix}}) \geq 1/N^\alpha$.

Let us remark that Theorem 2 constitutes a fairly complete description of the typical properties of QFI on various ensembles of isospectral density matrices. Typical properties of entanglement and its generalizations on sets of isospectral density matrices were analyzed in Ref. [37].

VI. ROBUSTNESS TO IMPERFECTIONS

Next, we underline the practical importance of the above results by showing that the usefulness of random symmetric states is robust against dephasing noise and particle loss.

A. Depolarizing noise

We first investigate how mixed σ_N may become while still providing a quantum advantage for metrology. To this aim, we consider a concrete ensemble of depolarized states:

Example 1 [Depolarized random symmetric states] Fix a local dimension d, single-particle Hamiltonian h, and $p\in[0,1]$. Let ψ_N be a pure state on $\mathcal{S}_N$, and set

$$\sigma_N(p) = (1-p)\psi_N + p\sigma_{\mathrm{mix}}. \quad (23)$$

Let $F_p(U) \coloneqq F(U\sigma_N(p)U^\dagger, H)$; then, for every $\epsilon > 0$,

$$\Pr_{U\sim\mu(\mathcal{S}_N)}(|F_p(U) - \mathbb{E}F_p| \geq \epsilon F_p) \leq \exp(-\epsilon^2\Theta(N^{d-1})), \quad (24)$$

where $\mathbb{E}F_p \coloneqq \mathbb{E}_{U\sim\mu(\mathcal{S}_N)}F_p$ is given by Eq. (19) with

$$\Lambda = \Lambda_p \coloneqq \frac{(1-p)^2}{1-p+2p/|\mathcal{S}_N|}. \quad (25)$$

The above example shows that for all values of $p<1$, the Heisenberg scaling of the QFI is typically retained. The QFI then still concentrates around its average, which is of order N^2. Moreover, we observe that for large N, the average value of QFI of random symmetric depolarized states decreases essentially linearly with p as $|\Lambda_p-(1-p)| \leq 2/|\mathcal{S}_N|$. Finally, Eq. (24) is a large deviation inequality for QFI on the ensemble of depolarized pure symmetric states, with the mean $\mathbb{E}F_p$. The average $\mathbb{E}F_p$ for the special case $p=0$ is given by Eq. (21).

B. Finite loss of particles

Next, we investigate whether the Heisenberg scaling of random symmetric states is robust under the loss of particles. We model the particle loss by the partial trace over $k\leq N$ particles, i.e., $\sigma_N \mapsto \mathrm{tr}_k(\sigma_N)$ for a given state σ_N. Note that because of the permutation symmetry of state σ_N considered, it does not matter which particles are lost. In particular, such a mechanism is equivalent to the situation in which one is capable of measuring only (as if distinguishable) $N-k$ of the particles. We are therefore interested in typical properties of $F(\mathrm{tr}_k(U\sigma_N U^\dagger), H_{N-k})$, where $H_{N-k} = \sum_{j=1}^{N-k} h^{(j)}$ and $U\sigma_N U^\dagger$ is a random isospectral state on $\mathcal{S}_N$.

For comparison, let us recall that the GHZ state, which is known to be optimal in quantum sensing [5], becomes completely useless upon the loss of just a single particle, as the remaining reduced state is separable. In contrast, sufficiently pure random bosonic states typically remain useful for sensing even when a constant number of particles has been lost. Even the Heisenberg scaling $\sim 1/N^2$ of the QFI is preserved:

Theorem 3. (Random isospectral symmetric states are typically useful under finite particle loss.) Fix a local dimension d, single-particle Hamiltonian h, and a state σ_N on $\mathcal{S}_N$ with eigenvalues $\{p_j\}_j$. Let $F_k(U) \coloneqq F(\mathrm{tr}_k(U\sigma_N U^\dagger), H_{N-k})$; then,

$$\Pr_{U\sim\mu(\mathcal{S}_N)}\left(F_k(U) < \|\sigma_N - \sigma_{\mathrm{mix}}\|_{\mathrm{HS}}^2 \Theta\left(\frac{N^2}{k^d}\right)\right) \leq \exp\left(-\|\sigma_N - \sigma_{\mathrm{mix}}\|_{\mathrm{HS}}^4 \Theta\left(\frac{N^{d-1}}{k^{2d}}\right)\right). \quad (26)$$

The main difficulty in the proof of the above theorem is that the random matrix ensemble induced by the partial trace of random isospectral states from $\mathcal{S}_N$ is not well studied. Hence, we cannot use standard techniques to compute the average of $F_k(U)$. We circumvent this problem by lower bounding the QFI by the asymmetry measure [73] and then the HS norm

$$F(\rho, H) \geq \|\rho, H\|_1^2 \geq \|[\rho, H]\|_{\mathrm{HS}}^2 \quad (27)$$

and then computing the average of the right-hand side instead. We provide all the technical details of the proof in Appendix D; here, we only consider the two-mode case as an example:

Example 2. For $d = 2$ and after fixing $\mathrm{tr}(h^2) = 1/2$, the following inequality holds:

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)} F(\mathrm{tr}_k(U\sigma_N U^\dagger), H_{N-k}) \geq \frac{1}{3}\frac{(N-k)(N+1)}{(k+1)(k+2)}\frac{(N+1)\mathrm{tr}\rho_N^2 - 1}{N} \quad (28)$$

(see Lemma 9 in Appendix C), which, for the pure states, further simplifies to

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)} F(\mathrm{tr}_k(U\psi_N U^\dagger), H_{N-k}) \geq \frac{1}{3}\frac{(N-k)(N+1)}{(k+1)(k+2)}. \quad (29)$$

It is interesting to note that without particle losses, i.e., $k = 0$, this formula gives

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)} F(U\psi_N U^\dagger, H_N) \geq \frac{1}{6}N(N+1), \quad (30)$$

a result that differs from the exact expression (21) only by a factor of $1/2$.

The above result is most relevant for atomic interferometry experiments [40–43], in which unit detection efficiencies can be achieved, and it is hence reasonably possible to limit the loss of particles to a small number. In contrast, current optical implementations are limited by inefficiencies of photon detectors that are adequately modeled with a fictitious beam-splitter model [3]. It is known in noisy quantum metrology that generic uncorrelated noises (in particular, the noise described by the beam-splitter model) constrain the ultimate precision to a constant factor beyond the SQL [8,9]. The beam splitter effectively fixes the loss-rate per particle, allowing *all* the particles to be lost with some probability. In fact, modeling losses with a beam splitter is equivalent to tracing out k particles with probability $\binom{N}{k}(1-\eta)^k\eta^{N-k}$, where η is the fictitious beam-splitter transitivity (see Appendix E for the proof). Thus, the number k of particles lost fluctuates according to a binomial distribution. As a result, the lower bound on the average QFI utilized in Sec. VI B must also be averaged over the fluctuations of k, and the superclassical scaling is lost. Our results on the robustness of random bosonic states against *finite* particle losses are hence fully consistent with the no-go theorems of Refs. [8,9]. Moreover, the fact that random states are much more robust against particle loss than, for example, N00N states, which lose their metrological usefulness upon the loss of a single particle, raises the hope that for finite N, they might still perform comparably well even under uncorrelated noise.

VII. ATTAINING THE HEISENBERG LIMIT WITH A SIMPLE MEASUREMENT

We have demonstrated that random bosonic states lead, in a robust manner, to superclassical scaling of the QFI. This proves that, in principle, they must allow one to locally sense the phase around any value with resolutions beyond the SQL. However, as previously explained in Sec. II, the phase-sensing scenario allows the measurement to be optimized for the particular parameter value considered. Moreover, such a measurement may also strongly depend on the state utilized in the protocol, so one may question whether it could be potentially implemented in a realistic experiment, as it then must theoretically be adjusted depending on the state drawn at random. Thus, it is *a priori* not clear if metrological usefulness of random symmetric states can actually be exploited in practice. Here, we show that this is indeed the case. For random symmetric states of two-mode bosons, a standard measurement in optical and atomic interferometry suffices to attain the Heisenberg scaling of precision when sensing the phase around any value.

In particular, we consider the detection of the distribution of the N bosons between two modes (interferometer arms) after a balanced beam-splitter transformation [3]. As depicted in Fig. 2, this corresponds, in optics, to the photon-number detection at two output ports of a Mach-Zehnder (MZ) interferometer [39]. Yet, such a setup also applies to experiments with atoms in double-well potentials, in which the beam-splitter transformation can be implemented via trap-engineering and atomic interactions [40,41], while number-resolving detection has recently been achieved via cavity-coupling [42] and fluorescence [43].

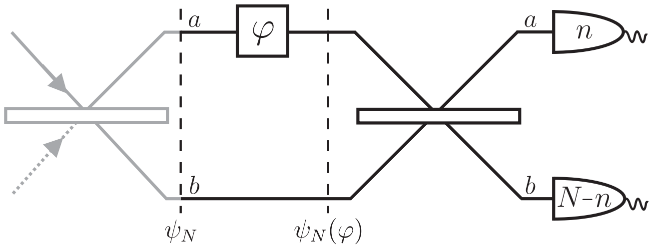

FIG. 2. Mach-Zehnder interferometer: We consider a bosonic pure state ψ_N of N photons in modes a and b inside the interferometer. The estimated phase φ is acquired because of the path difference in the arms. After a balanced beam splitter, the number of photons n and $N-n$ are measured in arms a and b, respectively.

One can directly relate the general protocol of distinguishable qubits (see Fig. 1) to the optical setup of photons in two modes (see Fig. 2) after acknowledging that the Dicke basis of general pure symmetric qubit states is nothing but their two-mode picture, in which a qubit in a state $|0\rangle$ ($|1\rangle$) describes a photon traveling in arm a (b) of the interferometer. In particular, a general pure bosonic state of N qubits can then be written as a superposition of Dicke states $\{|n, N-n\rangle\}_{n=0}^{N}$, where each $|n, N-n\rangle$ represents the situation in which n and $N-n$ photons are traveling in arm a and b, respectively. In Fig. 2, the estimated phase φ is acquired in between the interferometer arms, i.e., by the transformation $\exp(-\mathrm{i}\hat{J}_z\varphi)$ [74]. In the qubit picture, this corresponds to a single-particle unitary $\exp(-\mathrm{i}h\varphi)$ with $h=\sigma_z/2$. Moreover, the unitary balanced beam-splitter transformation of Fig. 2, commonly defined in the modal picture as $\hat{B} \coloneqq \exp(-\mathrm{i}\pi\hat{J}_x/2)$, is then equivalent to a local rotation $\exp(-\mathrm{i}\pi\sigma_x/4)$ of each particle in the qubit picture applied after the φ encoding.

Hence, the measurement of Fig. 2 with outcomes labeled by n (the number of photons detected in mode a) corresponds to a POVM $\{\Pi_n^N\}_{n=0}^N$ with elements $\Pi_n^N = \hat{B}^\dagger D_n^N \hat{B}$, where D_n^N are the projections onto Dicke states, $|D_n^N\rangle \coloneqq |n, N-n\rangle$. Given a general pure state ψ_N inside the interferometer (see Fig. 2), the state after acquiring the estimated phase reads $\psi_N(\varphi) \coloneqq \mathrm{e}^{-\mathrm{i}\hat{J}_z\varphi}\psi_N \mathrm{e}^{\mathrm{i}\hat{J}_z\varphi}$. Then, the probability of outcome n given that the unknown parameter was φ is just

$$p_{n|\varphi}(\psi_N) = \mathrm{tr}[\Pi_n^N \psi_N(\varphi)]. \tag{31}$$

Before we proceed, let us note that because of the identity $\hat{B}\mathrm{e}^{-\mathrm{i}\hat{J}_z\varphi}\hat{B}^\dagger = \mathrm{e}^{\mathrm{i}\hat{J}_y\varphi}$, it is possible to effectively map the above measurement scheme to the situation in which the initial state is already propagated through a beam splitter, $\tilde{\psi}_N = \hat{B}\psi_N\hat{B}^\dagger$, yet the parameter is encoded via a Hamiltonian in the y direction (via $\tilde{h} = -\sigma_y/2$). As a result, the measurement (POVM) elements then simplify to just projections onto the Dicke states D_n^N (see Appendix C 4 for details).

Having the explicit form of the measurement-outcome probability, we can compute the corresponding (classical) FI

$$F_{\mathrm{cl}}(\{p_{n|\varphi}(\psi_N)\}) = \sum_{n=0}^{N} \frac{\mathrm{tr}(\mathrm{i}[\Pi_n^N, \hat{J}_z]\psi_N(\varphi))^2}{\mathrm{tr}[\Pi_n^N \psi_N(\varphi)]}. \tag{32}$$

The unitary rotation $\mathrm{e}^{-\mathrm{i}\hat{J}_z\varphi}$ entering the definition of $\psi_N(\varphi)$ is responsible for the strong dependence (see also the numeric results in Sec. VIII) of the FI on the value of φ. However, let us note that when averaging (with respect to the Haar measure) the FI (32) over all bosonic states $\psi_N \in \mathcal{S}_N$, any unitary transformation of the state becomes irrelevant. In particular, owing to the parameter being unitary encoded, $\psi_N(\varphi)$ can then simply be replaced by ψ_N, so the average of Eq. (32) manifestly ceases to depend on φ.

This observation alone is not sufficient to deduce that the concentration behavior of FI (32) is independent of the value of the parameter φ. However, with the following theorem, we show not only this but actually a significantly stronger statement. We prove that the FI given in Eq. (32) evaluated on random symmetric states not only typically attains the Heisenberg scaling for a certain value of φ but typically does so for all values of φ *at the same time*.

Theorem 4. (Pure random symmetric qubit states typically attain the HL for all values of φ in the setup of Fig. 2.) Let ψ_N be a fixed pure state on $\mathcal{S}_N$ for $d=2$ modes and $p_{n|\varphi}(U\psi_N U^\dagger)$ the probability to obtain outcome n given that the value of the unknown phase parameter is φ and the interferometer state is $U\psi_N U^\dagger$ [see also Eq. (31)]. Let $F_{\mathrm{cl}}(U,\varphi) \coloneqq F_{\mathrm{cl}}(\{p_{n|\varphi}(U\psi_N U^\dagger)\})$ be the corresponding FI defined in Eq. (32); then,

$$\Pr_{U\sim\mu(\mathcal{S}_N)}(\exists_{\varphi\in[0,2\pi]} F_{\mathrm{cl}}(U,\varphi) \le \Theta(N^2)) \le \exp(-\Theta(N)). \tag{33}$$

In other words, the probability that for a random state there exists a value of the parameter φ for which $F_{\mathrm{cl}}(U,\varphi)$ does not achieve Heisenberg scaling is exponentially small in N. Hence, when dealing with typical two-mode bosonic states, one does not have to resort to LU optimization (similarly, as in Theorem 2) in order to reveal their metrological usefulness *even for a fixed measurement*. Note that such an optimization would make the problem φ independent. As the interferometric scheme of Fig. 2 is restricted to the symmetric subspace $\mathcal{S}_N$, one is only allowed to perform LU operations of the form $V^{\otimes N}$. However, setting $V = \exp(-\mathrm{i}\theta\sigma_z/2)$, one can then always shift $\varphi \to \varphi + \theta$ to any desired value.

We provide a detailed proof of Theorem 4 in Appendix D, where we also present its more precise and technical version. One of its constituents—the analysis of the average value of the FI (32)—can be found in Appendix C 4, where, by rigorously showing that

$$c_-N^2 \leq \underset{U\sim\mu(\mathcal{S}_N)}{\mathbb{E}} F_{\mathrm{cl}}(U,\varphi) \leq c_+N^2 + N, \qquad (34)$$

with constants $0 < c_- < c_+ < \infty$, we prove that the average FI indeed asymptotically follows the HL-like scaling. Although our derivation only allows us to bound the actual asymptotic constant factor, we conjecture that $\mathbb{E}_{U\sim\mu(\mathcal{S}_N)}F_{\mathrm{cl}}(U,\varphi) \to N^2/6$. In particular, we realize that such behavior is recovered after replacing the denominators of all terms in the sum of Eq. (32) by their average values; we also verify our conjecture numerically in Sec. VIII below.

The fact that random symmetric states typically lead to Heisenberg scaling *for all values* of φ in the simple setup of Fig. 2 has important consequences. If this were not the case, it could be possible that for a typical symmetric state ψ_N, there are values of φ for which the sensitivity is low. Our stronger result, however, is directly useful for realizable setups: In real interferometry experiments, one typically starts the phase-estimation protocol by calibrating the device [10,11]. This is done by taking control of φ and reconstructing the $p_{n|\varphi}$ *empirically* from measurements with different known values of φ. A tomography of the state ψ_N is then not necessary, and an efficient estimator (e.g., max-likelihood [57]) that saturates the CRB (1) can always be constructed after sufficiently many protocol repetitions. Crucially, this implies that one might randomly generate (for instance, following the protocol we present in Sec. VIII) many copies of some fixed random symmetric state ψ_N and, even without being aware of its exact form, one can still typically construct an estimator that attains the Heisenberg scaling, while sensing small fluctuations of the parameter around any given value of φ.

VIII. EFFICIENT GENERATION OF RANDOM SYMMETRIC STATES

We have shown that random symmetric states have very promising properties for quantum-sensing scenarios, but so far, we have not addressed the question of how to efficiently generate such states. In this section, we demonstrate that the random symmetric states can be simulated with help of short random circuits whose outputs indeed yield, on average, the Heisenberg scaling not only of the QFI but also of the FI for the measurement scheme depicted in Fig. 2.

Concretely, we consider random circuits over a set of gates that is universal on the special unitary group of the symmetric subspace and consists of four different gates: three beam splitters and a cross-Kerr nonlinearity. A set of gates is said to be universal on a certain unitary group if, by taking products of its elements, one can obtain *arbitrarily good* approximations (in trace norm) to any unitary operation in this group. We emphasize that the universality of the symmetric subspace $\mathcal{S}_N$ is not connected to the notion of universal quantum computation. This follows from the fact that the dimension of $\mathcal{S}_N$ scales polynomially (in the case of $d=2$ modes, linearly) in the number of particles N, and consequently, this space is not sufficient for universal quantum computation. We first construct a universal set of unitary gates on $\mathcal{S}_N$ for $d=2$, which is inspired by operations commonly available when dealing with bosonic (optical and atomic) systems. We present our results in the language of two-mode interferometry (see Fig. 2).

We start with the following set of gates:

$$V_1 := \frac{1}{\sqrt{5}}\begin{pmatrix} 1 & 2\mathrm{i} \\ 2\mathrm{i} & 1 \end{pmatrix}, \qquad V_2 := \frac{1}{\sqrt{5}}\begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix},$$
$$V_3 := \frac{1}{\sqrt{5}}\begin{pmatrix} 1+2\mathrm{i} & 0 \\ 0 & 1-2\mathrm{i} \end{pmatrix}, \qquad (35)$$

known to be a "fast" universal gate set for linear optics [75–79]. The above matrices reflect how gates act on a single particle (which can be either in mode a or in mode b). The action on $\mathcal{S}_N$ is then given by $\hat{V}_j = V_j^{\otimes N}$ for $j \in \{1,2,3\}$. We now supplement the above collection by a two-mode gate corresponding to a cross-Kerr nonlinearity (with effective action time $t=\pi/3$) [80]. Concretely, we take

$$\hat{V}_{\mathrm{XK}} := \exp\left(\frac{-\mathrm{i}\pi\hat{n}_a\hat{n}_b}{3}\right), \qquad (36)$$

where $\hat{n}_{a/b}$ are the particle-number operators of modes a/b marked in Fig. 2. For the general method of checking if a given gate promotes linear optics to universality in $\mathcal{S}_N$, see Ref. [81].

In atom optics, large cross-Kerr nonlinearities and phase shifts (XKPS) can be achieved in ultracold two-component Bose gases in the so-called two-mode approximation [82–84] (see also Refs. [34,85]). In optics, reaching large XKPS is more challenging [86–88], but there has recently been spectacular progress in this area, both on weak [89,90] and strong nonlinearities [91–93]. From the theoretical perspective, using the methods of geometric control theory [94,95] and ideas from representation theory of Lie algebra [81], it is possible to prove that the gates $\{\hat{V}_1,\hat{V}_2,\hat{V}_3,\hat{V}_{\mathrm{XK}}\}$ are universal for $\mathrm{SU}(\mathcal{S}_N)$. The gate $\hat{V}_{\mathrm{XK}}$ is not the only gate yielding universality when supplemented with gates that are universal for linear optics. The comprehensive characterization of (nonlinear) gates having this property will be presented in Ref. [81].

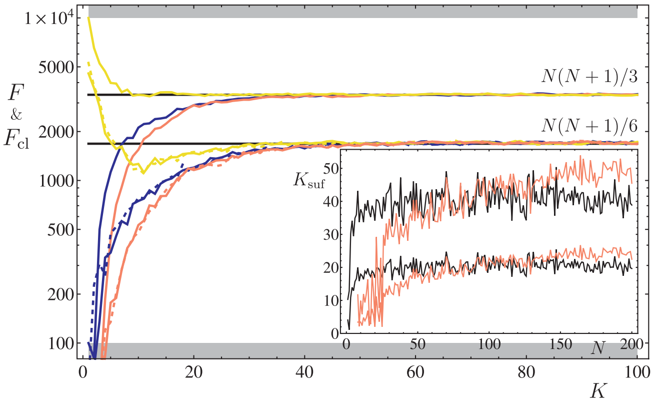

FIG. 3. Convergence of QFI and FI of random circuit states for increasing depth K. The main plot shows F and F_{cl} of $N = 100$ two-mode photons (indistinguishable qubits) for the measurement from Fig. 2 averaged over 150 realizations (sufficient to make the finite sample size irrelevant) of random circuits for different depths K. The starting states before the random circuit are of the form $|\psi_N\rangle = \sum_{n=0}^{N} \sqrt{x_n}\,|n, N-n\rangle$, with $x_0 = 1$ polarized (red line), $x_n = \binom{N}{n}/2^N$ balanced (blue line), or $x_0 = x_N = 1/2$ N00N (yellow line). Fast convergence to the values $N(N+1)/3$ and $N(N+1)/6$ (black horizontal lines) is evident in all cases. The shaded regions mark "worse than SQL" and "better than HL" precisions. The F_{cl} curves are plotted for $\varphi = \pi/2$ (solid line) and $\varphi = \pi/3$ (dotted line). The inset depicts the sufficient circuit depth K_{suf} as a function of N, such that the corresponding sample-averaged QFI (black curves) or FI (red curves) is at most 1% (top curves) or 10% (bottom curves) from its typical value. As K_{suf} grows, at most, mildly with N, for realistically achievable photon numbers [78], $K \approx 20$ may be considered sufficient.

A random circuit of depth K over this gate set is now obtained by picking, at random (according to a uniform distribution), K gates from the set $\{\hat{V}_1, \hat{V}_1^\dagger, \hat{V}_2, \hat{V}_2^\dagger, \hat{V}_3, \hat{V}_3^\dagger, \hat{V}_{\mathrm{XK}}, \hat{V}_{\mathrm{XK}}^\dagger\}$. We call states generated by applying such a circuit to some fixed symmetric state "random circuit states".

Our intuition that the above scheme should generate unitaries distributed approximately according to $\mu(\mathcal{S}_N)$ comes from the theory of the so-called ϵ-approximate unitary t-designs [54,55,96]. There are several essentially equivalent ways to define unitary designs [54]. One of them, introduced in Ref. [53], implies that an ϵ-approximate unitary t-design $\mu_{\epsilon,t}(\mathcal{H})$ is a distribution over the unitary group $\mathrm{SU}(\mathcal{H})$ acting on a Hilbert space $\mathcal{H}$ (of dimension $|\mathcal{H}|$) that efficiently approximates the Haar measure $\mu(\mathcal{H})$ in such a way that for all balanced monomials $f\colon \mathrm{SU}(\mathcal{H}) \to \mathbb{R}$ of degree t, it holds that

$$\left| \mathop{\mathbb{E}}_{U\sim\mu(\mathcal{H})} f(U) - \mathop{\mathbb{E}}_{U\sim\mu_{\epsilon,t}(\mathcal{H})} f(U) \right| \leq \epsilon/|\mathcal{H}|^t. \quad (37)$$

Moreover, if such a function satisfies a concentration inequality of the form given in Eq. (10) with respect to the Haar measure, then an (albeit weaker) concentration also holds with respect to the design [53]. All these statements carry over in a similar form to balanced polynomials. Their corresponding difference can always be bounded by the weighted sum of the differences of their constituting monomials. The isospectral QFI, $F(U) = F(U\sigma_N U^\dagger, H)$, introduced in Theorem 2 as a function of unitary rotations is precisely such a balanced polynomial of order two, which can be seen directly from Eq. (7).

For distinguishable qudits, there are efficient methods to generate approximate designs by using random circuits over local universal gate sets on $\mathcal{H}_N$ [54,55]. These constructions unfortunately do not immediately carry over to the symmetric subspace $\mathcal{S}_N$ of N qubits. However, one can use the fact proven in Ref. [55] (based on results of Ref. [52]) that in any Hilbert space $\mathcal{H}$, sufficiently long random circuits over a set of universal gates form an ϵ-approximate unitary t-design. More precisely, this holds whenever the gates employed in the circuit are nontrivial and have algebraic entries. The set of gates that are universal in the $\mathcal{S}_N$ given above satisfies this condition. For this to hold, it would actually be sufficient to replace all three gates $\hat{V}_1, \hat{V}_2, \hat{V}_3$ by a single nontrivial beam splitter [77] and the gate $\hat{V}_{\mathrm{XK}}$ by an essentially nontrivial nonlinear gate. The latter would not even have to be a reproducibly implementable nonlinear operation, but its strength could even be allowed to vary from invocation to invocation. However, it is mathematically very difficult to analytically bound the depth K of such circuits that is sufficient to achieve a given ϵ for a given t.

For this reason, we resort to a numerical analysis to verify how rapidly, with increasing K, the average QFI and the FI of random circuit states converge to the respective averages for random symmetric states. We consider the scenario of Fig. 2. In this two-mode case, $d = 2$, it holds that $|\mathcal{S}_N| = N + 1$, and from Eq. (32), we obtain concentration of the QFI around the value $\mathbb{E}F \coloneqq \mathbb{E}_{U\sim\mu(\mathcal{H})} F(U) = N(N+1)/3$ [97]. For the FI, we expect to find $\mathbb{E}F_{\mathrm{cl}} \coloneqq \mathbb{E}_{U\sim\mu(\mathcal{H})} F_{\mathrm{cl}}(U, \varphi) = N(N+1)/6$ (for details, see the discussion after Theorem 4 and Appendix C 4). In Fig. 3, we show explicitly that random circuit states generated according to our recipe indeed allow us to reach these values already for moderate K.

In Fig. 4, we further verify the behavior of the ultimate limits on the attainable precision via the relevant CRB [see Eq. (1)] dictated by the attained values for $1/F$ and $1/F_{\mathrm{cl}}$. We observe that the ultimate bounds predicted by the average $\mathbb{E}F$ and $\mathbb{E}F_{\mathrm{cl}}$ for random symmetric states are indeed quickly saturated (here, $K = 60$), and crucially, both reach the predicted Heisenberg scaling. This demonstrates that it is possible to generate states that share the favorable metrological properties of Haar-random symmetric states via the *physical* processes of applying randomly selected optical gates.

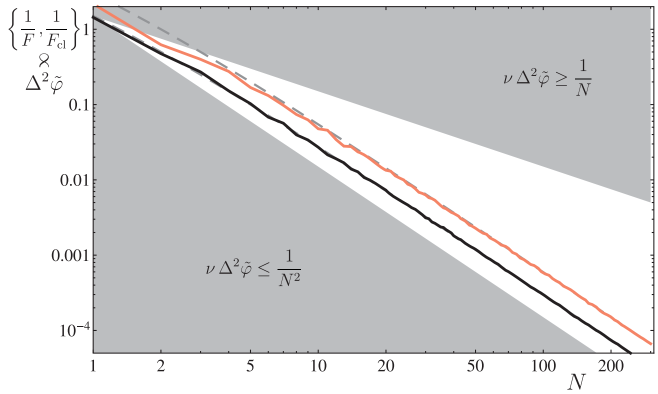

FIG. 4. Mean-squared error attained by random bosonic states generated by sufficiently deep random circuits. We depict the ultimate limit of the resolution $\nu\Delta^2\tilde{\varphi}$ attainable with random circuit states generated by applying a deep random circuit ($K = 60$) onto a two-mode balanced [$x_n = \binom{N}{n}\frac{1}{2^N}$ in Fig. 3] state for both the interferometric measurement of Fig. 2 (red line) and the theoretically optimal one yielding the QFI (black line). The corresponding sample-averaged FI and QFI quickly concentrate around the typical values $N^2/3$ and $N^2/6$ (dashed lines), respectively. The shaded regions mark the "worse than SQL" and "better than HL" precisions.

IX. CONCLUSIONS

In this work, we present a systematic study of the usefulness of random states for quantum metrology. We show that random states, sampled according to the Haar measure from the full space of states of distinguishable particles, are typically not useful for quantum enhanced metrology. In stark contrast, we prove that states from the symmetric subspace have many very promising properties for quantum metrology: They typically achieve Heisenberg scaling of the quantum Fisher information, and this scaling is robust against particle loss and equally holds for very mixed isospectral states. Moreover, we show that the high quantum Fisher information of such random states can actually be exploited with a single fixed measurement that is implementable with a beam splitter and particle-number detectors. Finally, we also demonstrate that states generated with short random circuits can be used as a resource to achieve classical Fisher information with the same scaling as the Heisenberg limit. Our results on random symmetric states open up new possibilities for quantum enhanced metrology.

Our work, which is a study initiating a new research direction, naturally raises a number of interesting questions: From the physical perspective, it would be important to investigate the impact of more realistic noise types, such as local (and correlated) dephasing, depolarization [56], and particle loss on the classical Fisher information in the interferometric scenario considered in Sec. VII, as well as quantum Fisher information, in general, for finite N. Further, it would be interesting to see whether bosonic random states are also useful for multiparameter sensing problems with noncommuting generators [72]. An important part of the quantum metrology research is devoted to infinite-dimensional optical systems with the *mean* number of particles—corresponding to the power of a light beam—being fixed [3], e.g., in squeezing-enhanced interferometry with strong laser beams of constant power [98]. Here, one could ask whether states prepared via random Gaussian transformations [99,100] or random circuits of gates that are universal for linear optics are typically useful for metrology. Another relevant problem beyond our analysis is the speed of convergence to the approximate designs when considering performance of the states prepared with random bosonic circuits discussed in Sec. VIII. Furthermore, it is interesting to study properties of the ensembles of random states generated from the Haar-random pure states and possibly particle loss of both bosons and fermions. Lastly, a natural question to be asked is whether the typical metrological usefulness of random bosonic states remains valid if one considers general phase-estimation scenarios, e.g., single-shot protocols with no prior knowledge assumed about the parameter value [68], for which Bayesian inference methods must be employed to quantify the attainable precision [69,70].

ACKNOWLEDGMENTS

We would like to thank Martin Kliesch for inspiring discussions. We acknowledge support from the European Research Council (ERC AdG OSYRIS and CoG QITBOX), the Axa Chair in Quantum Information Science, the John Templeton Foundation, EU (IP SIQS and QUIC), the Spanish National Plan FOQUS (FIS2013-46768), MINECO (Severo Ochoa Grant No. SEV-2015-0522), Fundació Privada Cellex, and Generalitat de Catalunya (Grants No. SGR 874 and No. 875). We are grateful to Centro de Ciencias de Benasque Pedro Pascual for hospitality during the "Quantum Information" conference in 2015. M. O. was supported by the START grant from the Foundation for Polish Science and NCN Grant No. DEC-2013/09/N/ST1/02772; J. K. and C. G. acknowledge funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Q-METAPP and NIMBqUS Grants No. 655161 and No. 700140. C. G. further acknowledges support by MPQ-ICFO and ICFOnest+ (FP7-PEOPLE-2013-COFUND). R. A. acknowledges funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant agreement No. 705109 (N-MuQuaS).

APPENDIX: CONTENTS

Here, we give the details that are needed to obtain the results given in the main text. In Appendix A, we discuss the concentration of measure on the special unitary group and give bounds on the Lipschitz constants of the relevant functions on this group. In Appendix B, we prove a lower bound on the QFI, which is useful when studying particle losses. In Appendix C, we give bounds for averages of FI and QFI on the relevant ensembles of density matrices that we consider—isospectral density matrices of distinguishable particles, random symmetric (bosonic) states of identical particles, and random bosonic states that underwent particle loss (Appendix C 3). In Appendix D, we use the previously derived technical results to prove Theorems 1, 2, 3, and 4 in the main text. In Appendix E, we prove the equivalence of the beam-splitter model of particle losses and the operation of taking partial traces over particles contained in two-mode bosonic systems.

Notation used throughout the paper, unless indicated differently.

Symbol/Acronym	Explanation
$\mathcal{H}_{\mathrm{loc}}$	Local Hilbert space
$d = \lvert\mathcal{H}_{\mathrm{loc}}\rvert$	Dimension of the local Hilbert space
$\mathcal{H}_N = \mathcal{H}^{\otimes N}$	Hilbert space of N distinguishable particles
$D = \lvert\mathcal{H}_N\rvert$	Dimension of the space of N distinguishable particles
$\mathcal{S}_N = \mathrm{span}\{\lvert\psi\rangle^{\otimes N} : \lvert\psi\rangle \in \mathcal{H}\}$	Hilbert space of N bosons
$\mathcal{H}$	General Hilbert space
$\lvert\mathcal{H}\rvert$	Dimension of the general Hilbert space
h	Local Hamiltonian encoding the phase φ
$H = H_N = \sum_{j=1}^{N} h^{(j)}$	Hamiltonian acting on N particles
$\mathcal{D}(\mathcal{H})$	Set of states on the Hilbert space $\mathcal{H}$
$\mathbb{1}$	Identity operator on the relevant Hilbert space
$\rho, \sigma, \ldots$	Symbols denoting (in general) mixed states
$\psi, \phi, \ldots$	Symbols denoting pure states
$F_{\mathrm{cl}}(\{p_{n\vert\varphi}\})$	Classical Fisher information associated with the family of probability distributions $\{p_{n\vert\varphi}\}$
$F(\rho, H)$	Fisher information computed for the state ρ with respect to the Hamiltonian H
$d_B(\rho, \sigma)$	Bures distance between states ρ and σ
$\mathcal{F}(\rho, \sigma)$	Uhlmann fidelity between states ρ and σ
$\mathbb{P}^{\mathrm{sym}}$	Orthogonal projector onto $\mathrm{Sym}^2(\mathcal{H})$
$\mathbb{P}^{\mathrm{asym}}$	Orthogonal projector onto $\bigwedge^2(\mathcal{H})$
$\mathrm{End}(\mathcal{H})$	Set of linear operators on $\mathcal{H}$
$\mathrm{Herm}(\mathcal{H})$	Set of Hermitian operators on $\mathcal{H}$
$\mathrm{SU}(\mathcal{H})$	Special unitary group on $\mathcal{H}$
$\mu(\mathcal{H})$	Haar measure on $\mathrm{SU}(\mathcal{H})$
$\mathbb{E}_{U\sim\mu(\mathcal{H})}$	Expectation value (average) with respect to $\mu(\mathcal{H})$
Ω	Set of isospectral density matrices in $\mathcal{H}$ (for the specified ordered spectrum)
QFI	Quantum Fisher information
FI	Classical Fisher information
GHZ	Greenberger-Horne-Zeilinger state
SQL	Standard quantum limit
HL	Heisenberg Limit
POVM	Positive operator-valued measure

APPENDIX A: CONCENTRATION OF MEASURE ON SPECIAL UNITARY GROUP AND LIPSCHITZ CONSTANTS FOR QFI AND FI

In this appendix, we first present a basic concentration of measure inequalities on the special unitary group $\mathrm{SU}(\mathcal{H})$. Then, we give bounds on the Lipschitz constants of various functions based on QFI and FI that appear naturally, while studying different statistical ensembles of states on $\mathcal{H}$—isospectral density matrices, partially traced isospectral density matrices, etc.

1. Concentration of measure on unitary group

We make extensive use of the concentration of measure phenomenon on the special unitary group $\mathrm{SU}(\mathcal{H})$. It will be convenient to use a metric tensor g_{HS} induced on $\mathrm{SU}(\mathcal{H})$ from the embedding of $\mathrm{SU}(\mathcal{H})$ in the set of all linear operators, $\mathrm{End}(\mathcal{H})$, equipped with the Hilbert-Schmidt inner product $\langle A, B\rangle = \mathrm{tr}(A^\dagger B)$. Let us write the formula for g_{HS} explicitly. The special unitary group is a Lie group; thus, for every $U \in \mathrm{SU}(\mathcal{H})$, we have an isomorphism $T_U\mathrm{SU}(\mathcal{H}) \approx \mathfrak{su}(\mathcal{H})$, where T_U is the tangent space to SU at U and $\mathfrak{su}(\mathcal{H})$ is the Lie algebra of the group consisting of Hermitian traceless operators on $\mathcal{H}$. The linear isomorphism is given by the following mapping:

$$\mathfrak{su}(\mathcal{H}) \ni X \mapsto \hat{X} = \frac{d}{d\varphi}\bigg|_{\varphi=0} \exp(-\mathrm{i}\varphi X)U \in T_U\mathrm{SU}(\mathcal{H}). \tag{A1}$$

Using the identification Eq. (A1) and treating the operator $\hat{X} = \frac{d}{d\varphi}\big|_{\varphi=0} \exp(-\mathrm{i}\varphi X)U = -\mathrm{i}XU$ as an element of $\mathrm{End}(\mathcal{H})$, we get

$$g_{\mathrm{HS}}(\hat{X},\hat{Y}) = \langle \mathrm{i}XU, \mathrm{i}YU\rangle = \mathrm{tr}([\mathrm{i}XU]^\dagger \mathrm{i}YU) = \mathrm{tr}(XY), \tag{A2}$$

where we have used $X^\dagger = X$, the identity $UU^\dagger = \mathbb{1}$, and the cyclic property of the trace. The gradient of a smooth function $f{:}\,\mathrm{SU}(\mathcal{H}) \to \mathbb{R}$ at point $U \in \mathrm{SU}(\mathcal{H})$ is defined by the condition

$$g_{\mathrm{HS}}(\nabla \hat{f}|_U, \hat{X}) = \frac{d}{d\phi}\bigg|_{\phi=0} f(\exp(-\mathrm{i}\phi X)U), \tag{A3}$$

which has to be satisfied for all $X \in \mathfrak{su}(\mathcal{H})$.

Fact 1. (Concentration of measure on $\mathrm{SU}(\mathcal{H})$. [50]) Consider a special unitary group $\mathrm{SU}(\mathcal{H})$ equipped with the Haar measure μ and the metric g_{HS}. Let

$$f{:}\,\mathrm{SU}(\mathcal{H}) \mapsto \mathbb{R} \tag{A4}$$

be a smooth function on $\mathrm{SU}(\mathcal{H})$ with the mean $\mathbb{E}_\mu f$, and let

$$L = \sqrt{\max_{U\in \mathrm{SU}(\mathcal{H})} g_{\mathrm{HS}}(\nabla f, \nabla f)} \tag{A5}$$

be the Lipschitz constant of f. Then, for every $\epsilon \geq 0$, the following concentration inequalities hold:

$$\Pr_{U\sim\mu(\mathcal{S}_N)}\left(f(U) - \mathop{\mathbb{E}}_{U\sim\mu(\mathcal{H})} f \geq \epsilon\right) \leq \exp\left(-\frac{D\epsilon^2}{4L^2}\right), \tag{A6}$$

$$\Pr_{U\sim\mu(\mathcal{S}_N)}\left(f(U) - \mathop{\mathbb{E}}_{U\sim\mu(\mathcal{H})} f \leq -\epsilon\right) \leq \exp\left(-\frac{D\epsilon^2}{4L^2}\right), \tag{A7}$$

where $D = |\mathcal{H}|$ is the dimension of $\mathcal{H}$.

Fact 2. Concentration inequalities Eq. (A7) also hold for the general (not necessarily smooth) L-Lipschitz functions $f{:}\,\mathrm{SU}(\mathcal{H}) \mapsto \mathbb{R}$ [50], that is, functions satisfying

$$|f(U) - f(V)| \leq L\mathrm{d}(U, V), \tag{A8}$$

where $\mathrm{d}(V, W)$ is the geodesic distance between unitaries U and V given by

$$\mathrm{d}(U, V) \coloneqq \inf_{\gamma:\gamma(0)=U,\gamma(1)=V} \mathrm{D}_\gamma, \tag{A9}$$

with $\mathrm{D}_\gamma \coloneqq \int_{[0,1]} \sqrt{g_{\mathrm{HS}}(\frac{d\gamma}{dt}, \frac{d\gamma}{dt})}$, and the infimum is over the (piecewise smooth) curves γ that start at U and end at V.

Remark 1. Because of the definition of the gradient ∇f (A3) and the structure of the tangent space $T_U\mathrm{SU}(\mathcal{H})$ for $U \in \mathrm{SU}(\mathcal{H})$ [see Eq. (A1)], we have

$$\mathrm{tr}(\nabla f|_U X) = \frac{d}{d\phi}\bigg|_{\phi=0} f(\exp(-\mathrm{i}\phi X)U), \tag{A10}$$

where $X = X^\dagger$ and $\mathrm{tr}X = 0$. Assume that for $C > 0$, we can find the bound

$$\left|\frac{d}{d\phi}\bigg|_{\phi=0} f(\exp(-\mathrm{i}\phi X)U)\right| \leq C\|X\|_{\mathrm{HS}}, \tag{A11}$$

which is valid for all $U \in \mathrm{SU}(\mathcal{H})$. Then, from Eq. (A10), we can conclude that C is an upper bound on the Lipschitz constant of f.

2. Lipschitz constants for quantum Fisher information for general Hamiltonian encoding

Recall that for unitary encodings, the quantum Fisher information for a state ρ with spectral decomposition $\sum_i p_i|e_i\rangle\langle e_i|$ is

$$F(\rho, H) = 2\sum_{i,j:\,p_i+p_j\neq 0} \frac{(p_i - p_j)^2}{p_i + p_j} |\langle e_i|H|e_j\rangle|^2, \tag{A12}$$

where H is the Hamiltonian generating the unitary evolution of mixed states,

$$\mathbb{R} \ni \varphi \mapsto \rho(\varphi) \coloneqq \exp(-i\varphi H)\rho\exp(i\varphi H) \in \mathcal{D}(\mathcal{H}). \tag{A13}$$

The QFI depends on both the state $\rho \in \mathcal{D}(\mathcal{H})$ and the Hamiltonian $H \in \mathrm{Herm}(\mathcal{H})$ encoding the phase φ. In what follows, without any loss of generality, we assume that $\mathrm{tr}(H) = 0$. We are interested in the behavior of $F(\rho, H)$ when H is fixed and ρ varies over some ensemble of (generally mixed) states. As we want to use concentration inequalities [of the type in Eq. (A7)], our aim here is to give bounds on Lipschitz constants of QFI on relevant sets of density matrices.

We first study QFI on the set of isospectral density matrices

$$\Omega_{(p_1,\ldots,p_D)} \coloneqq \{\rho \in \mathcal{D}(\mathcal{H}) | \mathrm{sp}_\uparrow(\rho) = \{p_j\}_j\}, \tag{A14}$$

where $\mathrm{sp}_\uparrow(\rho)$ denotes the vector on nonincreasingly ordered eigenvalues of ρ. In what follows, for the sake of simplicity, we use the shorthand notation $\Omega_{(p_1,\ldots,p_D)} \coloneqq \Omega$. Let

$$F_{\Omega,H}{:}\,\mathrm{SU}(\mathcal{H}) \ni U \mapsto F(U\rho_0 U^\dagger, H) \in \mathbb{R}, \tag{A15}$$

where ρ_0 is the arbitrarily chosen state belonging to Ω. Then, we can prove the following lemma.

Lemma 1. The Lipschitz constant (with respect to g_{HS}) of the function $F_{\Omega,H}$ defined by Eq. (A15) is upper bounded by

$$L_\Omega \leq \min\left\{1, 2\sqrt{2}\sqrt{d_B\left(\rho, \frac{\mathbb{1}}{D}\right)}\right\} 32\|H\|^2, \tag{A16}$$

where $\mathbb{1}/D$ is the maximally mixed state on $\mathcal{H}$.

Remark 2. The quantity $d_B(\rho, \mathbb{1}/D)$ depends only on the spectrum of ρ and thus is constant on the set of isospectral density matrices Ω.

Proof.——In a recent paper [30], the following inequality was proven:

$$|F(\rho, H) - F(\sigma, H)| \leq 32 d_B(\rho, \sigma)\|H\|^2, \quad \text{(A17)}$$

where $d_B(\rho,\sigma) = \sqrt{2[1-\mathcal{F}(\rho,\sigma)]}$ is the Bures distance between density matrices, with $\mathcal{F}(\rho,\sigma) = \mathrm{tr}\sqrt{\sigma^{1/2}\rho\sigma^{1/2}}$ denoting the fidelity. Inserting

$$\rho = U\rho_0 U^\dagger \quad \text{and} \quad \sigma = \exp(-i\phi X)U\rho_0 U^\dagger \exp(i\phi X) \quad \text{(A18)}$$

into Eq. (A17), one obtains

$$\begin{aligned}&|F_{\Omega,H}(\exp(-i\phi X)U) - F_{\Omega,H}(U)| \\ &\quad\leq 32 d_B(U\rho_0 U^\dagger, \exp(-i\phi X)U\rho_0 U^\dagger \exp(i\phi X))\|H\|^2.\end{aligned} \quad \text{(A19)}$$

Dividing the above by $|\phi|$ and taking the limit $\phi \to 0$, one arrives at

$$\begin{aligned}&\left|\frac{d}{d\phi}\bigg|_{\phi=0} F_{\Omega,H}(\exp(-i\phi X)U)\right| \\ &\quad\leq 32\|H\|^2 \lim_{\phi\to 0}\frac{1}{|\phi|} d_B(U\rho_0 U^\dagger, \exp(-i\phi X) \\ &\qquad\times U\rho_0 U^\dagger \exp(i\phi X)).\end{aligned} \quad \text{(A20)}$$

Then, Eq. (6) implies that

$$\lim_{\phi\to 0}\frac{1}{|\phi|} d_B(\rho, \exp(-i\phi X)\rho\exp(i\phi X)) = \frac{1}{2}\sqrt{F(\rho, X)}, \quad \text{(A21)}$$

and therefore,

$$\left|\frac{d}{d\phi}\bigg|_{\phi=0} F_{\Omega,H}(\exp(-i\phi X)U)\right| \leq 16\|H\|^2\sqrt{F(\rho, X)}. \quad \text{(A22)}$$

In addition, we have two upper bounds on the square root of the quantum Fisher information:

$$\sqrt{F(\rho, H)} \leq 2\|H\| \leq 2\|H\|_{\mathrm{HS}} \quad \text{(A23)}$$

and

$$\begin{aligned}\sqrt{F(\rho,H)} &\leq 4\sqrt{2}\sqrt{d_B\left(\rho, \frac{\mathbb{1}}{D}\right)}\|H\| \\ &\leq 4\sqrt{2}\sqrt{d_B\left(\rho, \frac{\mathbb{1}}{D}\right)}\|H\|_{\mathrm{HS}},\end{aligned} \quad \text{(A24)}$$

where Eq. (A23) follows from the fact that the maximal value of the QFI for the phase encoded via the Hamiltonian H is bounded from above by $4\|H\|^2$ [1], and Eq. (A24) follows from Eq. (A17) for $\sigma = \mathbb{1}/D$, for which the QFI trivially vanishes. Combining inequalities Eqs. (A23) and (A24) with Eq. (A22) and using Remark 1, one finally obtains Eq. (A16). ■

Remark 3. The upper bound on the Lipschitz constant of $F_{\Omega,H}$ given in Eq. (A16) depends explicitly on the spectrum of the considered set of isospectral density matrices. Specifically, the right-hand side of Eq. (A16) decreases as ρ becomes more mixed. For special cases of Haar-random pure states and random depolarized states (see below), we can get better bounds on the Lipschitz constant of the QFI.

Lemma 2. Consider the ensemble of Haar-random depolarized pure states,

$$\rho = (1-p)\psi + p\frac{\mathbb{1}}{D}, \quad \text{(A25)}$$

where ψ stands for the projector onto a Haar-random pure state $|\psi\rangle$ and $p \in [0,1]$. For fixed p, the states in Eq. (A25) form an ensemble of isospectral density matrices since, for any such ρ, we have

$$\mathrm{sp}_\uparrow(\rho) = \left(1 - p + \frac{p}{D}, \frac{p}{D}, \ldots, \frac{p}{D}\right). \quad \text{(A26)}$$

For this particular spectrum, the Lipschitz constant L_p (with respect to g_{HS}) of the function $F_p := F_{\Omega,H}$, defined by Eq. (A15), is upper bounded by

$$L_p \leq 16\frac{(1-p)^2}{1-p+\frac{2p}{D}}\|H\|^2. \quad \text{(A27)}$$

Proof.—Let us first note that for ρ given by Eq. (A25), the QFI takes the form [2]

$$F(\rho, H) = \frac{(1-p)^2}{1-p+\frac{2p}{D}}F(\psi, H), \quad \text{(A28)}$$

from which it directly follows that, for all $U \in \mathrm{SU}(\mathcal{H})$,

$$F_p(U) = \frac{(1-p)^2}{1-p+\frac{2p}{D}}F_p(U), \quad \text{(A29)}$$

and consequently,

$$L_p = \frac{(1-p)^2}{1-p+\frac{2p}{D}}L_0. \quad \text{(A30)}$$

One can estimate L_0 by exploiting the fact that, for pure states, the QFI is simply $F(\psi_0, H) = 4\{\mathrm{tr}(\psi_0 H^2) - [\mathrm{tr}(\psi_0 H)]^2\}$, which allows one to express $F_0(U)$ as

$$F_0(U) = \mathrm{tr}[(U \otimes U\psi_0 \otimes \psi_0 U^\dagger \otimes U^\dagger)V], \quad \text{(A31)}$$

where $V = 4(H^2 \otimes \mathbb{1} - H \otimes H)$. By the virtue of Lemma 6.1 of Ref. [38] (see also Ref. [37]), the Lipschitz constant of F_0 is bounded by $2\|V\| \leq 16\|H\|^2$. Combining this with Eq. (A30) yields Eq. (A27). ■

It is also possible to prove the Lipschitz continuity of the optimized version of QFI on Ω,

$$F^{\mathcal{V}}_{\Omega,H}: \mathrm{SU}(\mathcal{H}) \ni U \mapsto \sup_{V\in\mathcal{V}} F_{\Omega,H}(VU) \in \mathbb{R}, \quad \text{(A32)}$$

where $\mathcal{V} \subset \mathrm{SU}(\mathcal{H})$ is a compact class of unitary gates on $\mathcal{H}$.

Lemma 3. The Lipschitz constant $L^{\mathcal{V}}_\Omega$ (with respect to the geodesic distance) of the function $F^{\mathcal{V}}_\Omega$ defined by Eq. (A32) is upper bounded by the Lipschitz constant of F_Ω,

$$L^{\mathcal{V}}_\Omega \leq L_\Omega. \quad \text{(A33)}$$

Proof.—Let $U, U' \in \mathrm{SU}(\mathcal{H})$. Without loss of generality, we can assume $F^{\mathcal{V}}_{\Omega,H}(U) \geq F^{\mathcal{V}}_{\Omega,H}(U')$. Let $V_0 \in \mathcal{V}$ be the element such that

$$F_{\Omega,H}(V_0U) = F^{\mathcal{V}}_{\Omega,H}(U) = \sup_{V\in\mathcal{V}} F_{\Omega,H}(VU). \quad \text{(A34)}$$

Consequently, we have the following inequalities:

$$\begin{aligned} |F^{\mathcal{V}}_{\Omega,H}(U) - F^{\mathcal{V}}_{\Omega,H}(U')| &= F_{\Omega,H}(V_0U) - F^{\mathcal{V}}_{\Omega,H}(U') \\ &\leq F_{\Omega,H}(V_0U) - F_{\Omega,H}(V_0U') \\ &\leq L_\Omega \mathrm{d}(U,U'), \end{aligned} \quad \text{(A35)}$$

where, in the last inequality, we used Lipschitz continuity of $F_{\Omega,H}$, which is guaranteed by Lemma A 2. ■

Remark 4. For us, the case of greatest interest is $\mathcal{H} = \mathcal{H}_N$ and $\mathcal{V} = \mathrm{LU}$ (local unitary group on N distinguishable particles).

3. Lipschitz constants for quantum Fisher information with particle losses

We now give bounds on the Lipschitz constant of the QFI in the case of particle losses for bosonic states. Recall that, in this setting, the Hamiltonian acting on N particles is given by $H_N = \sum_{i=1}^N h^{(i)}$ and that the Hilbert space of the system is the totally symmetric space of N particles denoted by $\mathcal{S}_N$. Let us define a function

$$F^{[k]}_\Omega: \mathrm{SU}(\mathcal{S}_N) \ni U \mapsto F(\mathrm{tr}_k(U\rho U^\dagger), H_{N-k}) \in \mathbb{R}. \quad \text{(A36)}$$

Lemma 4. The Lipschitz constant (with respect to g_{HS}) of the function $F^{[k]}_\Omega$ defined by Eq. (A36) is upper bounded by

$$L^{[k]}_\Omega \leq \min\left\{1, 2\sqrt{2}\sqrt{d_B\left(\rho, \frac{\mathbb{P}^N_{\mathrm{sym}}}{|\mathcal{S}_N|}\right)}\right\} 32\|H_{N-k}\|^2, \quad \text{(A37)}$$

where $\mathbb{P}^N_{\mathrm{sym}}/|\mathcal{S}_N|$ is the maximally mixed state on $\mathcal{S}_N$ and $\mathbb{P}^N_{\mathrm{sym}}$ stands for the projector onto $\mathcal{S}_N$.

Proof.—We prove Eq. (A37) in an analogous way to Eq. (A16). Let $\rho' = \mathrm{tr}_k(\rho)$ and $\sigma' = \mathrm{tr}_k(\sigma)$ be two states on $\mathcal{S}_{N-k}$ obtained by tracing out k particles from ρ and σ, respectively. Applying the inequality Eq. (A17) to ρ' and σ' and the Hamiltonian $H = H_{N-k}$, one obtains

$$\begin{aligned} |F(\rho',H) - F(\sigma',H)| &\leq 32 d_B(\rho',\sigma')\|H_{N-k}\|^2 \\ &\leq 32 d_B(\rho,\sigma)\|H_{N-k}\|^2, \end{aligned} \quad \text{(A38)}$$

where the second inequality follows from the fact that the Bures distance does not increase under trace-preserving completely positive maps [66] [for us, the relevant TPCP map is the partial trace, $\rho \mapsto \mathrm{tr}_k(\rho)$]. We now set

$$\begin{aligned} \rho &= U\rho_0U^\dagger, \\ \sigma &= \exp(-i\varphi X)U\rho_0U^\dagger \exp(i\varphi X), \end{aligned} \quad \text{(A39)}$$

where $U \in \mathrm{SU}(\mathcal{S}_N)$ and $X \in \mathfrak{su}(\mathcal{S}_N)$, and the rest of the proof is exactly the same as that of Lemma 1. ■

4. Lipschitz constant of the classical Fisher information

We conclude this section by giving bounds on the Lipschitz constant of the classical Fisher information for the case of isospectral mixed states, fixed Hamiltonian encoding, and a fixed measurement setting. Recall that for the unitary encoding (A13), classical Fisher information is a function of the state $\rho \in \mathcal{D}(\mathcal{H})$, Hamiltonian H, the phase φ, and the POVM $\{\Pi_n\}$ used in the phase-estimation procedure. These three objects define a family of probability distributions

$$p_{n|\varphi}(\rho(\varphi)) = \mathrm{tr}(\Pi_n\rho(\varphi)), \quad \text{(A40)}$$

where $\rho(\varphi) = \exp(-\mathrm{i}\varphi H)\rho\exp(\mathrm{i}\varphi H)$. The classical Fisher information is then given by

$$F_{\mathrm{cl}}(\rho,H,\varphi,\{\Pi_n\}) \equiv F_{\mathrm{cl}}(\{p_{n|\varphi}\}) = \sum_n \frac{\mathrm{tr}(\mathrm{i}[\Pi_n,H]\rho(\varphi))^2}{\mathrm{tr}(\Pi_n\rho(\varphi))}, \quad \text{(A41)}$$

where the summation is over the range of indices labeling the outputs of a POVM $\{\Pi_n\}$ (for simplicity, we consider POVMs with a finite number of outcomes). Let us fix the Hamiltonian H, the phase φ, and the POVM $\{\Pi_n\}$. Let us define a function

$$F_{\mathrm{cl},\Omega,H} : \mathrm{SU}(\mathcal{H}) \ni U \mapsto F_{\mathrm{cl}}(U\rho_0 U^\dagger, H, \varphi, \{\Pi_n\}) \in \mathbb{R}, \tag{A42}$$

for some fixed state $\rho_0 \in \Omega$.

Lemma 5. The Lipschitz constant (with respect to g_{HS}) of the function $F_{\mathrm{cl},\Omega,H}$ defined by Eq. (A42) is upper bounded by

$$L_{\mathrm{cl},\Omega,H} \le 24\|H\|^2. \tag{A43}$$

Proof.—The strategy of the proof is analogous to the one presented in the other lemmas in this section. The idea is to find a bound for

$$\left| \frac{d}{d\phi}\bigg|_{\phi=0} F_{\mathrm{cl},\Omega,H}(\exp(-\mathrm{i}\phi X)U) \right|, \tag{A44}$$

in terms of the Hilbert-Schmidt norm of X. Let us first assume that at $U \in \mathrm{SU}(\mathcal{H})$ for all n,

$$\mathrm{tr}(\Pi_n \rho(\varphi)) \neq 0. \tag{A45}$$

Under the above condition, we have

$$\begin{aligned} \frac{d}{d\phi}\bigg|_{\phi=0} & F_{\mathrm{cl},\Omega,H}(\exp(-\mathrm{i}\phi X)U) \\ &= \sum_n \frac{\mathrm{tr}\{[H,\Pi_n]\rho_U(\varphi)\}\mathrm{tr}\{[H,\Pi_n][\mathrm{i}X,\rho_U(\varphi)]\}}{\mathrm{tr}(\Pi_n\rho_U(\varphi))} \\ &\quad + \sum_n \frac{\mathrm{tr}\{[H,\Pi_n]\rho_U(\varphi)\}^2 \mathrm{tr}\{[\mathrm{i}H,\Pi_n]\rho_U(\varphi)\}}{\mathrm{tr}(\Pi_n\rho_U(\varphi))^2}, \end{aligned} \tag{A46}$$

where $\rho_U(\varphi) = \exp(-\mathrm{i}\varphi H)U\rho_0 U^\dagger \exp(\mathrm{i}\varphi H)$. Let us introduce the auxiliary notation

$$A = \sum_n \left| \frac{\mathrm{tr}\{[H,\Pi_n]\rho_U(\varphi)\}\mathrm{tr}\{[H,\Pi_n][\mathrm{i}X,\rho_U(\varphi)]\}}{\mathrm{tr}(\Pi_n\rho_U(\varphi))} \right|, \tag{A47}$$

$$B = \sum_n \left| \frac{\mathrm{tr}\{[H,\Pi_n]\rho_U(\varphi)\}^2 \mathrm{tr}\{[\mathrm{i}H,\Pi_n]\rho_U(\varphi)\}}{\mathrm{tr}(\Pi_n\rho_U(\varphi))^2} \right|. \tag{A48}$$

Clearly, we have the inequality

$$\left| \frac{d}{d\phi}\bigg|_{\phi=0} F_{\mathrm{cl},\Omega,H}(\exp(-\mathrm{i}\phi X)U) \right| \le A + B. \tag{A49}$$

In order to bound A and B (from above), we observe that, for any state $\rho \in \mathcal{D}(\mathcal{H})$, we have

$$|\mathrm{tr}([H,\Pi_n]\rho)| \le 2\mathrm{tr}(\Pi_n\rho)\|H\|, \tag{A50}$$

$$|\mathrm{tr}([X,\Pi_n]\rho)| \le 2\mathrm{tr}(\rho\Pi_n)\|X\|, \tag{A51}$$

$$|\mathrm{tr}(\{[H,\Pi_n][\mathrm{i}X,\rho]\})| \le 4\mathrm{tr}(\rho\Pi_n)\|X\|\|H\|. \tag{A52}$$

In order to prove (A50), we first upper bound $|\mathrm{tr}(H\Pi_n\rho)|$,

$$|\mathrm{tr}(H\Pi_n\rho)| = \left|\mathrm{tr}\left(H\sqrt{\Pi_n}\sqrt{\Pi_n}\sqrt{\rho}\sqrt{\rho}\right)\right| \tag{A53}$$

$$\le \sqrt{\mathrm{tr}(\rho H^2\Pi_n)}\sqrt{\mathrm{tr}(\rho\Pi_n)} \tag{A54}$$

$$\le \sqrt{\sqrt{\mathrm{tr}(\rho\Pi_n^2)}\sqrt{\mathrm{tr}(\rho H^4)}}\sqrt{\mathrm{tr}(\rho\Pi_n)} \tag{A55}$$

$$\le \mathrm{tr}(\rho\Pi_n)\|H\|, \tag{A56}$$

where in (A53), we have used the non-negativity of operators Π_n and ρ. In Eq. (A55), we have repetitively used the Cauchy-Schwarz inequality, first for $P = \sqrt{\rho}H\sqrt{\Pi_n}$, $Q = \sqrt{M}\sqrt{\rho}$ and then for $P = \sqrt{\rho}H^2$, $Q = M\sqrt{\rho}$. The final inequality (A56) follows immediately from operator inequalities,

$$H^4 \le \|H\|^4 \mathbb{1}, \qquad \Pi_n \le \Pi_n^2. \tag{A57}$$

Using analogous reasoning, it is possible to prove $|\mathrm{tr}(H\Pi_n\rho)| \le tr(\rho\Pi_n)\|H\|$. This finishes the proof of Eq. (A50). Using essentially the same methodology, it is possible to prove the inequalities (A52) and (A51). By plugging inequalities (A50)–(A52) into Eq. (A49) for $\rho = \rho_U(\varphi)$ and using the normalization condition

$$\sum_n \mathrm{tr}(\Pi_n\rho_U(\varphi)) = 1, \tag{A58}$$

we obtain

$$\left| \frac{d}{d\phi}\bigg|_{\phi=0} F_{\mathrm{cl},\Omega,H}(\exp(-\mathrm{i}\phi X)U) \right| \le 24\|H\|^2\|X\| \tag{A59}$$

$$\le 24\|H\|^2\|X\|_{\mathrm{HS}}. \tag{A60}$$

By the virtue of Remark 1, we conclude that the Lipschitz constant of $F_{\mathrm{cl},\Omega,H}(U)$ is upper bounded by $24\|H\|$. The above derivation explicitly used the assumption (A45), which translates to assuming that denominators appearing in the definition of classical Fisher information do not vanish. However, with the help of inequalities (A50)–(A52), one can easily prove that the possible singularities coming from zeros of some denominators are actually removable and that $F_{\mathrm{cl},\Omega,H}(U)$ is actually a differentiable function of U. Consequently, inequality (A59) is actually satisfied for $U \in \mathrm{SU}(\mathcal{H})$, for which conditions (A45) are not satisfied. ∎

APPENDIX B: LOWER BOUNDS ON THE QFI

Lemma 6. Let ρ_φ be a one-parameter family of states on a Hilbert space $\mathcal{H}$. Then, the following lower bound for QFI holds (see also Ref. [101]):

$$F(\rho_\varphi, \dot{\rho}_\varphi) \geq \|\dot{\rho}\|_1^2. \tag{B1}$$

In particular, for $\rho_\varphi = \exp(-iH\varphi)\rho\exp(iH\varphi)$, we have

$$F(\rho, H) \geq \|[H,\rho]\|_1^2. \tag{B2}$$

Recall that the right-hand side of Eq. (B2), $\|[H,\rho]\|_1^2$, equals the measure of asymmetry introduced in Ref. [73].

Proof.——We give the proof only in the Hamiltonian case (B2). The proof of the general case is analogous to this. Recall that the quantum Fisher information is related to the Bures distance $d_\mathrm{B}(\rho,\sigma) = \sqrt{2[1-\mathcal{F}(\rho,\sigma)]}$ through

$$d_\mathrm{B}(\rho_\varphi, \rho_{\varphi+\delta\varphi}) = \frac{1}{2}\sqrt{F(\rho_\varphi, H)}|\delta\varphi| + O(\delta\varphi^2). \tag{B3}$$

At the same time, from the Fuchs–van der Graaf inequalities [102], we know that

$$\begin{aligned}\|\rho-\sigma\|_1 &\leq 2\sqrt{1-\mathcal{F}(\rho,\sigma)^2}\\ &\leq 2\sqrt{2[1-\mathcal{F}(\rho,\sigma)]}\\ &= 2d_B(\rho,\sigma),\end{aligned} \tag{B4}$$

with the second inequality stemming from the fact that $\mathcal{F}(\rho,\sigma) \leq 1$. By combining Eqs. (B3) and (B4), we obtain

$$\|\rho_\varphi - \rho_{\varphi+\delta\varphi}\|_1 \leq \sqrt{F(\rho_\varphi, H)}|\delta\varphi| + O(\delta\varphi^2). \tag{B5}$$

Dividing this by $|\delta\varphi|$ and then taking the limit $\delta\varphi \to 0$, we get

$$\lim_{\delta\varphi\to 0}\frac{1}{|\delta\varphi|}\|\rho_\varphi - \rho_{\varphi+\delta\varphi}\|_1 \leq \sqrt{F(\rho_\varphi, H)}, \tag{B6}$$

which, by virtue of the fact that

$$\lim_{\delta\varphi\to 0}\frac{1}{|\delta\varphi|}\|\rho_\varphi - \rho_{\varphi+\delta\varphi}\|_1 = \|\dot{\rho}\|_1, \tag{B7}$$

directly leads us to Eq. (B2). ∎

Remark 5. Using the standard inequality between trace and Hilbert-Schmidt norms, we obtain the weaker versions of inequalities (B1) and (B2),

$$F(\rho_\varphi, \dot{\rho}_\varphi) \geq \|\dot{\rho}\|_\mathrm{HS}^2, \tag{B8}$$

$$F(\rho, H) \geq \|[H,\rho]\|_\mathrm{HS}^2. \tag{B9}$$

Remark 6. For pure states, $F(\rho, H) = \|[H,\rho]\|_1^2$.

Proof.—Let us prove that for pure states, the right-hand side of Eq. (B2) is simply the QFI of ρ. To this end, let us denote $H|\psi\rangle = |\tilde{\varphi}\rangle$ and then

$$\begin{aligned}\|[H,\rho]\|_1^2 &= \||\tilde{\varphi}\rangle\langle\psi| - |\psi\rangle\langle\tilde{\varphi}|\|_1^2\\ &= \langle\psi|H^2|\psi\rangle\||\varphi\rangle\langle\psi| - |\psi\rangle\langle\varphi|\|_1^2,\end{aligned} \tag{B10}$$

where $|\varphi\rangle = |\tilde{\varphi}\rangle/\sqrt{\langle\psi|H^2|\psi\rangle}$. The matrix under the trace norm is manifestly anti-Hermitian and of rank two, so it is straightforward to compute its norm. To do this, let us write $|\varphi\rangle = \alpha|\psi\rangle + \beta|\psi^\perp\rangle$, where $|\alpha|^2 + |\beta|^2 = 1$ and $|\psi^\perp\rangle$ is some normalized vector orthogonal to $|\psi\rangle$. It also follows that $\alpha \in \mathbb{R}$ because $\alpha = \langle\psi|\varphi\rangle = \langle\psi|H|\psi\rangle/\sqrt{\langle\psi|H^2|\psi\rangle}$. All this implies that

$$|\varphi\rangle\langle\psi| - |\psi\rangle\langle\varphi| = \beta|\psi^\perp\rangle\langle\psi| - \beta^*|\psi\rangle\langle\psi^\perp|, \tag{B11}$$

and consequently, the eigenvalues of the above matrix are $\pm\mathrm{i}|\beta|$. Thus, its trace norm amounts to $2|\beta|$, giving

$$\|[H,\rho]\|_1^2 = 4\langle\psi|H^2|\psi\rangle|\beta|^2 = 4\langle\psi|H^2|\psi\rangle(1-\alpha^2) \tag{B12}$$

$$= 4\langle\psi|H^2|\psi\rangle\left(1 - \frac{\langle\psi|H|\psi\rangle^2}{\langle\psi|H^2|\psi\rangle}\right) \tag{B13}$$

$$= 4(\langle\psi|H^2|\psi\rangle - \langle\psi|H|\psi\rangle^2) = F(\rho, H). \tag{B14}$$

∎

APPENDIX C: AVERAGES AND BOUNDS ON AVERAGES OF QFI AND FI ON RELEVANT STATISTICAL ENSEMBLE

In this appendix, we compute and/or bound averages of the FI or QFI on the ensembles of mixed quantum states appearing in the main text.

1. Averages of QFI on ensembles of isospectral density matrices

We extensively use the following result concerning the integration on the special unitary group.

Fact 3. (Integration of a quadratic function on the unitary group.) Letting $V \in \mathrm{Herm}(\mathcal{H}\otimes\mathcal{H})$, the following equality holds [67],

$$\int_{\mathrm{SU}(\mathcal{H})} d\mu(U) U^{\otimes 2} V (U^\dagger)^{\otimes 2} = \alpha\mathbb{P}^\mathrm{sym} + \beta\mathbb{P}^\mathrm{asym}, \tag{C1}$$

where $\mathbb{P}^\mathrm{sym}$ and $\mathbb{P}^\mathrm{asym}$ are projectors onto the symmetric and antisymmetric subspaces of $\mathcal{H}\otimes\mathcal{H}$ which can be expressed as

$$\mathbb{P}^\mathrm{sym} = \frac{1}{2}(\mathbb{1}\otimes\mathbb{1} + \mathbb{S}), \quad \mathbb{P}^\mathrm{asym} = \frac{1}{2}(\mathbb{1}\otimes\mathbb{1} - \mathbb{S}), \tag{C2}$$

with $\mathbb{S}$ being the so-called swap operator satisfying $\mathbb{S}|x\rangle|y\rangle = |y\rangle|x\rangle$ for any pair $|x\rangle, |y\rangle \in \mathcal{H}$. Finally, the real coefficients α and β are given by

$$\alpha = \frac{1}{D_+}\mathrm{tr}(\mathbb{P}^{\mathrm{sym}}V),$$
$$\beta = \frac{1}{D_-}\mathrm{tr}(\mathbb{P}^{a\mathrm{sym}}V), \tag{C3}$$

where $D_\pm = D(D\pm 1)/2$ with $D = |\mathcal{H}|$.

We first consider the case in which both the Hilbert space $\mathcal{H}$ and the Hamiltonian H are fully general.

Lemma 7. Let $F_{\Omega,H}$ be defined as in Eq. (A15). Then, the following equality holds:

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{H})} F_{\Omega,H}(U) = \frac{2\mathrm{tr}(H^2)}{D^2-1}\sum_{i,j:\,p_i+p_j\neq 0}\frac{(p_i-p_j)^2}{p_i+p_j}. \tag{C4}$$

Proof.—We have the following sequence of equalities:

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{H})} F_{\Omega,H}(U) = 2\sum_{i,j:\,p_i+p_j\neq 0}\frac{(p_i-p_j)^2}{p_i+p_j}\left(\int_{\mathrm{SU}(\mathcal{H})}\mathrm{d}\mu(U)|\langle e_i|U^\dagger H U|e_j\rangle|^2\right) \tag{C5}$$

$$= 2\sum_{i,j:\,p_i+p_j\neq 0}\frac{(p_i-p_j)^2}{p_i+p_j}\left(\int_{\mathrm{SU}(\mathcal{H})}\mathrm{d}\mu(U)\mathrm{tr}[U\otimes UH\otimes HU^\dagger\otimes U^\dagger|e_i\rangle\langle e_j|\otimes|e_j\rangle\langle e_i|]\right) \tag{C6}$$

$$= \sum_{i,j:\,p_i+p_j\neq 0}\frac{(p_i-p_j)^2}{p_i+p_j}\mathrm{tr}\{[(\alpha_H+\beta_H)\mathbb{1}\otimes\mathbb{1}+(\alpha_H-\beta_H)\mathbb{S}]|e_i\rangle\langle e_j|\otimes|e_j\rangle\langle e_i|\}, \tag{C7}$$

where the third equality follows from Fact 3 for $V = H\otimes H$, and the real numbers α_H and β_H are given by

$$\alpha_H = \frac{1}{D_+}\mathrm{tr}(H\otimes H\mathbb{P}^{\mathrm{sym}}) = \frac{1}{2D_+}\mathrm{tr}(H^2),$$
$$\beta_H = \frac{1}{D_-}\mathrm{tr}(H\otimes H\mathbb{P}^{a\mathrm{sym}}) = -\frac{1}{2D_-}\mathrm{tr}(H^2). \tag{C8}$$

To obtain Eq. (C8), we also used the fact that $\mathrm{tr}(H) = 0$. Then, inserting Eq. (C8) into Eq. (C7) and using the identities

$$\mathrm{tr}(|e_i\rangle\langle e_j|\otimes|e_j\rangle\langle e_i|) = 0,$$
$$\mathrm{tr}(|e_i\rangle\langle e_j|\otimes|e_j\rangle\langle e_i|\mathbb{S}) = 1, \tag{C9}$$

one arrives at

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{H})} F_{\Omega,H}(U) = \frac{\mathrm{tr}(H^2)}{2}\left(\frac{1}{D_+}+\frac{1}{D_-}\right)\sum_{i,j:\,p_i+p_j\neq 0}\frac{(p_i-p_j)^2}{p_i+p_j}, \tag{C10}$$

which, by virtue of the definitions of $D_\pm$, leads us to Eq. (C4). ∎

The formula (C4) simplifies significantly for the case of pure states.

Remark 7. Let Ω_0 consist of pure states on $\mathcal{H}$. In this case, it is fairly easy to see that

$$\sum_{i,j:\,p_i+p_j\neq 0}\frac{(p_i-p_j)^2}{p_i+p_j} = 2(D-1), \tag{C11}$$

and consequently,

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{H})} F_{\Omega_0,H}(U) = \frac{4\mathrm{tr}(H^2)}{D+1}. \tag{C12}$$

By comparing Eqs. (C12) and (C4), one finds that the average QFI over any ensemble Ω of isospectral states can be easily related to the average QFI over pure states. Specifically, one has

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{H})} F_{\Omega,H}(U) = \frac{4\mathrm{tr}(H^2)}{D+1}\Lambda(\{p_j\}_j)$$
$$= \mathop{\mathbb{E}}_{U\sim\mu(\mathcal{H})} F_{\Omega_0,H}(U)\Lambda_\Omega, \tag{C13}$$

where

$$\Lambda(\{p_j\}_j) = \frac{1}{2(D-1)}\sum_{i,j:\,p_i+p_j\neq 0}\frac{(p_i-p_j)^2}{p_i+p_j}. \tag{C14}$$

Note that $\Lambda(\{p_j\}_j) = 1$ for pure states and $\Lambda(\{p_j\}_j) = 0$ for the maximally mixed state. Since, in general, the dependence on the spectrum in the above formula is quite complicated, it is desirable to have simple bounds on $\sum_{i,j:\,p_i+p_j\neq 0}\frac{(p_i-p_j)^2}{p_i+p_j}$. The following fact provides one such bound:

Fact 4. Let the numbers $p_1,\ldots,p_D$ satisfy $p_i\geq 0$ and $\sum_{i=1}^{D}p_i=1$. Then, the following inequality holds:

$$\sum_{i,j:\,p_i+p_j\neq 0}\frac{(p_i-p_j)^2}{p_i+p_j}\geq 2\left[D-\left(\sum_{i=1}^{D}\sqrt{p_i}\right)^2\right]$$
$$=2D\left[1-\mathcal{F}^2\left(\rho,\frac{\mathbb{1}}{D}\right)\right]. \tag{C15}$$

Proof.—First, by using the identity $(p_i-p_j)^2=(p_i+p_j)^2-4p_ip_j$, the left-hand side of the inequality (C15) can be rewritten as

$$\sum_{i,j:\,p_i+p_j\neq 0}\frac{(p_i-p_j)^2}{p_i+p_j}=\sum_{i,j:\,p_i+p_j\neq 0}\left[\frac{(p_i+p_j)^2}{p_i+p_j}-\frac{4p_ip_j}{p_i+p_j}\right], \tag{C16}$$

which, noting that the first sum in the above amounts to $2(D-1)$, can be rewritten as

$$\sum_{i,j:\,p_i+p_j\neq 0}\frac{(p_i-p_j)^2}{p_i+p_j}$$
$$=2\left((D-1)-\sum_{i,j:\,p_i+p_j\neq 0}\frac{2p_ip_j}{p_i+p_j}+1\right). \tag{C17}$$

To obtain the inequality in Eq. (C15), we apply the following well-known relation between the harmonic and geometric means,

$$\frac{2}{\frac{1}{p_i}+\frac{1}{p_j}}\leq\sqrt{p_ip_j}, \tag{C18}$$

to Eq. (C17). Then, to obtain the equality in Eq. (C15) and complete the proof, it suffices to notice that

$$\mathcal{F}\left(\rho,\frac{\mathbb{1}}{D}\right)=\mathrm{tr}\sqrt{\sqrt{\frac{\mathbb{1}}{D}}\rho\sqrt{\frac{\mathbb{1}}{D}}}=\frac{1}{\sqrt{D}}\mathrm{tr}\sqrt{\rho}=\frac{1}{\sqrt{D}}\sum_{i=1}^{D}\sqrt{p_i}\,. \tag{C19}$$

■

Remark 8. Note that the bound (C15) is tight. To be more precise, it is saturated for the maximally mixed state $\rho=\mathbb{1}/D$, for which both sides of the inequality (C15) simply vanish, and for pure states, for which they amount to $2(D-1)$.

2. Averages of QFI for N particles

We now discuss the average behavior of the QFI for ensembles consisting of states of distinguishable or bosonic particles (in the case when all particles evolve in the same manner under a local Hamiltonian). For the case of N distinguishable particles, we have

$$\mathcal{H}=\mathcal{H}_N=(\mathbb{C}^d)^{\otimes N}, \tag{C20}$$

where $\mathbb{C}^d$ is the Hilbert space of a single particle and N is the number of particles. Clearly, we have $D=|\mathcal{H}_N|=d^N$. The Hilbert space of N bosons in d modes is the completely symmetric subspace of $\mathcal{H}_N$,

$$\mathcal{H}=\mathcal{S}_N=\mathrm{span}_{\mathbb{C}}\{|\phi\rangle^{\otimes N}||\phi\rangle\in\mathbb{C}^d\}, \tag{C21}$$

of dimension $D=|\mathcal{S}_N|=\binom{N+d-1}{N}$. It will be convenient for us to use the orthonormal basis of $\mathcal{S}_N$ consisting of generalized Dicke states [103] (we also use them extensively in part of the Appendix, where we estimate the impact of particle losses on typical properties of QFI). Within the second quantization picture, $\mathcal{S}_N$ can be treated as a subspace of d-mode bosonic Fock space, and the generalized Dicke states are of the form

$$|\vec{k},N\rangle=\frac{\prod_{i=1}^{d}(a_i^\dagger)^{k_i}}{\sqrt{\prod_{i=1}^{d}k_i!}}|\Omega\rangle, \tag{C22}$$

where $|\Omega\rangle$ is the Fock vacuum, $a_i^\dagger$ are the standard creation operators and the vector $\vec{k}=(k_1,k_2,\ldots,k_d)$ consists of non-negative integers that count how many particles occupy each mode. Due to the fact that the number of particles is N, the vector $\vec{k}$ satisfies the normalization condition $|\vec{k}|\coloneqq\sum_{i=1}^{d}k_i=N$. Let us also notice that, in the particle picture, the Dicke states are given by

$$|\vec{k},N\rangle=\mathcal{N}(\vec{k},N)\mathbb{P}_{\mathrm{sym}}^{N}|\vec{k}\rangle, \tag{C23}$$

where $|\vec{k}\rangle$ is a vector from $(\mathbb{C}^d)^{\otimes N}$ given by $|\vec{k}\rangle=|1\rangle^{\otimes k_1}\otimes|2\rangle^{\otimes k_2}\otimes\ldots\otimes|d\rangle^{\otimes k_d}$, the constant $\mathcal{N}(\vec{k},N)$ is given by

$$\mathcal{N}(\vec{k},N)=\sqrt{\binom{N}{\vec{k}}} \tag{C24}$$

with

$$\binom{N}{\vec{k}}=\frac{N!}{\prod_{i=1}^{d}k_i!}, \tag{C25}$$

and by $\mathbb{P}_{\mathrm{sym}}^{N}$, we denote the orthonormal projector onto $\mathcal{S}_N\subset\mathcal{H}_N$.

The Hamiltonian used in the phase estimation is local and symmetric under the exchange of particles,

$$\begin{aligned} H &= H_N \\ &= h\otimes \mathbb{I}\otimes \ldots \otimes \mathbb{I} + \mathbb{I}\otimes h \otimes \mathbb{I}\otimes \ldots \otimes \mathbb{I} \\ &\quad + \cdots + \mathbb{I}\otimes \ldots \otimes \mathbb{I}\otimes h, \end{aligned} \tag{C26}$$

where h stands for the single-particle local Hamiltonian. In what follows, we assume, for simplicity, that $\mathrm{tr}(h)=0$. Note that the Hamiltonian H_N preserves the subspace $\mathcal{S}_N$.

Now, it follows from Eqs. (C4) and (C12) that the average behavior of the QFI on states supported on the subspace $\mathcal{W}\subset\mathcal{H}$ is dictated by the value of $\mathrm{tr}_{\mathcal{W}}(H_N^2)$. In the following lemma, we compute the latter in the cases $\mathcal{W}=\mathcal{H}_N$ and $\mathcal{W}=\mathcal{S}_N$.

Lemma 8. Let Hamiltonian H be given by Eq. (C26). Then, the following relations,

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{H}_N)} F_{\Omega,H} = \frac{4N\mathrm{tr}(h^2)}{d}\frac{|\mathcal{H}_N|}{|\mathcal{H}_N|+1}\Lambda(\{p_j\}_j) \tag{C27}$$

and

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)} F_{\Omega,H} = \frac{4N(N+d)\mathrm{tr}(h^2)}{d(d+1)}\frac{|\mathcal{S}_N|}{|\mathcal{S}_N|+1}\Lambda(\{p_j\}_j) \tag{C28}$$

are true. For pure qubits, Eq. (C28) simplifies to

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)} F_{(1,0,\ldots,0),H} = \frac{2}{3}N(N+1)\mathrm{tr}(h^2). \tag{C29}$$

Remark 9. The qualitative meaning of the above lemma is twofold. First, it shows that for uniformly distributed isospectral states from $\mathcal{H}_N$, the scaling of the QFI on average is at most linear in the number of particles N, and, second, it proves that for random pure symmetric states, the average QFI scales quadratically with N, both for fixed local Hamiltonian h and local dimension d. Thus, for symmetric states the average QFI attains the Heisenberg limit. This behavior still holds for random isospectral density matrices, provided their spectrum is sufficiently pure, with the "degree of purity" quantified by $\Lambda(\{p_j\}_j)$ defined by Eq. (C14).

Proof.—We start from the proof of Eq. (C27). Using the fact that the local Hamiltonian h is traceless, one obtains

$$\begin{aligned} \mathrm{tr}_{\mathcal{H}_N}(H^2) &= \sum_{i=1}^{N}\mathrm{tr}_{\mathcal{H}_N}[(h^{(i)})^2] \\ &= \sum_{i=1}^{N}\mathrm{tr}(h^2)d^{N-1} \\ &= N\mathrm{tr}(h^2)\frac{|\mathcal{H}_N|}{d}. \end{aligned} \tag{C30}$$

Inserting the above into Eq. (C4) (note that here $\mathcal{H}=\mathcal{H}_N$), we arrive at Eq. (C27).

The proof of Eq. (C28) is more involved, as it requires the computation of $\mathrm{tr}_{\mathcal{S}_N}(H^2)$. The final result reads

$$\mathrm{tr}_{\mathcal{S}_N}(H^2) = \frac{N(N+d)\mathrm{tr}(h^2)}{d(d+1)}|\mathcal{S}_N|, \tag{C31}$$

which when plugged into Eq. (C4) yields Eq. (C28).

To explicitly determine Eq. (C31), let us choose the basis $\{|i\rangle\}_{i=1}^d$ of the single-particle space as the eigenbasis of the local Hamiltonian h. Thus, we have $h|i\rangle=\lambda_i|i\rangle$ for $i=1,\ldots,d$. The corresponding generalized Dicke states [see Eq. (C23)] satisfy

$$H_N|\vec{k},N\rangle = (\vec{k}\,\vec{\lambda})|\vec{k},N\rangle, \tag{C32}$$

where $\vec{\lambda}=(\lambda_1,\ldots,\lambda_d)$ is the vector of eigenvalues of h and the standard inner product in $\mathbb{R}^d$. Now, Eq. (C32), together with the fact that the generalized Dicke states form a basis of $\mathcal{S}_N$, allows us to write

$$\mathrm{tr}_{\mathcal{S}_N}(H^2) = \sum_{\vec{k}:|\vec{k}|=N}(\vec{k}\,\vec{\lambda})^2 \tag{C33}$$

$$= \sum_{\vec{k}:|\vec{k}|=N}\sum_{i=1}^{d}\lambda_i^2k_i^2 + \sum_{\vec{k}:|\vec{k}|=N}\sum_{\substack{i,j=1\\ i\neq j}}^{d}(\lambda_i\lambda_j)(k_ik_j), \tag{C34}$$

where, to obtain the second equality, we explicitly squared all scalar products appearing under the sum. From the symmetry, we have

$$\begin{aligned} \sum_{\vec{k}:|\vec{k}|=N} k_i^2 &= \sum_{\vec{k}:|\vec{k}|=N} k_{i'}^2, \\ \sum_{\vec{k}:|\vec{k}|=N} k_ik_j &= \sum_{\vec{k}:|\vec{k}|=N} k_{i'}k_{j'}, \end{aligned} \tag{C35}$$

for all i, i' and for all pairs of different indices (i,j) and (i',j'). As a result, Eq. (C33) simplifies to

$$\mathrm{tr}_{\mathcal{S}_N}(H^2) = \sum_{\vec{k}:|\vec{k}|=N}\left(\left[\sum_{i=1}^{d}\lambda_i^2\right]k_1 + \left[\sum_{\substack{i,j=1\\ i\neq j}}^{d}\lambda_i\lambda_j\right]k_1k_2\right). \tag{C36}$$

The fact that h is traceless yields

$$\mathrm{tr}(h^2) = \sum_{i=1}^{d}\lambda_i^2 = -\sum_{\substack{i,j=1\\ i\neq j}}^{d}\lambda_i\lambda_j. \tag{C37}$$

Moreover, because of the condition $k_1+\cdots+k_d=N$, we have

$$\sum_{\vec{k}:|\vec{k}|=N} (k_1+\cdots+k_d)^2 = |\mathcal{S}_N|N^2. \tag{C38}$$

By exploiting the identities (C35), the left-hand side of the above equation can be rewritten as

$$\sum_{\vec{k}:|\vec{k}|=N} (k_1+\cdots+k_d)^2 = d\sum_{\vec{k}:|\vec{k}|=N} k_1^2 + d(d-1)\sum_{\vec{k}:|\vec{k}|=N} k_1k_2. \tag{C39}$$

As a result, one obtains

$$d(d-1)\sum_{\vec{k}:|\vec{k}|=N} k_1k_2 = |\mathcal{S}_N|N^2 - d\sum_{\vec{k}:|\vec{k}|=N} k_1^2. \tag{C40}$$

Using Eqs. (C36), (C37), and (C40), we finally arrive at

$$\mathrm{tr}_{\mathcal{S}_N}(H^2) = \mathrm{tr}(h^2)\left[(d+1)\left(\sum_{\vec{k}:|\vec{k}|=N} k_1^2\right) - N^2|\mathcal{S}_N|\right]. \tag{C41}$$

We compute the sum $\sum_{\vec{k}:|\vec{k}|=N} k_1^2$ by noting that

$$\#(\{\vec{k}||\vec{k}|=N, k_1=i\}) = \binom{N-i+d-2}{N-i}, \tag{C42}$$

where #() denotes the number of elements of a discrete set. The above equation follows from the fact that the number of elements of the set $\{\vec{k}||\vec{k}|=N, k_1=i\}$ is the same as the dimension of the Hilbert space of $N-i$ bosons in $d-1$ modes. Consequently, we get

$$\sum_{\vec{k}:|\vec{k}|=N} k_1^2 = \sum_{i=0}^{N} i^2\binom{N-i+d-2}{N-i} = \frac{N(2N+d-1)}{d(d+1)}|\mathcal{S}_N|. \tag{C43}$$

Inserting the above expression into Eq. (C41) yields Eq. (C31). The equality (C43) can be proven using standard combinatorial identities. Below, we sketch its proof for completeness. First, by virtue of the diagonal sum property of binomial coefficients [105], we have that

$$\binom{N-i+a}{N-i} = \sum_{k=0}^{N-i}\binom{a+k-1}{k}, \tag{C44}$$

where a is an arbitrary integer. Inserting Eq. (C44) (with $a=d=2$) to the left-hand side of Eq. (C43), we get

$$\sum_{i=0}^{N} i^2\binom{N-i+d-2}{N-i} = \sum_{i=0}^{N}\sum_{k=0}^{N-i} i^2\binom{d-2+k-1}{k} = \sum_{k=0}^{N}\left(\sum_{i=0}^{N-k} i^2\right)\binom{d-2+k-1}{k}. \tag{C45}$$

The sum $\sum_{i=0}^{N-k} i^2$ is a polynomial of degree 3 in k and can be easily computed. Therefore, in order to finish the computation, it suffices to know the moments

$$\sum_{k=0}^{N} k^j\binom{a-1+k}{k}, \tag{C46}$$

for the powers $j=1,2,3$. These can be found, for instance, on page 5 of Ref. [106]. ■

Remark 10. The most demanding part in the proof of Lemma C 2 is the computation of $\mathrm{tr}_{\mathcal{S}_N}(H^2)$, which can be simplified greatly by the use of group theoretic methods. This should allow one to perform an analogous analysis for other irreducible representations of the group SU(d), for instance, for the fermionic subspace of $\mathcal{H}_N$.

3. Average QFI for bosons with particle losses

Let $\rho' = \mathrm{tr}_k(\rho)$ be a mixed symmetric state on $N-k$ particles arising from tracing out k particles of some N-partite state $\rho \in \mathcal{D}(\mathcal{S}_N)$. Our aim in this section is to bound the average of the QFI over mixed states created in the above way, where ρ is a random isospectral state acting on $\mathcal{S}_N$. Recall that we are interested in the standard context of quantum metrology; i.e., the Hamiltonian H encoding the phase φ is given by Eq. (C26).

Lemma 9. Let $\rho \in \mathcal{D}(\mathcal{S}_N)$ be a state of N bosons with single-particle d-dimensional Hilbert space $\mathcal{H}_l$ and the spectrum $\{p_1, \ldots, p_{|\mathcal{S}_N|}\}$. Let us fix the local Hamiltonian h and a non-negative integer k. Then, the following inequality holds:

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)} F(\mathrm{tr}_k(U\rho U^\dagger), H^{[N-k]}) \geq 2\frac{(N-k)(N+d)}{(d+1)(d+k)}\frac{|\mathcal{S}_N|(|\mathcal{S}_N|\mathrm{tr}\rho_N^2 - 1)\mathrm{tr}h^2}{|\mathcal{S}_k|(|\mathcal{S}_N|^2-1)}. \tag{C47}$$

Proof.—Denoting $\sigma_U = \mathrm{tr}_k(U\rho U^\dagger)$, we notice that the inequality (B9) allows one to lower bound the QFI as

$$F(\sigma_U, H_{N-k}) \geq \|[\sigma_U, H_{N-k}]\|_{\mathrm{HS}}^2 \tag{C48}$$

$$= 2(\mathrm{tr}_{\mathcal{S}_{N-k}}\{\sigma_U^2 H_{N-k}^2\} - \mathrm{tr}_{\mathcal{S}_{N-k}}\{(\sigma_U H_{N-k})^2\}), \tag{C49}$$

where due to the fact that σ_U is symmetric, the trace is taken over the symmetric subspace $\mathcal{S}_{N-k}$. For the same reason, we can cut the Hamiltonian to the symmetric subspace on which it acts as

$$H_{N-k}|_{\mathcal{S}_{N-k}} \coloneqq \mathbb{P}_{\rm sym}^{N-k} H_{N-k} \mathbb{P}_{\rm sym}^{N-k} = \sum_{\vec{n}} \lambda_{\vec{n}}^{(N-k)} |\vec{n}, N-k\rangle\langle\vec{n}, N-k|, \qquad (C50)$$

where, as before, $|\vec{n}, N-k\rangle$ are $(N-k)$-partite generalized Dicke states and $\vec{n} = (n_0, \ldots, n_{d-1})$ is a vector of nonnegative integers such that $n_0 + \cdots + n_{d-1} = N-k$, and $\lambda_{\vec{n}}^{(N-k)}$ are the eigenvalues of H_{N-k}. By abuse of notation, in what follows, we denote both the Hamiltonian and its symmetric part (C50) by H_{N-k}.

Using the swap operator introduced in Fact 3 for $\mathcal{H} = \mathcal{S}_{N-k}$ and the fact that $\mathrm{tr}(\mathbb{S} A \otimes B) = \mathrm{tr}(AB)$ holds for any pair of operators acting on $\mathcal{S}_{N-k}$, we can rewrite Eq. (C48) as

$$F(\sigma_U, H_{N-k}) \geq 2\{\mathrm{tr}[(H_{N-k}\sigma_U \otimes \sigma_U H_{N-k})\mathbb{S}_{N-k}] - \mathrm{tr}[(\sigma_U H_{N-k} \otimes \sigma_U H_{N-k})\mathbb{S}_{N-k}]\} \qquad (C51)$$

$$= 2\{\mathrm{tr}[(\sigma_U \otimes \sigma_U)(\mathbb{P}_{\rm sym}^{N-k} \otimes H_{N-k}^2 - H_{N-k} \otimes H_{N-k})\mathbb{S}_{N-k}]\}, \qquad (C52)$$

where to obtain the second line we used the fact that σ_U acts on $\mathcal{S}_{N-k}$ and that $\mathbb{S}_{N-k}^2 = \mathbb{P}_{\rm sym}^{N-k} \otimes \mathbb{P}_{\rm sym}^{N-k}$, and, for simplicity, we dropped the subscript $\mathcal{S}_{N-k} \otimes \mathcal{S}_{N-k}$ in the trace.

Exploiting the fact that the symmetric projector $\mathbb{P}_{\rm sym}^N$ is diagonal in the Dicke basis, that is,

$$\mathbb{P}_{\rm sym}^N = \sum_{\vec{p}} |\vec{p}, N\rangle\langle\vec{p}, N|, \qquad (C53)$$

the representation of the Hamiltonian in Eq. (C48), and the definition of the swap operator, one arrives at the following formula:

$$(\mathbb{P}_{\rm sym}^{N-k} \otimes H_{N-k}^2 - H_{N-k} \otimes H_{N-k})\mathbb{S}_{N-k} = \sum_{\vec{n},\vec{m}} (\lambda_{\vec{n}}^2 - \lambda_{\vec{n}}\lambda_{\vec{m}}) |\vec{m}, N-k\rangle\langle\vec{n}, N-k| \otimes |\vec{n}, N-k\rangle\langle\vec{m}, N-k|, \qquad (C54)$$

which, when plugged into Eq. (C51), gives

$$F(\sigma_U, H_{N-k}) \geq 2\sum_{\vec{n},\vec{m}} (\lambda_{\vec{n}}^2 - \lambda_{\vec{n}}\lambda_{\vec{m}}) \mathrm{tr}[(\sigma_U \otimes \sigma_U)|\vec{m}, N-k\rangle \times \langle\vec{n}, N-k| \otimes |\vec{n}, N-k\rangle\langle\vec{m}, N-k|]. \qquad (C55)$$

We are now ready to lower bound the average $\mathbb{E}_{U\sim\mu(\mathcal{S}_N)} F(\sigma_U, H_{N-k})$. Using the fact that $\sigma_U = \mathrm{tr}_k(U\rho U^\dagger)$ and that $U\rho U^\dagger$ is symmetric, we obtain from inequality (C55) that

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)} F(\sigma_U, H_{N-k}) \geq 2\sum_{\vec{n},\vec{m}} (\lambda_{\vec{n}}^2 - \lambda_{\vec{n}}\lambda_{\vec{m}}) \int_{\mathrm{SU}(\mathcal{S}_N)} d\mu(U) \mathrm{tr}[(U\rho U^\dagger \otimes U\rho U^\dagger)|\vec{m}, N-k\rangle\langle\vec{n}, N-k| \otimes \mathbb{P}_{\rm sym}^k \otimes |\vec{n}, N-k\rangle\langle\vec{m}, N-k| \otimes \mathbb{P}_{\rm sym}^k], \qquad (C56)$$

where now the trace is performed over $\mathcal{S}_N \otimes \mathcal{S}_N$. Let us focus for a moment on the state

$$\int_{\mathrm{SU}(\mathcal{S}_N)} d\mu(U)(U\rho U^\dagger \otimes U\rho U^\dagger). \qquad (C57)$$

It follows from Fact 3 (for $\mathcal{H} = \mathcal{S}_N$) that after performing the integration, the above state assumes the following form:

$$\int_{\mathrm{SU}(\mathcal{S}_N)} d\mu(U)(U\rho U^\dagger \otimes U\rho U^\dagger) = \alpha\mathbb{P}_{\rm sym\wedge sym} + \beta\mathbb{P}_{\rm as\wedge as}. \qquad (C58)$$

For completeness, let us recall that $\mathbb{P}_{\rm sym\wedge sym}$ and $\mathbb{P}_{\rm as\wedge as}$ are the projectors onto the symmetric and antisymmetric subspaces of $\mathcal{S}_N \otimes \mathcal{S}_N$, respectively, and they are given by

$$\mathbb{P}_{\rm sym\wedge sym} = \frac{1}{2}(\mathbb{P}_{\rm sym}^N \otimes \mathbb{P}_{\rm sym}^N + \mathbb{S}_N), \quad \mathbb{P}_{\rm as\wedge as} = \frac{1}{2}(\mathbb{P}_{\rm sym}^N \otimes \mathbb{P}_{\rm sym}^N - \mathbb{S}_N). \qquad (C59)$$

Moreover, the real coefficients α and β are explicitly given by

$$\alpha = \frac{1}{2D_+(\mathcal{S}_N)}(1 + \mathrm{tr}\rho^2), \quad \beta = \frac{1}{2D_-(\mathcal{S}_N)}(1 - \mathrm{tr}\rho^2), \qquad (C60)$$

where $D_\pm(\mathcal{S}_N) = |\mathcal{S}_N|(|\mathcal{S}_N| \pm 1)/2$.

Plugging Eq. (C58) into Eq. (C56) and using Eq. (C59), one arrives at

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)} F(\sigma_U, H_{N-k}) \geq \sum_{\vec{n},\vec{m}} (\lambda_{\vec{n}}^2 - \lambda_{\vec{n}}\lambda_{\vec{m}}) \{ (\alpha+\beta) |\mathrm{tr}(\mathbb{P}_{\mathrm{sym}}^N |\vec{m}, N-k\rangle\langle\vec{n}, N-k| \otimes \mathbb{P}_{\mathrm{sym}}^k)|^2 + (\alpha-\beta)\mathrm{tr}[\mathbb{S}_N(|\vec{m}, N-k\rangle\langle\vec{n}, N-k| \otimes \mathbb{P}_{\mathrm{sym}}^k \otimes |\vec{n}, N-k\rangle\langle\vec{m}, N-k| \otimes \mathbb{P}_{\mathrm{sym}}^k)]\}. \tag{C61}$$

The right-hand side of this inequality can significantly be simplified if one observes that the first trace under the curly brackets is nonzero only if $\vec{m} = \vec{n}$, giving

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)} F(\sigma_U, H_{N-k}) \geq (\alpha-\beta) \sum_{\vec{n},\vec{m}} (\lambda_{\vec{n}}^2 - \lambda_{\vec{n}}\lambda_{\vec{m}}) \times \mathrm{tr}[\mathbb{S}_N(|\vec{m}, N-k\rangle\langle\vec{n}, N-k| \otimes \mathbb{P}_{\mathrm{sym}}^k \otimes |\vec{n}, N-k\rangle\langle\vec{m}, N-k| \otimes \mathbb{P}_{\mathrm{sym}}^k)]\}. \tag{C62}$$

Our aim now is to compute the remaining trace, which for further purposes, we denote $T_{\vec{m},\vec{n}}$. We use the fact that the projector $\mathbb{P}_{\mathrm{sym}}^k$ can be written as in Eq. (C53), which, together with the following identity,

$$\mathbb{P}_{\mathrm{sym}}^N |\vec{n}, N-k\rangle|\vec{p}, k\rangle = \frac{\sqrt{\binom{N-k}{\vec{n}}}\sqrt{\binom{k}{\vec{p}}}}{\sqrt{\binom{N}{\vec{n}+\vec{p}}}} |\vec{n}+\vec{p}, N\rangle, \tag{C63}$$

allows us to express $T_{\vec{m},\vec{n}}$ as

$$T_{\vec{m},\vec{n}} = \sum_{\vec{o}} \frac{\binom{N-k}{\vec{m}}\binom{k}{\vec{o}}}{\binom{N}{\vec{m}+\vec{o}}} \frac{\binom{N-k}{\vec{n}}\binom{k}{\vec{o}}}{\binom{N}{\vec{n}+\vec{o}}}. \tag{C64}$$

This gives

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)} F(\sigma_U, H_{N-k}) \geq (\alpha-\beta) \sum_{\vec{n},\vec{m}} (\lambda_{\vec{n}}^2 - \lambda_{\vec{n}}\lambda_{\vec{m}}) \times \sum_{\vec{o}} \frac{\binom{N-k}{\vec{m}}\binom{k}{\vec{o}}}{\binom{N}{\vec{m}+\vec{o}}} \frac{\binom{N-k}{\vec{n}}\binom{k}{\vec{o}}}{\binom{N}{\vec{n}+\vec{o}}} \tag{C65}$$

$$= 2\frac{|\mathcal{S}_N|\mathrm{tr}\rho^2 - 1}{|\mathcal{S}_N|(|\mathcal{S}_N|^2 - 1)} (L_{N,k} - L'_{N,k}), \tag{C66}$$

where we used the explicit expressions for α and β and denoted

$$L_{N,k} = \sum_{\vec{o}} \left[\sum_{\vec{m}} \frac{\binom{N-k}{\vec{m}}\binom{k}{\vec{o}}}{\binom{N}{\vec{m}+\vec{o}}} \right] \left[\sum_{\vec{n}} \lambda_{\vec{n}}^2 \frac{\binom{N-k}{\vec{n}}\binom{k}{\vec{o}}}{\binom{N}{\vec{n}+\vec{o}}} \right],$$

$$L_{N,k}' = \sum_{\vec{o}} \left[\sum\nolimits_{\vec{m}} \lambda_{\vec{m}} \frac{\binom{N-k}{\vec{m}}\binom{k}{\vec{o}}}{\binom{N}{\vec{m}+\vec{o}}} \right]^2. \tag{C67}$$

We now compute each sum separately. To this end, let us first notice that it follows from Eq. (C63) that

$$\sum_{\vec{m}} \frac{\binom{N-k}{\vec{m}}\binom{k}{\vec{o}}}{\binom{N}{\vec{m}+\vec{o}}} = \sum_{\vec{m}} \mathrm{tr}[\mathbb{P}_{\mathrm{sym}}^N |\vec{m}, N-k\rangle\langle\vec{m}, N-k| \otimes |\vec{o}, k\rangle\langle\vec{o}, k|] \tag{C68}$$

$$= \mathrm{tr}[\mathbb{P}_{\mathrm{sym}}^N(\mathbb{P}_{\mathrm{sym}}^{N-k} \otimes |\vec{o}, k\rangle\langle\vec{o}, k|)] \tag{C69}$$

$$= \frac{|\mathcal{S}_N|}{|\mathcal{S}_k|}, \tag{C70}$$

where to get the second equality, we used Eq. (C53), while to obtain the third one, we used the fact that the partial trace of $\mathbb{P}_{\mathrm{sym}}^N$ over $N-k$ subsystems is given by

$$\mathrm{tr}_{N-k}(\mathbb{P}_{\mathrm{sym}}^N) = \frac{|\mathcal{S}_N|}{|\mathcal{S}_k|} \mathbb{P}_{\mathrm{sym}}^k. \tag{C71}$$

With the aid of formula Eq. (C68), we can write $L_{N,k}$ as

$$L_{N,k} = \frac{|\mathcal{S}_N|}{|\mathcal{S}_k|} \sum_{\vec{n}} \lambda_{\vec{n}}^2 \sum_{\vec{o}} \frac{\binom{N-k}{\vec{n}}\binom{k}{\vec{o}}}{\binom{N}{\vec{n}+\vec{o}}}. \tag{C72}$$

Then, exploiting Eqs. (C53) and Eq. (C63) and the form of the Hamiltonian H_{N-k}, this equation can be further rewritten as

$$L_{N,k} = \frac{|\mathcal{S}_N|}{|\mathcal{S}_k|} \mathrm{tr}[\mathbb{P}_{\mathrm{sym}}^N(H_{N-k}^2 \otimes \mathbb{P}_{\mathrm{sym}}^k)] \tag{C73}$$

$$= \frac{|\mathcal{S}_N|^2}{|\mathcal{S}_k||\mathcal{S}_{N-k}|} \mathrm{tr}_{\mathcal{S}_{N-k}}(H_{N-k}^2), \tag{C74}$$

where the second equality stems from Eq. (C71).

To compute $L'_{N,k}$, we follow more or less the same strategy. First, using Eqs. (C63) and (C50), we can rewrite it as

$$L'_{N,k} = \sum_{\vec{o}} \{\mathrm{tr}[\mathbb{P}^N_{\mathrm{sym}}(H_{N-k} \otimes |\vec{o}, k\rangle\langle\vec{o}, k|)]\}^2. \quad \text{(C75)}$$

Then, we use the fact that in the full Hilbert space $(\mathbb{C}^d)^{\otimes(N-k)}$, H_{N-k} assumes the form given in Eq. (C26), which gives

$$\begin{aligned}
&\mathrm{tr}[\mathbb{P}^N_{\mathrm{sym}}(H_{N-k} \otimes |\vec{o}, k\rangle\langle\vec{o}, k|)] \\
&\quad = (N-k)\mathrm{tr}[\mathbb{P}^N_{\mathrm{sym}}(h \otimes \mathbb{P}^{N-k-1}_{\mathrm{sym}} \otimes |\vec{o}, k\rangle\langle\vec{o}, k|)] \quad \text{(C76)} \\
&\quad = (N-k)\frac{|\mathcal{S}_N|}{|\mathcal{S}_{k+1}|}\mathrm{tr}[\mathbb{P}^{k+1}_{\mathrm{sym}}(h \otimes |\vec{o}, k\rangle\langle\vec{o}, k|)], \quad \text{(C77)}
\end{aligned}$$

where the second line follows from Eq. (C71). To compute the remaining trace, we expand h in its eigenbasis as $h = \sum_{n=0}^{d-1} \xi_n |n\rangle\langle n|$ (where ξ_i are the eigenvalues of h), which can also be written using the "mode representation" as

$$h = \sum_{\vec{n}} \xi_{\vec{n}} |\vec{n}\rangle\langle\vec{n}|, \quad \text{(C78)}$$

where $\vec{n} = (i_0, \ldots, i_{d-1})$ is now a d-dimensional vector whose components are such that $n_i = 0,\ 1$ and $n_0 + \cdots + n_{d-1} = 1$. In this representation, a number $n = i \in \{0, \ldots, d-1\}$ is represented by a vector $\vec{n}$ whose ith component $n_i = 1$, and the remaining ones are zero. Using Eq. (C63), one obtains

$$\begin{aligned}
\mathrm{tr}[\mathbb{P}^{k+1}_{\mathrm{sym}}(h \otimes |\vec{o}, k\rangle\langle\vec{o}, k|)] &= \sum_{\vec{n}} \xi_{\vec{n}} \frac{\binom{1}{\vec{n}}\binom{k}{\vec{o}}}{\binom{k+1}{\vec{n}+\vec{o}}} \\
&= \sum_{\vec{n}} \xi_{\vec{n}} \frac{\binom{k}{\vec{o}}}{\binom{k+1}{\vec{n}+\vec{o}}}, \quad \text{(C79)}
\end{aligned}$$

where the summation is taken over vectors $\vec{n}$ specified above (there is d such vectors). The second equality straightforwardly stems from the fact that $\binom{1}{\vec{n}} = 1$. We then exploit the fact that $\binom{k+1}{\vec{n}+\vec{o}} = \frac{k+1}{o_n+1}\binom{k}{\vec{o}}$ and the assumption that $\mathrm{tr}h = 0$ to get

$$\begin{aligned}
\mathrm{tr}[\mathbb{P}^{k+1}_{\mathrm{sym}}(h \otimes |\vec{o}, k\rangle\langle\vec{o}, k|)] &= \frac{1}{k+1}\sum_{n=0}^{d-1} \xi_n o_n \\
&= \frac{1}{k+1}\lambda^{(k)}_{\vec{o}}, \quad \text{(C80)}
\end{aligned}$$

where we recall that $\lambda^{(k)}_{\vec{o}}$ is the eigenvalue of the k-partite Hamiltonian H_k [compare Eq. (C50)]. Combining the above identity with Eqs. (C76) and (C75), one finds that

$$L'_{N,k} = \left(\frac{N-k}{k+1}\frac{|\mathcal{S}_N|}{|\mathcal{S}_{k+1}|}\right)^2 \mathrm{tr}_{\mathcal{S}_k}(H_k^2). \quad \text{(C81)}$$

Plugging Eqs. (C73) and (C81) into Eq. (C56), one eventually finds that the average QFI is lower bounded as

$$\begin{aligned}
&\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)} F(\sigma_U, H_{N-k}) \\
&\quad \geq 2\frac{|\mathcal{S}_N|}{|\mathcal{S}_k|}\frac{|\mathcal{S}_N|\mathrm{tr}\rho_N^2 - 1}{|\mathcal{S}_N|^2 - 1}\frac{(N-k)(N+d)}{(d+1)(d+k)}\mathrm{tr}h^2. \quad \text{(C82)}
\end{aligned}$$

■

Remark 11. It is worth mentioning that using similar techniques, one can also provide an upper bound on the average QFI for bosons in the case of particle losses. To be more precise, in what follows, we will derive such a bound for multiqubit states. As the QFI is upper bounded by the variance, one has

$$\begin{aligned}
F_Q(\sigma_U, H_{N-k}) &\leq 4\Delta^2_{\sigma_U} H_{N-k} \\
&= 4\{\mathrm{tr}(\sigma_U H_{N-k}^2) - [\mathrm{tr}(\sigma_U H_{N-k})]^2\} \\
&\leq 4\mathrm{tr}(\sigma_U H_{N-k}^2). \quad \text{(C83)}
\end{aligned}$$

Then using the fact that the right-hand side can be rewritten as $\mathrm{tr}(\sigma_U H_{N-k}^2) = \mathrm{tr}[\rho(H_{N-k}^2 \otimes \mathbb{P}^k_{\mathrm{sym}})]$ and that

$$\int_{\mathrm{SU}(\mathcal{S}_N)} \mathrm{d}\mu(U) U\rho U^\dagger = \frac{\mathbb{P}^N_{\mathrm{sym}}}{N+1}, \quad \text{(C84)}$$

one obtains

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)} F(\sigma_U, H_{N-k}) \leq \frac{4}{N+1}\mathrm{tr}[\mathbb{P}^N_{\mathrm{sym}}(H_{N-k}^2 \otimes \mathbb{P}^k_{\mathrm{sym}})]. \quad \text{(C85)}$$

With the aid of Eqs. (C71) and (C31), we eventually get

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)} F(\sigma_U, H_{N-k}) \leq \frac{1}{3}(N-k)(N-k+2). \quad \text{(C86)}$$

Notice that for $k = 0$, this bound gives $N(N+2)/3$, which differs from the exact value for qubits by a factor linear in N. In general, however, this bound is not very informative because even for significant particle losses, e.g., $k = \eta N$ with $0 < \eta < 1$, the right-hand side of Eq. (C86) scales quadratically with N.

4. Average FI of random two-mode bosonic states in the interferometric setup

In this part, we study the interferometric setup introduced in Sec. VII and depicted in Fig. 2. Recall that the classical Fisher information (FI) associated with such a measurement scheme is given by

$$F_{\rm cl}(\{p_{n|\varphi}(\psi)\}) = \sum_{n=0}^{N} \frac{\mathrm{tr}(\mathrm{i}[\Pi_n^N, \hat{J}_z]\psi(\varphi))^2}{\mathrm{tr}(\Pi_n^N \psi(\varphi))}, \qquad \text{(C87)}$$

where by $\hat{J}_\alpha \coloneqq \frac{1}{2}\sum_{i=0}^{N} \sigma_\alpha^{(i)}$ $(\alpha = \{x, y, z\})$, we denote the angular momentum operators, $\Pi_n^N = \hat{B} D_n^N \hat{B}^\dagger$, where $\hat{B} \coloneqq \exp(-\mathrm{i}\pi \hat{J}_x/2)$, $D_n^N = |D_n^N\rangle\langle D_n^N|$, and $|D_n^N\rangle \coloneqq |n, N-n\rangle$ are the projections onto the Dicke states propagated through a balanced beam splitter, and $\psi(\varphi) \coloneqq \exp(-\mathrm{i}\hat{J}_z\varphi)\psi \exp(\mathrm{i}\hat{J}_z\varphi)$, with ψ some pure state in $\mathcal{S}_N$ with $d = 2$ modes.

Similarly as in Theorem 4 of Sec. VII, after fixing ψ_N to be a particular pure state on $\mathcal{S}_N$, we can then define

$$F_{\rm cl}(U, \varphi) \coloneqq F_{\rm cl}(\{p_{n|\varphi}(U\psi_N U^\dagger)\}, \qquad \text{(C88)}$$

where $U \in \mathrm{SU}(\mathcal{S}_N)$ and $\varphi \in [0, 2\pi]$.

Lemma 10. Let $F_{\rm cl}(U, \varphi)$ be defined as above. Then, the following inequalities hold:

$$c_- N^2 \leq \mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)} F_{\rm cl}(U, \varphi) \leq c_+ N^2 + N, \qquad \text{(C89)}$$

where

$$c_- = \frac{1}{36} - \frac{4}{3\mathrm{e}^5} \approx 0.0244,$$
$$c_+ = -\frac{5}{6} + \frac{3}{\mathrm{e}} \approx 0.270. \qquad \text{(C90)}$$

Proof.—The main difficulty in the proof comes from the fact that $F_{\rm cl}(U, \varphi)$ is a complicated, nonlinear function of U. Let us first note that by using the relation $\hat{B}\mathrm{e}^{-\mathrm{i}\hat{J}_z\varphi}\hat{B}^\dagger = \mathrm{e}^{\mathrm{i}\hat{J}_y\varphi}$, it is possible to rewrite the FI in Eq. (C87) as

$$F_{\rm cl}(\{p_{n|\varphi}(\psi)\}) = \sum_{n=0}^{N} \frac{\mathrm{tr}(\mathrm{i}[D_n^N, \hat{J}_y]\tilde{\psi}(\varphi))^2}{\mathrm{tr}(D_n^N \tilde{\psi}(\varphi))}, \qquad \text{(C91)}$$

where $\tilde{\psi}(\varphi) = \exp(\mathrm{i}\varphi\hat{J}_y)\psi \exp(-\mathrm{i}\varphi\hat{J}_y)$. Let us introduce the auxiliary notation

$$f_n(U, \varphi) = \{\mathrm{tr}(\mathrm{i}[D_n^N, \hat{J}_y] \exp(\mathrm{i}\varphi\hat{J}_y) U\psi U^\dagger \exp(-\mathrm{i}\varphi\hat{J}_y))\}^2, \qquad \text{(C92)}$$

$$g_n(U, \varphi) = \mathrm{tr}(D_n^N \exp(\mathrm{i}\varphi\hat{J}_y) U\psi U^\dagger \exp(-\mathrm{i}\varphi\hat{J}_y)). \qquad \text{(C93)}$$

Using the above formulas, we obtain the compact expression for $F_{\rm cl}(U, \varphi)$,

$$F_{\rm cl}(U, \varphi) = \sum_{n=0}^{N} \frac{f_n(U, \varphi)}{g_n(U, \varphi)}. \qquad \text{(C94)}$$

In what follows, we will make use of the inequality

$$f_n(U, \varphi) \leq N^2 g_n(U, \varphi)^2, \qquad \text{(C95)}$$

which follows directly from Eq. (A50) applied to the considered setting. In order to obtain bounds on the average $\mathbb{E}_{U\sim\mu(\mathcal{S}_N)} F_{\rm cl}(U, \varphi)$, we use the following subsets of the $\mathrm{SU}(\mathcal{S}_N)$,

$$\mathcal{G}_{+,\alpha}^n = \{U \in \mathrm{SU}(\mathcal{S}_n) | g_n(U, \varphi) \geq \alpha\}, \qquad \text{(C96)}$$

$$\mathcal{G}_{-,\alpha}^n = \{U \in \mathrm{SU}(\mathcal{S}_n) | g_n(U, \varphi) \leq \alpha\}, \qquad \text{(C97)}$$

where $n = 0, \ldots, N$ and $\alpha \in [0, 1]$. Because of the unitary invariance of the Haar measure and the fact that projectors D_n^N have rank one, the distribution of the random variable $g_n(U, \varphi)$ is identical to the distribution of the random variable $X(V) = \mathrm{tr}(\psi V \psi V^\dagger)$, where V is the Haar distributed unitary on $\mathbb{C}^{N+1}$ and ψ is a pure state on this Hilbert space. The distribution of $X(V)$ is known [see, for instance, Eq. (9) in Ref. [107]] and is given by

$$p(X) = N(1 - X)^{N-1}, \qquad X \in [0, 1]. \qquad \text{(C98)}$$

Lower bound. Let us first derive the lower bound for the average of FI. Consider first the average of a single term in a sum (C94). For $\alpha > 0$, we have the following chain of (in) equalities:

$$\mathbb{E}_{U\sim\mu(\mathcal{S}_N)} \frac{f_n(U, \varphi)}{g_n(U, \varphi)} \geq \int_{U\in\mathcal{G}_{-,\alpha}^n} \mathrm{d}\mu(U) \frac{f_n(U, \varphi)}{\alpha} + \int_{U\in\mathcal{G}_{+,\alpha}^n} \mathrm{d}\mu(U) \frac{f_n(U, \varphi)}{g_n(U, \varphi)} \qquad \text{(C99)}$$

$$= \frac{1}{\alpha}\int_{U\in\mathrm{SU}(\mathcal{S}_N)} \mathrm{d}\mu(U) f_n(U, \varphi) - \int_{U\in\mathcal{G}_{+,\alpha}^n} \mathrm{d}\mu(U) f_n(U, \varphi) \frac{(g_n(U, \varphi) - \alpha)}{g_n(U, \varphi)\alpha} \qquad \text{(C100)}$$

$$\geq \frac{1}{\alpha}\int_{U\in\mathrm{SU}(\mathcal{S}_N)} \mathrm{d}\mu(U) f_n(U, \varphi) - \frac{N^2}{\alpha}\int_{U\in\mathcal{G}_{+,\alpha}^n} \mathrm{d}\mu(U) g_n(U, \varphi)(g_n(U, \varphi) - \alpha) \qquad \text{(C101)}$$

$$=\frac{1}{\alpha}\int_{U\in\mathrm{SU}(\mathcal{S}_N)}\mathrm{d}\mu(U)f_n(U,\varphi)-\frac{N^2}{\alpha}\int_{\alpha}^{1}dXp(X)X(X-\alpha) \tag{C102}$$

$$=\frac{1}{\alpha}\int_{U\in\mathrm{SU}(\mathcal{S}_N)}\mathrm{d}\mu(U)f_n(U,\varphi)-\frac{N^2(1-\alpha)^{N+1}(2+\alpha N)}{\alpha(1+N)(2+N)}. \tag{C103}$$

In the above sequence of (in)equalities, Eq. (C99) follows from the definitions of sets $\mathcal{G}^n_{\pm,\alpha}$; Eq. (C101) follows from the non-negativity of $g_n(U,\varphi)-\alpha$ on $\mathcal{G}^n_{+,\alpha}$ and from Eq. (C95). Equation (C102) follows from the definition of the random variable X presented in the discussion above Eq. (C98). Finally, Eq. (C103) follows directly from Eq. (C98). Summing over n, we obtain the inequality

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)}F_{\mathrm{cl}}(U,\varphi)\geq\frac{1}{\alpha}\sum_{n=0}^{N}\left(\int_{U\in\mathrm{SU}(\mathcal{S}_N)}\mathrm{d}\mu(U)f_n(U,\varphi)\right)-\frac{N^2(1-\alpha)^{N+1}(2+\alpha N)}{\alpha(2+N)}. \tag{C104}$$

Using the integration techniques analogous to the ones used in preceding sections, it is possible to show that

$$\int_{U\in\mathrm{SU}(\mathcal{S}_N)}\mathrm{d}\mu(U)f_n(U,\varphi)=\frac{\mathrm{tr}(-[D^N_n,\hat{J}_y]^2)}{(N+1)(N+2)}. \tag{C105}$$

Making use of the fact that $\mathrm{tr}(\hat{J}_yD^N_n)=0$, we obtain

$$\sum_{n=0}^{N}\left(\int_{U\in\mathrm{SU}(\mathcal{S}_N)}\mathrm{d}\mu(U)f_n(U,\varphi)\right)=\sum_{n=0}^{N}\frac{2\mathrm{tr}(D^N_n,\hat{J}^2_y)}{(N+1)(N+2)}=\frac{2\mathrm{tr}(\hat{J}^2_y)}{(N+1)(N+2)}=\frac{N}{6}. \tag{C106}$$

In the last equality of Eq. (C106), we have used Eq. (C31) and the fact that $\hat{J}_y$ originates in a single-particle Hamiltonian satisfying $\mathrm{tr}(h^2)=\frac{1}{2}$. Plugging Eq. (C106) into Eq. (C104), for all $\alpha>0$, we obtain

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)}F_{\mathrm{cl}}(U,\varphi)\geq\frac{N}{6\alpha}-\frac{N^2(1-\alpha)^{N+1}(2+\alpha N)}{\alpha(2+N)}. \tag{C107}$$

By setting $\alpha=\frac{\Delta}{N}$, where Δ is a fixed positive parameter, and by using the inequality $(1-\frac{\Delta}{N})^{N+1}\leq\exp(-\Delta)$, we obtain

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)}F_{\mathrm{cl}}(U,\varphi)\geq\frac{N^2}{6\Delta}-\frac{N^2\exp(-\Delta)(2+\Delta)}{\Delta}. \tag{C108}$$

Finding the maximal value of the right-hand side of Eq. (C108) (treated as a function of Δ) is difficult. Numerical investigation shows that the maximal value is obtained very close to $\Delta=6$, which finally gives

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)}F_{\mathrm{cl}}(U,\varphi)\geq c_-N^2, \tag{C109}$$

where $c_-=\frac{1}{36}-\frac{4}{3e^5}\approx0.0244$.

Upper bound. The proof of the upper bound of the average Fisher information is analogous. For $\alpha>0$, we have the following chain of (in)equalities:

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)}\frac{f_n(U,\varphi)}{g_n(U,\varphi)}\leq\int_{U\in\mathcal{G}^n_{+,\alpha}}\mathrm{d}\mu(U)\frac{f_n(U,\varphi)}{\alpha}+\int_{U\in\mathcal{G}^n_{-,\alpha}}\mathrm{d}\mu(U)\frac{f_n(U,\varphi)}{g_n(U,\varphi)} \tag{C110}$$

$$=\frac{1}{\alpha}\int_{U\in\mathrm{SU}(\mathcal{S}_N)}\mathrm{d}\mu(U)f_n(U,\varphi)+\int_{U\in\mathcal{G}^n_{-,\alpha}}\mathrm{d}\mu(U)f_n(U,\varphi)\frac{(\alpha-g_n(U,\varphi))}{g_n(U,\varphi)\alpha} \tag{C111}$$

$$\leq\frac{1}{\alpha}\int_{U\in\mathrm{SU}(\mathcal{S}_N)}f_n(U,\varphi)+\frac{N^2}{\alpha}\int_{U\in\mathcal{G}^n_{-,\alpha}}g_n(U,\varphi)(\alpha-g_n(U,\varphi)) \tag{C112}$$

$$=\frac{1}{\alpha}\int_{U\in\mathrm{SU}(\mathcal{S}_N)}\mathrm{d}\mu(U)f_n(U,\varphi)+\frac{N^2}{\alpha}\int_0^{\alpha}dXp(X)X(\alpha-X) \tag{C113}$$

$$=\frac{1}{\alpha}\int_{U\in\mathrm{SU}(\mathcal{S}_N)}\mathrm{d}\mu(U)f_n(U,\varphi)+\frac{N^2(\alpha(2+N)+(1-\alpha)^{N+1}(2+\alpha N)-2)}{\alpha(1+N)(2+N)}. \tag{C114}$$

In the above sequence of (in)equalities, Eq. (C110) follows from the definitions of sets $\mathcal{G}^n_{\pm,\alpha}$; Eq. (C112) follows from the non-negativity of $\alpha-g_n(U,\varphi)$ on $\mathcal{G}^n_{-,\alpha}$ and from Eq. (C95). Equation (C102) follows from the definition of the random variable X presented in the discussion above Eq. (C98). Finally, Eq. (C114) follows directly from Eq. (C98). Summing over n, we obtain the inequality

$$\begin{aligned}\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)} F_{\mathrm{cl}}(U,\varphi) &\le \frac{1}{\alpha}\sum_{n=0}^{N}\left(\int_{U\in\mathrm{SU}(\mathcal{S}_N)} \mathrm{d}\mu(U) f_n(U,\varphi)\right) \\ &\quad + \frac{N^2(\alpha(2+N)+(1-\alpha)^{N+1}(2+\alpha N)-2)}{\alpha(2+N)}. \end{aligned} \tag{C115}$$

Let Δ be a fixed positive number. By setting $\alpha=\frac{\Delta}{N}$, and by using Eq. (C106), we obtain the upper bound

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)} F_{\mathrm{cl}}(U,\varphi) \le \frac{N^2}{6\Delta}+\frac{N^2}{\Delta}(\Delta-2+(\Delta+2)\exp(-\Delta))+N. \tag{C116}$$

Finding the minimal value of the right-hand side of Eq. (C116) (treated as a function of Δ) is difficult. Numerical investigation shows that the minimal value is obtained very close to $\Delta=1$. Inserting this into Eq. (C116) gives

$$\mathbb{E}_{U\sim\mu(\mathcal{S}_N)} F_{\mathrm{cl}}(U,\varphi) \le c_+ N^2 + N, \tag{C117}$$

where $c_+ = -\frac{5}{6}+\frac{3}{e}\approx 0.270$. ∎

APPENDIX D: PROOFS OF MAIN THEOREMS

In this section, we use the technical results developed in the preceding appendixes to prove the main theorems from the main manuscript, where we have used, for the sake of simplicity, the Θ notation, which allowed us to hide the presence of complicated constants in the concentration inequalities. In what follows, we will present technical versions of these theorems, explicitly giving all the relevant constants. Proofs of Theorems 1, 2, 3, and Example 1 are analogous in the sense that they all rely the concentration inequality (A7) and on

(i) upper bounds on the Lipschitz constants of the relevant functions on $\mathrm{SU}(\mathcal{H})$;
(ii) bounds or explicit values on the average of these functions on $\mathrm{SU}(\mathcal{H})$.

The proof of Theorem 4 is slightly more complicated and relies on the regularity of $F_{\mathrm{cl}}(U,\varphi)$ viewed as a function of the parameter φ.

Let us start with an immediate corollary of Fact 1 describing the concentration of measure on $\mathrm{SU}(\mathcal{H})$.

Corollary 2. Let $f:\mathrm{SU}(\mathcal{H})\mapsto\mathbb{R}$ be a function on $\mathrm{SU}(\mathcal{H})$. Let $D=|\mathcal{H}|$ be the dimension of $\mathcal{H}$. Assume the function f with the Lipschitz constant L satisfying $L\le\tilde{L}$ for some non-negative scalar $\tilde{L}$. Assume that the expectation value of f is upper bounded as $\mathbb{E}_{U\sim\mu(\mathcal{H})}f\le F_+$. Then, for every $\epsilon\ge 0$, the following large deviation bound holds,

$$\Pr_{U\sim\mu(\mathcal{H})}(f(U)\ge F_+ + \epsilon) \le \exp\left(-\frac{D\epsilon^2}{4\tilde{L}^2}\right). \tag{D1}$$

Assume that the expectation value of f is lower bounded as $\mathbb{E}_{U\sim\mu(\mathcal{H})}f\ge F_-$. Then, for every $\epsilon\ge 0$, the following large deviation bound holds,

$$\Pr_{U\sim\mu(\mathcal{H})}(f(U)\le F_- - \epsilon) \le \exp\left(-\frac{D\epsilon^2}{4\tilde{L}^2}\right). \tag{D2}$$

We use Corollary 2 to prove technical versions of Theorems 1, 2, 3 and Example 1 from the main text.

Theorem 5. (Technical version of Theorem 1 from the main text.) Fix a single-particle Hamiltonian h, local dimension d, and a pure state ψ_N on $\mathcal{H}_N$. Let $F^{\mathrm{LU}}(U) \coloneqq F^{\mathrm{LU}}(U\psi_N U^\dagger, H)$; then, for every $\epsilon\ge 0$,

$$\begin{aligned}&\Pr_{U\sim\mu(\mathcal{H}_N)}\left(F^{\mathrm{LU}}(U)\ge 4N\|h\|^2\left(1+\frac{(N-1)d^2}{\sqrt{d^N}}\right)+\epsilon\right)\\ &\quad\le \exp\left(-\frac{\epsilon^2 d^N}{4096\|h\|^4 N^4}\right),\end{aligned} \tag{D3}$$

$$\begin{aligned}&\Pr_{U\sim\mu(\mathcal{H}_N)}\left(F^{\mathrm{LU}}(U)\le \frac{4N\mathrm{tr}(h^2)d^N}{d(d^N+1)}-\epsilon\right)\\ &\quad\le \exp\left(-\frac{\epsilon^2 d^N}{4096\|h\|^4 N^4}\right).\end{aligned} \tag{D4}$$

Setting $\epsilon = 2N\|h\|^2(1+\frac{(N-1)d^2}{\sqrt{d^N}})$ and $\epsilon=\frac{2N\mathrm{tr}(h^2)d^N}{d(d^N+1)}$ in Eqs. (D3) and (D4), respectively, yields Theorem 1.

Proof.—The proof of Theorem 5 follows directly from Corollary 2 and results proved previously. From Lemmas 1 and 3, one can infer that the Lipschitz constant of F^{LU} is upper bounded by $\tilde{L}=32\|H\|^2=32N^2\|h\|^2$. From Eq. (17), we have the upper bound on $\mathbb{E}_{U\sim\mu(\mathcal{H})}F^{\mathrm{LU}}$. Using this bound in Eq. (D1) gives Eq. (D3). The lower bound $\mathbb{E}_{U\sim\mu(\mathcal{H})}F^{\mathrm{LU}}$ can be obtained by noting that the unoptimized QFI is a lower bound to its optimized version. Therefore,

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{H})} F^{\mathrm{LU}} \ge \mathop{\mathbb{E}}_{U\sim\mu(\mathcal{H})} F(U\psi_N U^\dagger, H) = \frac{4N\mathrm{tr}(h^2)d^N}{d(d^N+1)}, \tag{D5}$$

where, in the last equality, we used Eq. (C27). Plugging Eq. (D5) into Eq. (D2) yields Eq. (D4). ∎

Theorem 6. (Technical version of Theorem 2 from the main text.) Fix a single-particle Hamiltonian h, local dimension d, and a state σ_N from the symmetric subspace $\mathcal{S}_N$ with eigenvalues $\{p_j\}_j$. Let σ_{mix} be the maximally mixed state on $\mathcal{S}_N$. Let $F(U)\coloneqq F(U\sigma_N U^\dagger, H)$; then, for every $\epsilon\ge 0$,

$$\Pr_{U\sim\mu(\mathcal{S}_N)}\left(F(U)\le d_B(\sigma_N,\sigma_{\rm mix})^2\frac{2N(N+d)\mathrm{tr}(h^2)}{d(d+1)}\times\frac{|\mathcal{S}_N|^2}{|\mathcal{S}_N|^2-1}-\epsilon\right)\le\exp\left(-\frac{\epsilon^2|\mathcal{S}_N|}{4096C\|h\|^4N^4}\right),\quad (D6)$$

where $|\mathcal{S}_N|=\binom{N+d-1}{N}$ and $C=\min\{1,8d_B(\sigma_N,\sigma_{\rm mix})\}$. Setting $\epsilon=d_B(\sigma_N,\sigma_{\rm mix})^2\frac{N(N+d)\mathrm{tr}(h^2)}{d(d+1)}\frac{|\mathcal{S}_N|^2}{|\mathcal{S}_N|^2-1}$ in Eq. (D6) yields Theorem 2.

Proof.—The proof is analogous to the proof of Theorem 5. From Lemma 1, we infer that the Lipschitz constant of $F(U)$ is upper bounded by $\tilde{L}=32\|H\|^2\min\{1,8d_B(\sigma_N,\sigma_{\rm mix})\}=32N^2\|h\|^2\min\{1,8d_B(\sigma_N,\sigma_{\rm mix})\}$. From Eq. (C28) in Lemma 8, we get

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)}F(U)=\frac{4N(N+d)\mathrm{tr}(h^2)}{d(d+1)}\frac{|\mathcal{S}_N|}{|\mathcal{S}_N|+1}\Lambda(\{p_j\}_j).\quad (D7)$$

Using the inequality (C15) and the Fuch–van de Graaf inequality [102], $1-\mathcal{F}^2(\sigma_N,\sigma_{\rm mix})\le\frac{1}{2}d_B(\sigma_N,\sigma_{\rm mix})$, we obtain

$$\Lambda(\{p_j\}_j)\ge\frac{|\mathcal{S}_N|}{2(|\mathcal{S}_N|-1)}d_B(\sigma_N,\sigma_{\rm mix})^2.\quad (D8)$$

Inserting this inequality into Eq. (D7) gives

$$\mathbb{E}_{U\sim\mu(\mathcal{H}_N)}F(U)\ge d_B(\sigma_N,\sigma_{\rm mix})^2\frac{2N(N+d)\mathrm{tr}(h^2)}{d(d+1)}\frac{|\mathcal{S}_N|^2}{|\mathcal{S}_N|^2-1},\quad (D9)$$

which, together with the bound on the Lipschitz constant of $F(U)$ and Corollary 2, results in Eq. (D6). ∎

Example 3 (Technical version of Example VI A from the main manuscript). Fix a local dimension d, single-particle Hamiltonian h, and $p\in[0,1]$. Let ψ_N be a pure state on $\mathcal{S}_N$, and set

$$\sigma_N(p)=(1-p)\psi_N+p\sigma_{\rm mix}.\quad (D10)$$

Let $F_p(U)\coloneqq F(U\sigma_N(p)U^\dagger,H)$; then, for every $\epsilon>0$,

$$\Pr_{U\sim\mu(\mathcal{S}_N)}\left(|F_p(U)-\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)}F_p|\ge\epsilon\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)}F_p\right)\le 2\exp\left(-\frac{\epsilon^2\mathrm{tr}(h^2)^2(N+d)^2|\mathcal{S}_N|^2}{64\|h\|^4(d(d+1)N(1+|\mathcal{S}_N|))^2}|\mathcal{S}_N|\right),\quad (D11)$$

where $|\mathcal{S}_N|=\binom{N+d-1}{N}$ and

$$\mathop{\mathbb{E}}_{U\sim\mu(\mathcal{S}_N)}F_p=\frac{4N(N+d)\mathrm{tr}(h^2)}{d(d+1)}\frac{|\mathcal{S}_N|}{|\mathcal{S}_N|+1}\frac{(1-p)^2}{(1-p+2p/|\mathcal{S}_N|)}.\quad (D12)$$

Equation (D12) implies Example VI A, as for fixed local dimension d, we have $|\mathcal{S}_N|\in\Theta(N^{d-1})$.

Sketch of the proof.—The proof of Example 3 parallels proofs of Theorems 5 and 6 and relies on Fact 1. The bound of the Lipschitz constant of $F_p(U)$ is provided by Lemma 3. The expression for the average of $F_p(U)$ is given in Lemma 8. The inequality (D11) follows directly from concentration inequalities from Fact 1 by setting $\epsilon=\tilde{\epsilon}\mathbb{E}_{U\sim\mu(\mathcal{S}_N)}F_p$. ∎

Theorem 7. (Technical version of Theorem 3 from the main manuscript.) Fix a single-particle Hamiltonian h, local dimension d, non-negative integer k, and a state σ_N on $\mathcal{S}_N$ with eigenvalues $\{p_j\}_j$. Let $\sigma_{\rm mix}$ be the maximally mixed state on $\mathcal{S}_N$. Let $F_k(U)\coloneqq F(\mathrm{tr}_k(U\sigma_NU^\dagger),H_{N-k})$; then, for every $\epsilon\ge 0$,

$$\Pr_{U\sim\mu(\mathcal{S}_N)}\left(F_k(U)\le 2\frac{(N-k)(N+d)}{(d+1)(d+k)}\times\frac{|\mathcal{S}_N|(|\mathcal{S}_N|\mathrm{tr}\rho_N^2-1)\mathrm{tr}h^2}{|\mathcal{S}_k|(|\mathcal{S}_N|^2-1)}-\epsilon\right)\le\exp\left(-\frac{\epsilon^2|\mathcal{S}_N|}{4096C\|h\|^4(N-k)^4}\right),\quad (D13)$$

where $|\mathcal{S}_N|=\binom{N+d-1}{N}$ and $C=\min\{1,8d_B(\sigma_N,\sigma_{\rm mix})\}$. Setting $\epsilon=\frac{(N-k)(N+d)}{(d+1)(d+k)}\frac{|\mathcal{S}_N|(|\mathcal{S}_N|\mathrm{tr}\rho_N^2-1)\mathrm{tr}h^2}{|\mathcal{S}_k|(|\mathcal{S}_N|^2-1)}$ in Eq. (D13) yields Theorem 3.

Sketch of the proof.—The proof of Theorem 7 parallels proofs of Theorems 5 and 6 and relies on Corollary 2. The bound of the Lipschitz constant of $F_k(U)$ is provided by Lemma 4. The lower bound for the average of $F_k(U)$ is given in Lemma 9. ∎

Theorem 8. (Technical version of Theorem 4 from the main manuscript.) Let ψ_N be a fixed pure state on $\mathcal{S}_N$ with $d=2$ bosonic modes. Let $p_{n|\varphi}(U\psi_NU^\dagger)$ be the probability to obtain outcome n in the interferometric scheme defined in Sec. VII, given that the value of the unknown phase parameter is φ and the input state is $U\psi_NU^\dagger$ [see also Eq. (31)]. Let $F_{\rm cl}(U,\varphi)\coloneqq F_{\rm cl}(\{p_{n|\varphi}(U\psi_NU^\dagger)\})$ be the corresponding FI according to Eq. (32) [or Eq. (C87)]. Then, for every $\epsilon\ge 0$ and every $\varphi\in[0,2\pi]$, we have

$$\Pr_{U\sim\mu(\mathcal{S}_N)}(F_{\rm cl}(U,\varphi)\le\mathbb{E}_{U\sim\mu(\mathcal{S}_N)}F_{\rm cl}(U,\varphi)-\epsilon)\le\exp\left(-\frac{\epsilon^2}{144N^4}(N+1)\right),\quad (D14)$$

$$\Pr_{U\sim\mu(\mathcal{S}_N)}(F_{\rm cl}(U,\varphi)\ge\mathbb{E}_{U\sim\mu(\mathcal{S}_N)}F_{\rm cl}(U,\varphi)+\epsilon)\le\exp\left(-\frac{\epsilon^2}{144N^4}(N+1)\right).\quad (D15)$$

In the equations above, $\mathbb{E}_{U\sim\mu(\mathcal{S}_N)}F_{\mathrm{cl}}(U,\varphi)$ satisfies inequalities

$$c_{-}N^2 \leq \mathbb{E}_{U\sim\mu(\mathcal{S}_N)}F_{\mathrm{cl}}(U,\varphi) \leq c_{+}N^2 + N, \tag{D16}$$

where

$$c_{-} = \frac{1}{36} - \frac{4}{3e^5} \approx 0.0244, \qquad c_{+} = -\frac{5}{6} + \frac{3}{e} \approx 0.270. \tag{D17}$$

Moreover, we have the following inequality:

$$\Pr_{U\sim\mu(\mathcal{S}_N)}\left(\exists\varphi \in [0,2\pi] F_{\mathrm{cl}}(U,\varphi) \leq \frac{c_{-}}{4}N^2\right) \leq \left\lceil\frac{12\pi N}{c_{-}}\right\rceil \exp\left(-\frac{c_{-}^2}{576}(N+1)\right), \tag{D18}$$

where $\lceil x\rceil$ stands for the smallest integer that is not less than x. Equation (D18) yields exactly Eq. (33) from Theorem 4.

Proof.—Equations (D14) and (D15) follow directly from Fact 1 and the bounds of the Lipschitz constant of $F_{\mathrm{cl}}(U,\varphi)$, treated as a function of U (for fixed φ), given in Lemma 5. From Lemma 5, it follows that the Lipschitz constant of $F_{\mathrm{cl}}(U,\varphi)$ is bounded above as

$$L \leq 24\|H\|^2 = 24\|\hat{J}_z\|^2 = 6N^2. \tag{D19}$$

Inequalities from Eq. (D16) follow from Lemma 10. The nontrivial part of the proof is the justification of Eq. (D18). Let us first introduce the discretization of the interval $[0,2\pi]$ by M equally spaced points:

$$\varphi_i = (i-1)\frac{2\pi}{M}, \qquad i = 1,\ldots,M. \tag{D20}$$

Moreover, let us notice that from Eq. (A59), it follows that $F_{\mathrm{cl}}(U,\varphi)$ is Lipschitz continuous for fixed U and varying φ:

$$\left|\frac{d}{d\varphi}F_{\mathrm{cl}}(U,\varphi)\right| \leq 24\|H\|^3 = 3N^3, \tag{D21}$$

where in Eq. (A59), we set $X = H = \hat{J}_z$. From Eq. (D21), it follows that for fixed $U \in \mathrm{SU}(\mathcal{S}_N)$ and for $\varphi,\tilde{\varphi} \in [0,2\pi]$, we have

$$|F_{\mathrm{cl}}(U,\varphi) - F_{\mathrm{cl}}(U,\tilde{\varphi})| \leq 3N^3|\varphi - \tilde{\varphi}|. \tag{D22}$$

When the points in the discretization (D20) are separated by $\Delta = \frac{2\pi}{M}$, the distance on any $\varphi \in [0,2\pi]$ to the closest φ_i does not exceed $\Delta' = \frac{\Delta}{2} = \frac{\pi}{M}$. Using the union bound, Eq. (D14), and the lower bound in Eq. (D16), we obtain

$$\Pr_{U\sim\mu(\mathcal{S}_N)}(\exists\varphi_i F_{\mathrm{cl}}(U,\varphi_i) \leq c_{-}N^2 - \epsilon) \leq M\exp\left(-\frac{\epsilon^2}{144N^4}(N+1)\right). \tag{D23}$$

Using Eq. (D22) and the discussion following it, we obtain

$$\Pr_{U\sim\mu(\mathcal{S}_N)}(\exists\varphi F_{\mathrm{cl}}(U,\varphi) \leq c_{-}N^2 - 3N^3\Delta' - \epsilon) \leq M\exp\left(-\frac{\epsilon^2}{144N^4}(N+1)\right). \tag{D24}$$

Now, by setting $M = \lceil\frac{12\pi N}{c_{-}}\rceil$ (this is the smallest integer M such that $3N^3\Delta' \leq \frac{c_i}{4}N^2$) and $\epsilon = \frac{c_{-}}{2}N^2$ in the above equation, we obtain Eq. (D18). ∎

APPENDIX E: PARTIAL-TRACE AND BEAM-SPLITTER MODELS OF PARTICLE LOSSES

In this appendix, we prove the equivalence of the beam-splitter model of particle losses and the operation of taking the partial trace over the constituent particles in the system of N bosons in $d = 2$ modes. A general pure state ψ_N of N bosons in two modes a and b can be written as

$$|\psi_N\rangle = \sum_{n=0}^{N}\alpha_n|n, N-n\rangle = \sum_{n=0}^{N}\alpha_n|D_n^N\rangle, \tag{E1}$$

with the complex coefficients $\{\alpha_n\}_{n=0}^N$ satisfying $\sum_{n=0}^N|\alpha_n|^2 = 1$. Each Dicke state $|D_n^N\rangle$ can be written in the basis of *particle basis* $|\rangle_{\mathsf{p}}$ as

$$|D_n^N\rangle = \frac{1}{\sqrt{\binom{N}{n}}}\sum_{\mathbf{x}\in\{0,1\}^N}\delta_{x,n}|\mathbf{x}\rangle_{\mathsf{p}}, \tag{E2}$$

where $x \coloneqq |\mathbf{x}| \coloneqq \sum_i x_i$ denotes the Hamming weight of any binary string $\mathbf{x} = [x_1,\ldots,x_N]$, whose consecutive entries specify the state of each qubit. As a result, we may write a general bosonic pure state (E1) in the particle basis as

$$|\psi_N\rangle = \sum_{\mathbf{x}\in\{0,1\}^N}c_{\mathbf{x}}|\mathbf{x}\rangle_{\mathsf{p}} = \sum_{\mathbf{x}\in\{0,1\}^N}c_{\mathbf{x}}|x_1\rangle|x_2\rangle\ldots|x_N\rangle, \tag{E3}$$

with the coefficients $c_{\mathbf{x}}$ then given by $c_{\mathbf{x}} = \frac{1}{\sqrt{\binom{N}{n}}}\sum_{n=0}^N\alpha_n\delta_{x,n}$.

1. Tracing out k particles

Let us define the notation in which we can split any binary string $\mathbf{x}$ (describing N qubits) into two strings, $\mathbf{x}'$ and $\mathbf{u}$ (describing $N-k$ and k qubits, respectively) so that $\mathbf{x} = [\mathbf{x}',\mathbf{u}] = [x_1',\ldots,x_{N-k}',u_1,\ldots,u_k]$. Then, we can generally write the bosonic state (E3) in the particle basis after tracing out the last k qubits as

$$\varrho^{\mathrm{tr}}_{N-k} := \mathrm{tr}_k\{\psi_N\} \tag{E4}$$

$$= \mathrm{tr}_k\left\{\sum_{\mathbf{x},\mathbf{y}=0^N}^{1^N} c_{\mathbf{x}} c^{\star}_{\mathbf{y}} |\mathbf{x}\rangle_{\mathsf{p}}\langle\mathbf{y}|\right\}$$

$$= \sum_{\mathbf{x},\mathbf{y}=0^N}^{1^N} c_{\mathbf{x}} c^{\star}_{\mathbf{y}} \mathrm{tr}_k\{|\mathbf{x}\rangle_{\mathsf{p}}\langle\mathbf{y}|\} \tag{E5}$$

$$= \sum_{\mathbf{x}',\mathbf{x}'=0^{N-k}}^{1^{N-k}} [\varrho^{\mathrm{tr}}_{N-k}]_{\mathbf{x}'\mathbf{y}'} |\mathbf{x}'\rangle_{\mathsf{p}}\langle\mathbf{y}'|, \tag{E6}$$

where the above matrix entries of $\varrho^{\mathrm{tr}}_{N-k}$ are given by

$$[\varrho^{\mathrm{tr}}_{N-k}]_{\mathbf{x}'\mathbf{y}'} = \sum_{\mathbf{u},\mathbf{w}=0^k}^{1^k} c_{[\mathbf{x}',\mathbf{u}]} c^{\star}_{[\mathbf{y}',\mathbf{w}]} \delta_{\mathbf{u}\mathbf{w}} = \sum_{\mathbf{u}=0^k}^{1^k} c_{[\mathbf{x}',\mathbf{u}]} c^{\star}_{[\mathbf{y}',\mathbf{u}]} \tag{E7}$$

$$= \sum_{\mathbf{u}=0^k}^{1^k} \left(\sum_{n=0}^{N} \frac{\alpha_n \delta_{x'+u,n}}{\sqrt{\binom{N}{n}}}\right)\left(\sum_{m=0}^{N} \frac{\alpha^{\star}_m \delta_{y'+u,m}}{\sqrt{\binom{N}{m}}}\right)$$

$$= \sum_{u=0}^{k} \binom{k}{u} \frac{\alpha_{x'+u}\alpha^{\star}_{y'+u}}{\sqrt{\binom{N}{x'+u}}\sqrt{\binom{N}{y'+u}}}. \tag{E8}$$

In the mode basis, we can equivalently write

$$\varrho^{\mathrm{tr}}_{N-k} = \sum_{n,m=0}^{N-k} [\varrho^{\mathrm{tr}}_{N-k}]_{nm} |D^{N-k}_n\rangle\langle D^{N-k}_m|, \tag{E9}$$

and with the help of Eqs. (E2) and (E6), we can explicitly evaluate the corresponding density-matrix entries:

$$[\varrho^{\mathrm{tr}}_{N-k}]_{nm} = \sum_{\mathbf{x}',\mathbf{y}'=0^{N-k}}^{1^{N-k}} [\varrho^{\mathrm{tr}}_{N-k}]_{\mathbf{x}'\mathbf{y}'} \langle D^{N-k}_m||\mathbf{x}'\rangle_{\mathsf{p}}\langle\mathbf{y}'||D^{N-k}_n\rangle \tag{E10}$$

$$= \sum_{\mathbf{x}',\mathbf{y}'=0^{N-k}}^{1^{N-k}} [\varrho^{\mathrm{tr}}_{N-k}]_{\mathbf{x}'\mathbf{y}'} \frac{\delta_{x',m}}{\sqrt{\binom{N-k}{m}}}\frac{\delta_{y',n}}{\sqrt{\binom{N-k}{n}}} \tag{E11}$$

$$= \sum_{x',y'=0}^{N-k} \binom{N-k}{x'}\binom{N-k}{y'} \sum_{u=0}^{k}\binom{k}{u} \times \frac{\alpha_{x'+u}\alpha^{\star}_{y'+u}}{\sqrt{\binom{N}{x'+u}}\sqrt{\binom{N}{y'+u}}} \frac{\delta_{x',m}}{\sqrt{\binom{N-k}{m}}}\frac{\delta_{y',n}}{\sqrt{\binom{N-k}{n}}} \tag{E12}$$

$$= \sum_{u=0}^{k} \alpha_{m+u}\alpha^{\star}_{n+u} \binom{k}{u} \sqrt{\frac{\binom{N-k}{m}\binom{N-k}{n}}{\binom{N}{m+u}\binom{N}{n+u}}}. \tag{E13}$$

2. Beam-splitter model of mode-asymmetric particle losses

In quantum optics, photonic losses are modeled by adding fictitious beam splitters (BSs) of fixed transmittance into the light transmission modes [39]. In this way, by impinging a vacuum state on the other input port of any such BS and tracing out its unobserved output port, one obtains a model depicting loss of photons. In the case of the two-mode N-photon bosonic state (E1), after fixing the transmissivity of the fictitious BS introduced in mode a (b) to η_a (η_b), the density matrix then describing the observed modes generally reads [3]

$$\varrho^{\mathrm{BS}}_N := \Lambda^{\mathrm{BS}}_{\eta_a,\eta_b}[\psi_N] \tag{E14}$$

$$= \sum_{l_a=0}^{N}\sum_{l_b=0}^{N-l_a} p_{l_a,l_b} |\xi_{l_a,l_b}\rangle_{\mathsf{m}}\langle\xi_{l_a,l_b}|, \tag{E15}$$

where $\Lambda^{\mathrm{BS}}_{\eta_a,\eta_b}$ is the effective quantum channel representing the action of fictitious BSs in the two modes, while indices l_a and l_b denote the number of photons lost in modes a and b, respectively. The states

$$|\xi_{l_a,l_b}\rangle_{\mathsf{m}} := \frac{1}{\sqrt{p_{l_a,l_b}}} \sum_{n=l_a}^{N-l_b} \alpha_n \sqrt{b^{(l_a,l_b)}_n} |n-l_a, N-n-l_b\rangle \tag{E16}$$

are generally nonorthogonal, and their coefficients contain generalized binomial factors:

$$b^{(l_a,l_b)}_n := \binom{n}{l_a} \eta_a^{n-l_a}(1-\eta_a)^{l_a} \binom{N-n}{l_b} \eta_b^{N-n-l_b}(1-\eta_b)^{l_b}. \tag{E17}$$

The probability of losing l_a and l_b photons in modes a and b, respectively, then reads

$$p_{l_a,l_b} = \sum_{n=l_a}^{N-l_b} |\alpha_n|^2 b^{(l_a,l_b)}_n. \tag{E18}$$

On the other hand, after reindexing Eq. (E15) by l—the total number of photons lost in both modes—the output two-mode mixed state can equivalently be rewritten as

$$\varrho^{\mathrm{BS}}_N = \bigoplus_{l=0}^{N} p_l \varrho^{\mathrm{BS}}_{N,l}, \tag{E19}$$

where

$$\varrho_{N,l}^{\mathrm{BS}} \coloneqq \frac{1}{p_l}\sum_{l_a=0}^{l} p_{l_a,l-l_a}|\xi_{l_a,l-l_a}\rangle_{\mathsf{m}}\langle\xi_{l_a,l-l_a}| \tag{E20}$$

$$= \frac{1}{p_l}\sum_{l_a=0}^{l}\sum_{n,m=l_a}^{N-l+l_a} \times \alpha_n\alpha_m^\star\sqrt{b_n^{(l_a,l-l_a)}b_m^{(l_a,l-l_a)}}|n-l_a,N-n-l+l_a\rangle_{\mathsf{m}} \times\langle m-l_a,N-m-l+l_a| \tag{E21}$$

belong to orthogonal subspaces and represent the state after loss of l photons, which may occur with probability

$$p_l = \sum_{l_a=0}^{l} p_{l_a,l-l_a} = \sum_{l_a=0}^{l}\sum_{n=l_a}^{N-l+l_a}|\alpha_n|^2 b_n^{(l_a,l-l_a)} = \sum_{n=0}^{N}|\alpha_n|^2\sum_{l_a=\max\{0,n-N+l\}}^{\min\{l,n\}} b_n^{(l_a,l-l_a)}. \tag{E22}$$

3. Equivalence of the partial-trace and beam-splitter models in the case of equal losses in the two modes

Lemma 11. For equal photonic losses in both modes, $\eta \coloneqq \eta_a = \eta_b$, the fictitious BS model is equivalent to tracing out k particles with k distributed according to a binomial distribution, i.e.,

$$\forall_{\psi_n\in\mathcal{S}_N}:\ \Lambda_{\eta,\eta}^{\mathrm{BS}}[\psi_N] = \bigoplus_{k=0}^{N} p_k \mathrm{tr}_k\{\psi_N\} \quad \text{with} \quad p_k = \binom{N}{k}\eta^{N-k}(1-\eta)^k. \tag{E23}$$

Proof.—In the case of mode-symmetric losses, $\eta \coloneqq \eta_a = \eta_b$, the overall probability of losing l photons becomes independent of the state ψ_N [i.e., its coefficients α_n of Eq. (E1)], as Eq. (E22) then simplifies to

$$p_l = \sum_{n=0}^{N}|\alpha_n|^2\sum_{l_a=\max\{0,n-N+l\}}^{\min\{l,n\}}\binom{n}{l_a} \times\binom{N-n}{l-l_a}\eta^{N-l}(1-\eta)^l = \binom{N}{l}\eta^{N-l}(1-\eta)^l. \tag{E24}$$

Furthermore, the state (E21) in each orthogonal subspace indexed by l then takes a simpler form,

$$\varrho_{N,l}^{\mathrm{BS}} = \frac{1}{p_l}\sum_{l_a=0}^{l}\sum_{n,m=0}^{N-l}\alpha_{n+l_a}\alpha_{m+l_a}^\star\sqrt{b_{n+l_a}^{(l_a,l-l_a)}} \times\sqrt{b_{m+l_a}^{(l_a,l-l_a)}}|n,N-l-n\rangle_{\mathsf{m}}\langle m,N-l-m| \tag{E25}$$

$$= \sum_{n,m=0}^{N-l}[\varrho_{N,l}^{\mathrm{BS}}]_{nm}|D_n^{N-l}\rangle\langle D_m^{N-l}|, \tag{E26}$$

where we have shifted the indices $n\to n+l_a$ and $m\to m+l_a$ to explicitly rewrite the state in the Dicke basis, in which its matrix entries then read

$$[\varrho_{N,l}^{\mathrm{BS}}]_{nm} = \frac{1}{p_l}\sum_{l_a=0}^{l}\alpha_{n+l_a}\alpha_{m+l_a}^\star\sqrt{b_{n+l_a}^{(l_a,l-l_a)}}\sqrt{b_{m+l_a}^{(l_a,l-l_a)}} \tag{E27}$$

$$= \frac{1}{p_l}\sum_{l_a=0}^{l}\alpha_{n+l_a}\alpha_{m+l_a}^\star\eta^{N-l}(1-\eta)^l \times\sqrt{\binom{n+l_a}{l_a}\binom{N-n-l_a}{l-l_a}\binom{m+l_a}{l_a}\binom{N-m-l_a}{l-l_a}} \tag{E28}$$

$$= \sum_{l_a=0}^{l}\alpha_{n+l_a}\alpha_{m+l_a}^\star\sqrt{\frac{\binom{n+l_a}{l_a}\binom{N-n-l_a}{l-l_a}}{\binom{N}{l}}}\sqrt{\frac{\binom{m+l_a}{l_a}\binom{N-m-l_a}{l-l_a}}{\binom{N}{l}}}. \tag{E29}$$

However, using $\binom{n}{m}\binom{m}{k} = \binom{n}{k}\binom{n-k}{m-k}$ and $\binom{n}{k} = \binom{n}{n-k}$, we get

$$\sqrt{\frac{\binom{n+l_a}{l_a}\binom{N-n-l_a}{l-l_a}}{\binom{N}{l}}} = \sqrt{\binom{l}{l_a}}\sqrt{\frac{\binom{N-l}{n}}{\binom{N}{n+l_a}}}. \tag{E30}$$

Finally, this allows us to write the matrix entries specified in the Dicke basis as

$$[\varrho_{N,l}^{\mathrm{BS}}]_{nm} = \sum_{l_a=0}^{l}\alpha_{n+l_a}\alpha_{m+l_a}^\star\binom{l}{l_a}\sqrt{\frac{\binom{N-l}{n}}{\binom{N}{n+l_a}}}\sqrt{\frac{\binom{N-l}{m}}{\binom{N}{m+l_a}}}. \tag{E31}$$

Comparing the above expression with Eq. (E13) and relabeling the indices $l\to k$ and $l_a\to u$, one observes that, independently of ψ_N, indeed $\varrho_{N,l}^{\mathrm{BS}} = \varrho_{N-l}^{\mathrm{tr}} = \mathrm{tr}_l\{\psi_N\}$. Hence, Eq. (E19) yields Eq. (E23) with binomially distributed p_l according to Eq. (E24). ∎

[1] V. Giovannetti, S. Lloyd, and L. Maccone, *Advances in Quantum Metrology*, Nat. Photonics **5**, 222 (2011).
[2] G. Tóth and I. Apellaniz, *Quantum Metrology from a Quantum Information Science Perspective*, J. Phys. A Math. Theor. **47**, 424006 (2014).

[3] R. Demkowicz-Dobrzański, M. Jarzyna, and J. Kołodyński, *Quantum Limits in Optical Interferometry, in Progress in Optics*, edited by E. Wolf (Elsevier, New York, 2015), Vol. 60, pp. 345–435.

[4] J. P. Dowling and K. P. Seshadreesan, *Quantum Optical Technologies for Metrology, Sensing, and Imaging*, J. Lightwave Technol. **33**, 2359 (2015).

[5] V. Giovannetti, S. Lloyd, and L. Maccone, *Quantum-Enhanced Measurements: Beating the Standard Quantum Limit*, Science **306**, 1330 (2004).

[6] V. Giovannetti, S. Lloyd, and L. Maccone, *Quantum Metrology*, Phys. Rev. Lett. **96**, 010401 (2006).

[7] S. F. Huelga, C. Macchiavello, T. Pellizzari, A. K. Ekert, M. B. Plenio, and J. I. Cirac, *Improvement of Frequency Standards with Quantum Entanglement*, Phys. Rev. Lett. **79**, 3865 (1997).

[8] B. M. Escher, R. L. de Matos Filho, and L. Davidovich, *General Framework for Estimating the Ultimate Precision Limit in Noisy Quantum-Enhanced Metrology*, Nat. Phys. **7**, 406 (2011).

[9] R. Demkowicz-Dobrzański, J. Kołodyński, and M. Guţă, *The Elusive Heisenberg Limit in Quantum-Enhanced Metrology*, Nat. Commun. **3**, 1063 (2012).

[10] M. W. Mitchell, J. S. Lundeen, and A. M. Steinberg, *Super-Resolving Phase Measurements with a Multi-Photon Entangled State*, Nature (London) **429**, 161 (2004).

[11] T. Nagata, R. Okamoto, J. L. O'Brien, K. Sasaki, and S. Takeuchi, *Beating the Standard Quantum Limit with Four Entangled Photons*, Science **316**, 726 (2007).

[12] LIGO Collaboration, *A Gravitational Wave Observatory Operating Beyond the Quantum Shot-Noise Limit*, Nat. Phys. **7**, 962 (2011); *Enhanced Sensitivity of the LIGO Gravitational Wave Detector by Using Squeezed States of Light*, Nat. Photonics **7**, 613 (2013).

[13] D. Leibfried, M. D. Barrett, T. Schaetz, J. Britton, J. Chiaverini, W. M. Itano, J. D. Jost, C. Langer, and D. J. Wineland, *Toward Heisenberg-Limited Spectroscopy with Multiparticle Entangled States*, Science **304**, 1476 (2004).

[14] C. F. Roos, M. Chwalla, K. Kim, M. Riebe, and R. Blatt, *Designer Atoms for Quantum Metrology*, Nature (London) **443**, 316 (2006).

[15] W. Wasilewski, K. Jensen, H. Krauter, J. J. Renema, M. V. Balabas, and E. S. Polzik, *Quantum Noise Limited and Entanglement-Assisted Magnetometry*, Phys. Rev. Lett. **104**, 133601 (2010).

[16] R. J. Sewell, M. Koschorreck, M. Napolitano, B. Dubost, N. Behbood, and M. W. Mitchell, *Magnetic Sensitivity Beyond the Projection Noise Limit by Spin Squeezing*, Phys. Rev. Lett. **109**, 253605 (2012).

[17] J. Appel, P. J. Windpassinger, D. Oblak, U. Busk Hoff, N. Kjærgaard, and E. S. Polzik, *Mesoscopic Atomic Entanglement for Precision Measurements Beyond the Standard Quantum Limit*, Proc. Natl. Acad. Sci. U.S.A. **106**, 10960 (2009).

[18] A. Louchet-Chauvet, J. Appel, J. J. Renema, D. Oblak, N. Kjaergaard, and E. S. Polzik, *Entanglement-Assisted Atomic Clock Beyond the Projection Noise Limit*, New J. Phys. **12**, 065032 (2010).

[19] L. Pezzé and A. Smerzi, *Entanglement, Nonlinear Dynamics, and the Heisenberg Limit*, Phys. Rev. Lett. **102**, 100401 (2009).

[20] M. F. Riedel, P. Böhi, Y. Li, T. W. Hänsch, A. Sinatra, and P. Treutlein, *Atom-Chip-Based Generation of Entanglement for Quantum Metrology*, Nature (London) **464**, 1170 (2010).

[21] R. Krischek, C. Schwemmer, W. Wieczorek, H. Weinfurter, P. Hyllus, L. Pezzé, and A. Smerzi, *Useful Multiparticle Entanglement and Sub-Shot-Noise Sensitivity in Experimental Phase Estimation*, Phys. Rev. Lett. **107**, 080504 (2011).

[22] H. Strobel, W. Muessel, D. Linnemann, T. Zibold, D. B. Hume, L. Pezzè, A. Smerzi, and M. K. Oberthaler, *Fisher Information and Entanglement of Non-Gaussian Spin States*, Science **345**, 424 (2014).

[23] B. Lücke, J. Peise, G. Vitagliano, J. Arlt, L. Santos, G. Tóth, and C. Klempt, *Detecting Multiparticle Entanglement of Dicke States*, Phys. Rev. Lett. **112**, 155304 (2014).

[24] A. S. Sørensen and K. Mølmer, *Entanglement and Extreme Spin Squeezing*, Phys. Rev. Lett. **86**, 4431 (2001).

[25] P. Hyllus, W. Laskowski, R. Krischek, C. Schwemmer, W. Wieczorek, H. Weinfurter, L. Pezzé, and A. Smerzi, *Fisher Information and Multiparticle Entanglement*, Phys. Rev. A **85**, 022321 (2012); G. Tóth, *Multipartite Entanglement and High-Precision Metrology*, Phys. Rev. A **85**, 022322 (2012).

[26] P. Hyllus, O. Gühne, and A. Smerzi, *Not All Pure Entangled States Are Useful for Sub-Shot-Noise Interferometry*, Phys. Rev. A **82**, 012337 (2010).

[27] I. Apellaniz, B. Lücke, J. Peise, C. Klempt, and G. Tóth, *Detecting Metrologically Useful Entanglement in the Vicinity of Dicke States*, New J. Phys. **17**, 083027 (2015).

[28] F. Benatti and D. Braun, *Sub—Shot-Noise Sensitivities without Entanglement*, Phys. Rev. A **87**, 012340 (2013).

[29] Ł. Czekaj, A. Przysiężna, M. Horodecki, and P. Horodecki, *Quantum Metrology: Heisenberg Limit with Bound Entanglement*, Phys. Rev. A **92**, 062303 (2015).

[30] R. Augusiak, J. Kołodyński, A. Streltsov, M. N. Bera, A. Acín, and M. Lewenstein, *Asymptotic Role of Entanglement in Quantum Metrology*, Phys. Rev. A **94**, 012339 (2016).

[31] D. N. Page, *Average Entropy of a Subsystem*, Phys. Rev. Lett. **71**, 1291 (1993).

[32] S. K. Foong and S. Kanno, *Proof of Page's Conjecture on the Average Entropy of a Subsystem*, Phys. Rev. Lett. **72**, 1148 (1994).

[33] S. Sen, *Average Entropy of a Quantum Subsystem*, Phys. Rev. Lett. **77**, 1 (1996).

[34] M. Lewenstein, A. Sanpera, and V. Ahufinger, *Ultracold Atoms in Optical Lattices: Simulating Quantum Many Body Physics* (Oxford University Press, New York, 2012).

[35] P. Hayden, D. W. Leung, and A. Winter, *Aspects of Generic Entanglement*, Commun. Math. Phys. **265**, 95 (2006).

[36] D. Gross, S. T. Flammia, and J. Eisert, *Most Quantum States Are Too Entangled to Be Useful as Computational Resources*, Phys. Rev. Lett. **102**, 190501 (2009).

[37] M. Oszmaniec and M. Kuś, *Fraction of Isospectral States Exhibiting Quantum Correlations*, Phys. Rev. A **90**, 010302 (2014).
[38] M. Oszmaniec, Ph.D. thesis, University of Warsaw, 2014.
[39] H.-A. Bachor and T. C. Ralph, *A Guide to Experiments in Quantum Optics* (Wiley, New York, 2004).
[40] T. Schumm, S. Hofferberth, L. Mauritz Andersson, S. Wildermuth, S. Groth, I. Bar-Joseph, J. Schmiedmayer, and P. Krüger, *Matter-Wave Interferometry in a Double Well on an Atom Chip*, Nat. Phys. **1**, 57 (2005).
[41] J. Sebby-Strabley, B. L. Brown, M. Anderlini, P. J. Lee, W. D. Phillips, J. V. Porto, and P. R. Johnson, *Preparing and Probing Atomic Number States with an Atom Interferometer*, Phys. Rev. Lett. **98**, 200405 (2007).
[42] H. Zhang, R. McConnell, S. Ćuk, Q. Lin, M. H. Schleier-Smith, I. D. Leroux, and V. Vuletić, *Collective State Measurement of Mesoscopic Ensembles with Single-Atom Resolution*, Phys. Rev. Lett. **109**, 133603 (2012).
[43] I. Stroescu, D. B. Hume, and M. K. Oberthaler, *Double-Well Atom Trap for Fluorescence Detection at the Heisenberg Limit*, Phys. Rev. A **91**, 013412 (2015).
[44] F. Fröwis and W. Dür, *Measures of Macroscopicity for Quantum Spin Systems*, New J. Phys. **14**, 093039 (2012).
[45] F. Fröwis and W. Dür, *Are Cloned Quantum States Macroscopic?*, Phys. Rev. Lett. **109**, 170401 (2012).
[46] M. C. Tichy, C.-Y. Park, M. Kang, H. Jeong, and K. Mølmer, *Is Macroscopic Entanglement a Typical Trait of Many-Particle Quantum States?*, Phys. Rev. A **93**, 042314 (2016).
[47] G. Chiribella, Y. Yang, and A. C.-C. Yao, *Quantum Replication at the Heisenberg Limit*, Nat. Commun. **4**, 2915 (2013).
[48] W. Dür, P. Sekatski, and M. Skotiniotis, *Deterministic Superreplication of One-Parameter Unitary Transformations*, Phys. Rev. Lett. **114**, 120503 (2015).
[49] M. Ledoux, *The Concentration of Measure Phenomenon* (American Mathematical Society, Providence, 2005), Vol. 89.
[50] G. W. Anderson, A. Guionnet, and O. Zeitouni, *An Introduction to Random Matrices* (Cambridge University Press, Cambridge, England, 2010), p. 118.
[51] B. Collins and I. Nechita, *Random Matrix Techniques in Quantum Information Theory*, J. Math. Phys. (N.Y.) **57**, 015215 (2016).
[52] J. Bourgain and A. Gamburd, *A Spectral Gap Theorem in $SU(d)$*, arXiv:1108.6264.
[53] R. A. Low, *Large Deviation Bounds for k-designs*, Proc. R. Soc. A **465**, 3289 (2009).
[54] R. A. Low, Ph.D. thesis, University of Bristol, 2010.
[55] F. G. S. L. Brandao, A. W. Harrow, and M. Horodecki, *Local Random Quantum Circuits Are Approximate Polynomial-Designs*, Phys. Rev. Lett. **116**, 170502 (2016).
[56] F. Fröwis, M. Skotiniotis, B. Kraus, and W. Dür, *Optimal Quantum States for Frequency Estimation*, New J. Phys. **16**, 083010 (2014).
[57] E. L. Lehmann and G. Casella, *Theory of Point Estimation* (Springer, New York, 1998), Vol. 31.
[58] R. D. Gill and B. Y. Levit, *Applications of the van Trees Inequality: A Bayesian Cramér-Rao Bound*, Bernoulli **1**, 59 (1995).
[59] S. L. Braunstein and C. M. Caves, *Statistical Distance and the Geometry of Quantum States*, Phys. Rev. Lett. **72**, 3439 (1994).
[60] S. Boixo, S. T. Flammia, C. M. Caves, and J. M. Geremia, *Generalized Limits for Single-Parameter Quantum Estimation*, Phys. Rev. Lett. **98**, 090401 (2007).
[61] Á. Rivas and A. Luis, *Precision Quantum Metrology and Nonclassicality in Linear and Nonlinear Detection Schemes*, Phys. Rev. Lett. **105**, 010403 (2010).
[62] M. Zwierz, C. A. Pérez-Delgado, and P. Kok, *General Optimality of the Heisenberg Limit for Quantum Metrology*, Phys. Rev. Lett. **105**, 180402 (2010).
[63] A. Fujiwara, *Strong Consistency and Asymptotic Efficiency for Adaptive Quantum Estimation Problems*, J. Phys. A **39**, 12489 (2006).
[64] E. M. Kessler, I. Lovchinsky, A. O. Sushkov, and M. D. Lukin, *Quantum Error Correction for Metrology*, Phys. Rev. Lett. **112**, 150802 (2014).
[65] J. B. Brask, R. Chaves, and J. Kołodyński, *Improved Quantum Magnetometry Beyond the Standard Quantum Limit*, Phys. Rev. X **5**, 031010 (2015).
[66] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, Cambridge, England, 2006).
[67] S. Popescu, A. J. Short, and A. Winter, *Entanglement and the Foundations of Statistical Mechanics*, Nat. Phys. **2**, 754 (2006).
[68] D. W. Berry and H. M. Wiseman, *Optimal States and Almost Optimal Adaptive Measurements for Quantum Interferometry*, Phys. Rev. Lett. **85**, 5098 (2000).
[69] E. Bagan, M. Baig, and R. Muñoz Tapia, *Aligning Reference Frames with Quantum States*, Phys. Rev. Lett. **87**, 257903 (2001).
[70] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, *Reference Frames, Superselection Rules, and Quantum Information*, Rev. Mod. Phys. **79**, 555 (2007).
[71] A. T. Rezakhani and S. Alipour, *On Continuity of Quantum Fisher Information*, arXiv:1507.01736.
[72] T. Baumgratz and A. Datta, *Quantum Enhanced Estimation of a Multidimensional Field*, Phys. Rev. Lett. **116**, 030801 (2016).
[73] I. Marvian and R. W. Spekkens, *Extending Noether's Theorem by Quantifying the Asymmetry of Quantum States*, Nat. Commun. **5**, 3821 (2014).
[74] Note that $\hat{J}_x = \frac{1}{2}(\hat{a}^\dagger \hat{b} + \hat{b}^\dagger \hat{a})$, $\hat{J}_y = \frac{1}{2}(\hat{b}^\dagger \hat{a} - \hat{b}^\dagger \hat{a})$, and $\hat{J}_z = \frac{1}{2}(\hat{a}^\dagger \hat{a} - \hat{b}^\dagger \hat{b})$ are the standard two-mode angular-momentum operators defined via the Jordan-Schwinger map [3].
[75] A. Lubotzky, R. Phillips, and P. Sarnak, *Hecke Operators and Distributing Points on the Sphere I*, Commun. Pure Appl. Math. **39**, S149 (1986).
[76] A. W. Harrow, B. Recht, and I. L. Chuang, *Efficient Discrete Approximations of Quantum Gates*, J. Math. Phys. (N.Y.) **43**, 4445 (2002).
[77] A. Bouland and S. Aaronson, *Generation of Universal Linear Optics by Any Beam Splitter*, Phys. Rev. A **89**, 062316 (2014).
[78] J. Carolan, C. Harrold, C. Sparrow, E. Martin-Lopez, N. J. Russell, J. W. Silverstone, P. J. Shadbolt, N. Matsuda,

M. Oguma, M. Itoh *et al.*, *Universal Linear Optics*, Science **349**, 711 (2015).

[79] A. Sawicki, *Universality of Beamsplitters*, Quantum Inf. Comput. **16**, 0291 (2016).

[80] N. Imoto, H. A. Haus, and Y. Yamamoto, *Quantum Nondemolition Measurement of the Photon Number via the Optical Kerr Effect*, Phys. Rev. A **32**, 2287 (1985).

[81] M. Oszmaniec and Z. Zimborás, *Universal Extensions of Restricted Classes of Quantum Operations* (unpublished).

[82] M. Albiez, R. Gati, J. Fölling, S. Hunsmann, M. Cristiani, and M. K. Oberthaler, *Direct Observation of Tunneling and Nonlinear Self-Trapping in a Single Bosonic Josephson Junction*, Phys. Rev. Lett. **95**, 010402 (2005).

[83] M. Melé-Messeguer, B. Juliá-Díaz, M. Guilleumas, A. Polls, and A. Sanpera, *Weakly Linked Binary Mixtures of $f=1$ 87-rb Bose-Einstein Condensates*, New J. Phys. **13**, 033012 (2011).

[84] B. Juliá-Díaz, T. Zibold, M. K. Oberthaler, M. Melé-Messeguer, J. Martorell, and A. Polls, *Dynamic Generation of Spin-Squeezed States in Bosonic Josephson Junctions*, Phys. Rev. A **86**, 023615 (2012).

[85] L. Pitaevskii and S. Stringari, *Bose-Einstein Condensation* (Clarendon Press, Oxford, 2003).

[86] P. Kok and B. W. Lovett, *Introduction to Optical Quantum Information Processing* (Cambridge University Press, Cambridge, England, 2010).

[87] A. Furusawa and P. van Loock, *Quantum Teleportation and Entanglement: A Hybrid Approach to Optical Quantum Information Processing* (Wiley-VCH, Weinheim, 2011).

[88] X.-W. Wang, D.-Y. Zhang, S.-Q. Tang, L.-J. Xie, Z.-Y. Wang, and L.-M. Kuang, *Photonic Two-Qubit Parity Gate with Tiny Cross-Kerr Nonlinearity*, Phys. Rev. A **85**, 052326 (2012).

[89] K. Nemoto and W. J. Munro, *A Near Deterministic Linear Optical CNOT Gate*, Phys. Rev. Lett. **93**, 250502 (2004).

[90] B. He, Q. Lin, and C. Simon, *Cross-Kerr Nonlinearity Between Continuous-Mode Coherent States and Single Photons*, Phys. Rev. A **83**, 053826 (2011).

[91] D. E. Chang, V. Vuletić, and M. D. Lukin, *Quantum Nonlinear Optics—Photon by Photon*, Nat. Photonics **8**, 685 (2014).

[92] P. Lodahl, S. Mahmoodian, and S. Stobbe, *Interfacing Single Photons and Single Quantum Dots with Photonic Nanostructures*, Rev. Mod. Phys. **87**, 347 (2015).

[93] A. Javadi, I. Söllner, M. Arcari, S. L. Hansen, L. Midolo, S. Mahmoodian, G. Kiršanskė, T. Pregnolato, E. H. Lee, J. D. Song, S. Stobbe, and P. Lodahl, *Single-Photon Nonlinear Optics with a Quantum Dot in a Waveguide*, Nat. Commun. **6**, 8655 (2015).

[94] Z. Zimborás, R. Zeier, M. Keyl, and T. Schulte-Herbrüggen, *A Dynamic Systems Approach to Fermions and Their Relation to Spins*, Eur. Phys. J. Quantum Technology **1**, 11 (2014).

[95] Z. Zimborás, R. Zeier, T. Schulte-Herbrüggen, and D. Burgarth, *Symmetry Criteria for Quantum Simulability of Effective Interactions*, Phys. Rev. A **92**, 042309 (2015).

[96] P. Cwiklinski, M. Horodecki, M. Mozrzymas, L. Pankowski, and S. Michal, *Local Random Quantum Circuits Are Approximate Polynomial-Designs: Numerical Results*, J. Phys. A **46**, 305301 (2013).

[97] This confirms the scaling obtained numerically in Ref. [46] and is consistent with $2N(N+2)/3$ found in Ref. [56], in which the states were taken to be mode symmetric.

[98] M. D. Lang and C. M. Caves, *Optimal Quantum-Enhanced Interferometry Using a Laser Power Source*, Phys. Rev. Lett. **111**, 173601 (2013).

[99] A. Monras and F. Illuminati, *Information Geometry of Gaussian Channels*, Phys. Rev. A **81**, 062326 (2010).

[100] O. Pinel, J. Fade, D. Braun, P. Jian, N. Treps, and C. Fabre, *Ultimate Sensitivity of Precision Measurements with Intense Gaussian Quantum Light: A Multimodal Approach*, Phys. Rev. A **85**, 010101 (2012); O. Pinel, P. Jian, N. Treps, C. Fabre, and D. Braun, *Quantum Parameter Estimation Using General Single-Mode Gaussian States*, Phys. Rev. A **88**, 040102 (2013).

[101] M. Jarzyna and J. Kołodyński, *Geometrical Approach to Quantum Statistical Inference* (unpublished).

[102] C. A. Fuchs and J. Van De Graaf, *Cryptographic Distinguishability Measures for Quantum-Mechanical States*, IEEE Trans. Inf. Theory **45**, 1216 (1999).

[103] P. Migdał, J. Rodríguez-Laguna, M. Oszmaniec, and M. Lewenstein, *Multiphoton States Related via Linear Optics*, Phys. Rev. A **89**, 062329 (2014).

[105] I. N. Bronštejn and K. A. Semendjaev, *Handbook of Mathematics* (Springer, New York, 2013).

[106] H. Gould, Tables of Combinatorial Identities, edited by J. Quaintance (2010), Vol. **7**, http://www.math.wvu.edu/~gould.

[107] K. Zyczkowski and H.-J. Sommers, *Truncations of Random Unitary Matrices*, J. Phys. A **33**, 2045 (2000).