Protocol for global multiphase estimation

Giovanni Chesi ⬤ and Alberto Riccardi
National Institute for Nuclear Physics, Sezione di Pavia, Via Agostino Bassi 6, 27100 Pavia, Italy

Roberto Rubboli
Centre for Quantum Technologies, National University of Singapore, 117543 Singapore

Lorenzo Maccone and Chiara Macchiavello
Dipartimento di Fisica, Università degli Studi di Pavia, Via Agostino Bassi 6, 27100 Pavia, Italy
and National Institute for Nuclear Physics, Sezione di Pavia, Via Agostino Bassi 6, 27100 Pavia, Italy

Global estimation strategies allow one to extract information on a phase or a set of phases without any prior knowledge, which is required for local estimation strategies. We devise a global multiphase protocol based on Holevo's estimation theory and apply it to the case of digital estimation, i.e., we estimate the phases in terms of the mutual information between them and the corresponding estimators. We test the protocol in two particular cases, i.e., the single-phase and the double-phase estimation. In the single-phase scenario, the protocol encompasses two specific known optimal strategies. We extend them to the simultaneous estimation of two phases and evaluate their performance. Then we retrieve the ultimate digital bound on precision when a generic number of phases is simultaneously estimated. This bound is again expressed in terms of mutual information and is general for any digital multiphase estimation protocol. We show that in the multiphase strategy there is only a constant quantum advantage with respect to a sequence of independent single-phase estimations. This extends a recent similar result, which settled a controversy on the search for the multiphase enhancement.

I. INTRODUCTION

Quantities that in quantum mechanics are not described by operators require the search for good estimators revealing the value of the corresponding parameters [1]. The efficiency of an estimation strategy is assessed by the Heisenberg bound [2,3], which provides the ultimate achievable precision.

In this context, we have a general framework for local estimation, where an approximate knowledge of the parameter at least is assumed through, for instance, preliminary coarse measurements that the local estimation aims to improve [4,5]. The idea is to study the dispersion of the estimator in the neighborhood of the parameter, typically through the Fisher information and the Cramér-Rao bound. The theory has been successfully specified in the quantum formalism and extended to the multiparameter estimation, where more than one parameter is simultaneously estimated [4,6–8].

The same cannot be said for global estimation theories, since generally global models are exactly solved only in special cases [9]. In global estimation theories, no *a priori* knowledge of the parameter is assumed and the goal is to investigate the quality of an estimation procedure defined by a measurement on an input state where the parameter is encoded. One can do this by describing the parameter and estimator as phases and averaging a suitable cost function over the whole range. Much effort has been spent to apply a Bayesian approach to local theories, thus accounting for uncertainty on the knowledge of the parameter [9,10]. This approach produced many useful bounds on specific cost

functions (typically the mean-square error), such as the Ziv-Zakai bounds, originated in Ref. [11]. Still, all these criteria are based on preliminary knowledge about the parameter to be estimated. Useful group-covariant methods have been developed to assess a relation between Cramér-Rao bounds and the asymptotic bound of the error between the estimator and parameter [12–15].

As concerns genuine global Bayesian estimation theories, the seminal work by Holevo [16] defined the framework and the first main results for a restricted but significant class of cost functions in the case of a nondegenerate generator and single-parameter estimation. Then the case of a degenerate generator was considered and accomplished in Ref. [17]. A first extension to the multiparameter scenario grounded in Holevo's global theory was developed in Ref. [18]. Only a few investigations followed in this direction. A generalization of the quantum phase estimation algorithm (QPEA) to the multiphase context was recently provided in Ref. [19]. A Bayesian covariant estimation was provided in Ref. [9], where Demkowicz-Dobrzaski *et al.* found the positive operator-valued measures (POVMs) optimizing cost functions related to the quantum state fidelity in the specific case of pure and mixed qubits as probes. Interestingly, a couple of years later Górecki and Dobrzański exploited a minimax approach to generalize a local phase estimation to the global covariant scenario [20]. In particular, they retrieved the Heisenberg bound scaling on the sum of the squared errors in the multiphase case and proved that the advantage of simultaneous multiphase estimation with respect to independent single-phase

estimations amounts to a constant factor. This result created a controversy, since it contradicts the multiphase enhancement previously claimed both in local frameworks [6–8], where an intrinsic advantage scaling with the number of simultaneously estimated phases was found, and in the global framework of the multiphase QPEA [19].

Here we devise a protocol for global multiphase estimation with commuting generators based on the extension of Holevo's optimal POVM [16] to the multiphase case exploited in Ref. [18]. The protocol is more general than other procedures, such as the ones in Refs. [18,20], because it encompasses diverse well-known global single-phase estimation strategies. Indeed, the protocol fixes the POVM only so that, by suitably changing the input state, it outputs the statistics related to the corresponding strategy, defined by the desired cost function. Estimation strategies can be generally differentiated according to the application of the encoding unitaries into parallel and sequential [3,21]. In parallel configurations, we encode the parameter on a number N of inputs employed all together, while in the case of a sequential strategy we use the encoding unitary N times on the same input, possibly coupled to ancillary systems. As a preliminary test for our protocol, we show how a suitable choice of the probes can reduce it to known optimal parallel and sequential strategies.

In principle, if the estimator is a continuous variable, our protocol sets no limitations on the cost function to be optimized. If the estimator is discrete, the protocol needs to satisfy several assumptions about the cost function, which are detailed in Appendix A, in order to keep the Bayesian average of the cost function invariant under discretization. Here we take a discrete estimator and apply the protocol to the mutual information between the parameter ϕ and the estimator $\tilde{\phi}$, which can be defined as [22]

$$I(\tilde{\phi} : \phi) = H(\tilde{\phi}) - H(\tilde{\phi}|\phi), \qquad (1)$$

where $H(\tilde{\phi})$ is the entropy of the estimator and $H(\tilde{\phi}|\phi)$ is the conditional entropy of the estimator with respect to the parameter. Every estimation strategy is intrinsically dependent on the figure of merit one wants to optimize, since it is the indicator of the quality of the estimation itself. For a global estimation, the mutual information is a good candidate because it does not depend on the parameter and has a clear physical meaning: It is the number of bits shared by the parameter and estimator. An estimation theory based on the mutual information is known as digital estimation. Digital quantum estimation for sequential and parallel strategies was developed in the single-parameter case in Ref. [21]. In Ref. [23] the relative entropy was used as a precision criterion for the generation of approximating states in the case of mixed states. Another well-established case where the mutual information is used as a figure of merit for phase estimation is the rate-distortion theory [24,25]. Note that in rate-distortion theories typically the cost function employed is the so-called distortion, defined as a function of the difference between the estimator and parameter, and the minimum of the mutual information over the conditional distribution $p(\tilde{\phi}|\phi)$ is used to lower bound the distortion through Shannon's joint source-channel coding theorem [22]. Here instead we will use the conditional entropy defining the mutual information itself as a cost function. As noted by Hall in Ref. [26], it is quite

surprising that seldom has a quantum metrology problem been expressed in information-theoretic terms, since the mutual information, differently from the Cramér-Rao bound, is deeply and directly connected to the original information theory as introduced by Shannon. Hall and Wiseman found interesting equivalences between the information-theoretic framework for phase estimation and the Fisher-information formalism [27]. In particular, they successfully exploited the Holevo information bound as an Heisenberg limit on the mutual information and found that the latter is limited from above by the entropy of the generator for the probe state.

Finally and more importantly, our protocol can be naturally extended to the multiparameter case, thus providing a benchmark for multiphase estimation. Here we use it to investigate the simplest nontrivial case, which is the simultaneous estimation of two phases, but we remark that it can be applied to an arbitrary number of phases. We find that double-phase estimation slightly outperforms independent single-phase procedures. Hence, we move to the case of a generic number of phases and assess the ultimate limit on multiphase estimation in information-theoretic terms, i.e., we derive Heisenberg bounds on the mutual information as a consequence of the Holevo information bound, which was exploited in this context also in Refs. [21,26,27]. We remark that the bound is not derived from a specific estimation strategy, i.e., it is not limited to our estimation protocol. This last result is relevant because it helps to clarify the controversy about the multiphase enhancement mentioned above. On the one hand, we find that it is true that the maximum achievable information grows with the number of simultaneously estimated phases. On the other hand, by inspecting the effective information gain for each estimated phase, we find that asymptotically it amounts to a constant factor, as claimed in Ref. [20].

The paper is organized as follows. In Sec. II we review the basic framework of a global estimation theory and how it can be extended to account for degenerate eigenstates of the generator. Then we present our protocol in Sec. III. In Sec. IV we focus on the single-parameter estimation and retrieve two existing strategies from our protocol. Our analysis in this context shows a trade-off between optimal sequential and parallel-separable strategies. We move to the double-phase estimation scenario in Sec. V, where we derive a direct generalization of the optimal sequential and parallel protocols retrieved in the preceding section and compare them. Moreover, we evaluate them with respect to the single-phase estimation case. Sections IV and V are both divided in two parts. In Secs. IV A and V A we identify the input states and the POVM exploited, while in Secs. IV B and V B we show what is obtained in terms of mutual information by applying our protocol. In Sec. VI we consider the general case of estimating k phases with N encoding operations and investigate how the ultimate bound on precision depends on these parameters. We summarize in Sec. VII.

II. GLOBAL ESTIMATION THEORY WITH DEGENERATE STATES

Here we review Holevo's global quantum estimation theory [16]. The aim is to find a POVM minimizing a given functional describing the cost of the estimation, i.e., how

much information is lost in the estimation process. In particular, we focus on the so-called Bayesian uniform mean deviation [16], i.e.,

$$C = \int_0^{2\pi} \frac{d\phi}{2\pi} \int_0^{2\pi} \frac{d\tilde{\phi}}{2\pi} c(\phi, \tilde{\phi}) \text{Tr}(\rho_\phi \Pi_{\tilde{\phi}}), \qquad (2)$$

where ϕ is the parameter to be estimated, $\tilde{\phi}$ is the estimator, $c(\phi, \tilde{\phi})$ is a cost function, and the conditional probability density of measuring $\tilde{\phi}$ given ϕ is

$$\text{Pr}(\tilde{\phi}|\phi) = \text{Tr}(\rho_\phi \Pi_{\tilde{\phi}}), \qquad (3)$$

with ρ_ϕ the state where the phase is encoded and $\Pi_{\tilde{\phi}}$ an element of the POVM $\Pi = \{\Pi_{\tilde{\phi}}\}_{\tilde{\phi}}$ describing the measurement.

Holevo's POVM is optimal in the context of covariant phase estimation problems, i.e., in the case where both the state and the measurement are covariant. In particular, the state ρ_{ϕ_0} is transformed by a unitary representation of the phase-shift group U_ϕ as follows:

$$U_\phi \rho_{\phi_0} U_\phi^\dagger = \rho_{\phi+\phi_0}. \qquad (4)$$

The covariance of the POVM, on the other hand, implies

$$\text{Tr}(\rho_{\phi_0} \Pi_{\tilde{\phi}}) = \text{Tr}(\rho_{\phi_0+\phi} \Pi_{\tilde{\phi}+\phi}) \, \forall \phi \in [0, 2\pi]. \qquad (5)$$

Hence, note that in a covariant estimation problem the conditional probability density $\text{Pr}(\tilde{\phi}|\phi)$ depends just on the difference $\tilde{\phi} - \phi$ since

$$\text{Tr}(\rho_\phi \Pi_{\tilde{\phi}}) = \text{Tr}(\rho_0 \Pi_{\tilde{\phi}-\phi}). \qquad (6)$$

Furthermore, the optimality of Holevo's POVM holds for a wide but specific class of cost functions, which we call here the Holevo class. These functions are periodic, even, and again depend on the difference $\tilde{\phi} - \phi$ only. In other words, their Fourier expansion is given by

$$c(\phi, \tilde{\phi}) = c_0 + \sum_{k=0}^\infty c_k \cos[k(\tilde{\phi} - \phi)], \qquad (7)$$

with the following further assumptions on the coefficients: $c_0 \geqslant 0$ and $c_k \leqslant 0 \, \forall k \geqslant 1$.

Note that, given this set of conditions, the Bayesian mean deviation defined in Eq. (2) reduces to an integral over a single variable. However, we will keep it in its general form (2) since in the following we will not need all of the conditions that guarantee the optimality of Holevo's POVM.

Holevo showed in Ref. [16] that, if the estimation problem is covariant and the cost function belongs to the Holevo class, the differential POVM minimizing the cost functional in Eq. (2) is

$$d\mu(\tilde{\phi}) = \frac{d\tilde{\phi}}{2\pi} |e(\tilde{\phi})\rangle\langle e(\tilde{\phi})|, \qquad (8)$$

where

$$|e(\tilde{\phi})\rangle \equiv \sum_{n=0}^\infty e^{in\tilde{\phi}} |n\rangle \qquad (9)$$

are the Susskind-Glogower vectors and $\{|n\rangle\}_n$ is the set of eigenstates of the generator H of the transformations U_ϕ. The optimal POVM in Eq. (8) can be generalized to the case where the generator is degenerate through the projection method

[17]. We can easily show in which sense the generator can be degenerate and how the method works if we focus on the equatorial qubit state

$$|\psi_0\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \qquad (10)$$

and take N identical copies of it as a probe $|\tilde{\psi}_0\rangle$, i.e., $|\tilde{\psi}_0\rangle \equiv |\psi_0\rangle^{\otimes N}$. In such a case, an encoding operation as $U_\phi^{\otimes N}$ with generator $H = |1\rangle\langle 1|$ would provide the same eigenvalues for different states defined by the same number of $|1\rangle$ states. We can remove the degeneracy as follows. We define $\mathcal{H} \equiv \mathbb{C}^{\otimes N}$ as the Hilbert space where $|\tilde{\psi}_0\rangle$ is defined. We consider the nondegenerate subspace $\mathcal{H}_\parallel$ of $\mathcal{H}$ such that $\mathcal{H} = \mathcal{H}_\parallel \otimes \mathcal{H}_\perp$ and $\mathcal{H}_\parallel$ is spanned by the normalized vectors $|n_0, n_1\rangle_\parallel \propto P_{n_0, n_1} |\Psi_0\rangle$, where n_0 and n_1 are the numbers of zeros and of ones, respectively, $|\Psi_0\rangle$ is an arbitrary pure initial state, and P_{n_0, n_1} is a projector onto the degenerate eigenspace of dimension $\binom{N}{n_1}$ generated from all the states defined by the tensor product of n_1 states $|1\rangle$ and $n_0 = N - n_1$ states $|0\rangle$. Then the vectors $|n_0, n_1\rangle_\parallel$ spanning $\mathcal{H}_\parallel$ can be defined as

$$|n_0, n_1\rangle_\parallel = |N - n, n\rangle_\parallel \equiv |n\rangle_\parallel$$
$$\equiv \frac{1}{\sqrt{\lambda_n}} \sum_{\nu_1=0}^1 \cdots \sum_{\nu_N=0}^1 \delta\left(\sum_k \nu_k - n\right) \bigotimes_{k=1}^N |\nu_k\rangle, \qquad (11)$$

where $\lambda_n \equiv \binom{N}{n}$ is a symmetrization factor accounting for the multiplicity of the corresponding eigenvalue. Note that here we relabeled $n_1 = n$, i.e., n is the number of ones. The POVM can be chosen in a block-diagonal form on $\mathcal{H}$, i.e., $d\mu(\tilde{\phi}) = d\mu_\parallel(\tilde{\phi}) \bigoplus d\mu_\perp(\tilde{\phi})$, where $d\mu_\perp(\tilde{\phi})$ can be an arbitrary POVM on $\mathcal{H}_\perp$, since, being defined on the states orthogonal to $|\Psi_0\rangle$, it does not contribute to $\text{Pr}(\tilde{\phi}|\phi)$. Therefore, the problem is reduced to finding the optimal $d\mu_\parallel(\tilde{\phi})$. It was proved in Ref. [17] that the form of the POVM is identical to the one in the nondegenerate case with the eigenstates of H replaced by the basis of $\mathcal{H}_\parallel$ defined in Eq. (11), i.e.,

$$d\mu_\parallel(\tilde{\phi}) = \frac{d\tilde{\phi}}{2\pi} |E(\tilde{\phi})\rangle\langle E(\tilde{\phi})|, \qquad (12)$$

with

$$|E(\tilde{\phi})\rangle \equiv \sum_{n=0}^N e^{in\tilde{\phi}} |n\rangle_\parallel. \qquad (13)$$

III. MULTIPHASE GLOBAL ESTIMATION PROTOCOL

A. Protocol

The global estimation protocol we are introducing here fixes the measurement to the Holevo POVM and optimizes a given figure of merit over the set of input states. It can be used for the estimation of a single phase and, most importantly, for the joint estimation of an arbitrary number of phases. In this regard, it is a multiphase estimation protocol. The algorithm can be described as follows.

(i) Fix a set of commuting generators H_j defining the unitary representations U_{ϕ_j} of the phase-shift group with the corresponding set of phases to be estimated $\{\phi_j\}_{j=1}^k \equiv \phi$.

(ii) Using the procedure described above, build the Holevo POVM $d\mu_\parallel(\tilde{\phi})$ on the nondegenerate set of eigenstates of the generators for the set of estimators $\{\tilde{\phi}_j\}_{j=1}^k \equiv \tilde{\phi}$.

(iii) Choose a figure of merit for the estimation process.

(iv) Optimize the figure of merit over the set of input states $\rho_0 \in \mathcal{S}(\mathcal{H}_\parallel)$, with $\mathcal{S}(\cdot)$ the set of quantum states, through the Bayesian mean deviation, i.e.,

$$C = \min_{\rho_0}\left[\int d\boldsymbol{\phi}\,\text{Tr}\left(\int d\mu_\parallel(\tilde{\boldsymbol{\phi}})c(\phi,\tilde{\boldsymbol{\phi}})U_{\boldsymbol{\phi}}^{\otimes N}\rho_0(U_{\boldsymbol{\phi}}^\dagger)^{\otimes N} \right) \right], \tag{14}$$

where $U_{\boldsymbol{\phi}} \equiv \exp(i\sum_j H_j\phi_j)$ and the figure of merit is established by the cost function $c(\boldsymbol{\phi}, \tilde{\boldsymbol{\phi}})$.

We remark that Holevo's POVM optimizes a specific class of cost functions for a generic pure input state over the set of POVMs [16]. Our protocol applies Holevo's POVM to a generic figure of merit and finds the input state optimizing that figure of merit.

B. Application to the mutual information

As mentioned above, here our figure of merit is the mutual information $I(\tilde{\boldsymbol{\phi}} : \boldsymbol{\phi})$. The mutual information has the opposite meaning of a cost function since it provides the number of bits of the parameters gained through the estimation procedure. However, it is built by definition on a Bayesian mean deviation, with the so-called surprise as a cost function, i.e.,

$$c(\boldsymbol{\phi}, \tilde{\boldsymbol{\phi}}) = -\log_2 \text{Pr}(\tilde{\boldsymbol{\phi}}|\boldsymbol{\phi}). \tag{15}$$

Note that, despite being covariant, in general it does not belong to the Holevo class, since it is generally not periodic or even. The Bayesian mean deviation whose cost function is the surprise is the conditional entropy $H(\tilde{\boldsymbol{\phi}}|\boldsymbol{\phi})$, which provides the number of bits that we have to pay for the estimation. The mutual information is a gain function, i.e., it is the difference between the maximum number of phase bits $H(\tilde{\boldsymbol{\phi}})$ that we can get (fixed by the number of resources N that we exploit for the estimation) and the number of bits lost in the estimation procedure, provided by the conditional entropy.

The minimization of the conditional entropy over the set of input states is not a simple task. However, we can get an idea of the optimal pure state in the single-phase case through a local minimization of the cost function in Eq. (15) for $\tilde{\phi} \sim \phi$, i.e., we can look for the state $|\Psi_0\rangle$ such that

$$c(\phi, \tilde{\phi} \sim \phi) \geqslant \min_{\Psi_0}[-\log_2 |\langle E(\tilde{\phi})|U_\phi|\Psi_0\rangle|^2]. \tag{16}$$

Since the logarithm is a monotonic function, this minimization is equivalent to maximizing the distribution $\text{Pr}(\tilde{\phi} \sim \phi|\phi) = |\langle E(\tilde{\phi})|U_\phi|\Psi_0\rangle|^2$, which, via Lagrange multipliers, readily gives as the optimal state a uniform superposition of the nondegenerate states $\{|n\rangle_\parallel\}_{n=0}^N$, known as the Holland-Burnett state [28], i.e.,

$$|\Psi_0\rangle_{\text{opt}} \equiv \frac{1}{\sqrt{N+1}}\sum_{n=0}^N |n\rangle_\parallel. \tag{17}$$

We stress the fact that this state is locally optimal for the specific cost function defining the mutual information, i.e., the surprise in Eq. (15). Different cost functions which have been extensively explored in the literature, such as the dispersion and the maximum likelihood [12,13,29] or a generic Fourier expansion of Holevo cost functions [30,31], require one to consider different inputs to achieve the Heisenberg scaling. Our protocol can be applied to those contexts as well. Here we test it in the case of the digital estimation.

As mentioned in the Introduction, on the one hand, our protocol based on Holevo's POVM allows us to retrieve known global single-phase estimation strategies and, on the other hand, it paves the way for multiparameter global estimation.

IV. SINGLE-PHASE ESTIMATION

A. Theory

Here we show that, by suitably changing the probe state, our protocol allows us to retrieve as specific cases two known optimal single-parameter parallel and sequential estimation strategies and, to this aim, we introduce a notion of equivalence between estimation protocols.

Definition 1. Two estimation protocols *A* and *B* are equivalent if the output probability distribution of the estimator $\tilde{\phi}$ conditioned on the parameter ϕ is the same, i.e., if

$$\text{Pr}_A(\tilde{\phi}|\phi) = \text{Pr}_B(\tilde{\phi}|\phi). \tag{18}$$

In particular, we test our protocol with two different probes. By comparing the resulting conditional probabilities with the ones of known single-parameter estimation protocols, we will establish an equivalence through Definition 1. Hence, we can appreciate the generality of this approach.

1. Preparation

We consider first a probe state which is known to achieve the standard quantum limit and second a state saturating the Heisenberg bound [21]. The former is a probe which is very commonly exploited in the parallel estimation strategies, which we mentioned in the Introduction. It is defined by N copies of the equatorial state $|\psi_0\rangle$ in Eq. (10), i.e., the separable state

$$|\Psi_0\rangle^{(1)} \equiv |\psi_0\rangle^{\otimes N} = \frac{1}{\sqrt{2^N}}\sum_{n=0}^N \sqrt{\binom{N}{n}}|n\rangle_\parallel. \tag{19}$$

As for the latter, we take the Holland-Burnett states in Eq. (17) since we found in the previous section that they locally optimize the mutual information, namely,

$$|\Psi_0\rangle^{(2)} \equiv \frac{1}{\sqrt{N+1}}\sum_{n=0}^N |n\rangle_\parallel. \tag{20}$$

The Holland-Burnett states are entangled (see Appendix B), which implies that we need the N probe states to be entangled to achieve the Heisenberg bound. The phase is encoded through $U_\phi^{\otimes N}|\Psi_0\rangle^{(x)} \equiv |\Psi_\phi\rangle^{(x)}$ with $x = 1, 2$ and $U_\phi \equiv \exp(iH\phi)$ with generator $H = |1\rangle\langle 1|$.

2. Measurement

Our protocol provides an estimation of the parameter ϕ through Holevo's POVM. The probability density of the estimator $\tilde{\phi}$ conditioned on the value of the parameter ϕ can be

retrieved from

$$\Pr(\tilde{\phi}|\phi)\frac{d\tilde{\phi}}{2\pi} = \text{Tr}\big[d\mu(\tilde{\phi})_\| \rho_\phi^{(x)}\big], \qquad (21)$$

with $\rho_\phi^{(x)} = |\Psi_\phi\rangle\langle\Psi_\phi|^{(x)}$ and $x = 1, 2$.

If the phase is encoded on $\rho_\phi^{(1)}$, we find that the distribution of the estimator conditioned on the parameter is

$$\Pr^{(1)}(\tilde{\phi}|\phi) = \left|\frac{1}{\sqrt{2^N}}\sum_{n=0}^{N}\sqrt{\binom{N}{n}}e^{i(\tilde{\phi}-\phi)n}\right|^2. \qquad (22)$$

On the other hand, if we take $\rho_\phi^{(2)}$, we find

$$\Pr^{(2)}(\tilde{\phi}|\phi) = \frac{1}{N+1}\frac{\sin^2[(N+1)\pi(\phi-\tilde{\phi})]}{\sin^2[\pi(\phi-\tilde{\phi})]}. \qquad (23)$$

Since our figure of merit is the mutual information, here we consider the case of a discrete estimator $\tilde{\phi}/2\pi = m/(N+1)$ with $m \in [0, N]$. We show in Appendix A how the discretization of the distributions and of the cost functional works. In particular, we find that the conditional probability density is modified according to the prescription

$$p(m|\phi) = \frac{1}{N+1}\Pr\left(\tilde{\phi} = \frac{2\pi m}{N+1}\bigg|\phi\right), \qquad (24)$$

with the normalizations $\sum_{m=0}^{N}p(m|\phi) = 1$ and $\int_0^{2\pi}\frac{d\tilde{\phi}}{2\pi}\Pr(\tilde{\phi}|\phi) = 1$.

B. Results

1. Equivalence

By applying the prescription in Eq. (24), we find that the conditional probability densities related to the initial states $|\Psi_0\rangle^{(1)}$ and $|\Psi_0\rangle^{(2)}$ for a discrete estimator m are

$$p^{(1)}(m|\phi) = \frac{1}{N+1}\left|\frac{1}{\sqrt{2^N}}\sum_{n=0}^{N}\sqrt{\binom{N}{n}}e^{i[m/(N+1)-\phi]n}\right|^2 \qquad (25)$$

and

$$p^{(2)}(m|\phi) = \frac{1}{(N+1)^2}\frac{\sin^2[(N+1)\pi\phi]}{\sin^2\{\pi[\phi-m/(N+1)]\}}. \qquad (26)$$

We can recognize in these equations the probability densities of two well-known estimation protocols.

The first one $p^{(1)}(m|\phi)$ is the conditional distribution that describes optimized parallel-separable strategies [21]. In parallel strategies the N probes are employed jointly, i.e., the transformation U_ϕ is applied to all the probes together [3]. If the input state is not entangled, we say that the strategy is separable. It is not surprising that we end with the distribution of the optimal parallel-separable strategy since the probe

state in Eq. (19) and the Holevo POVM exactly describe that strategy. Indeed, it is simple to see that the POVM $\{\Pi_{\tilde{\phi}}\}_{\tilde{\phi}}$ with $\Pi_{\tilde{\phi}} \equiv |E(\tilde{\phi})\rangle\langle E(\tilde{\phi})|$, defined after the differential POVM in Eq. (33), is a particular case of Davies' POVM [21,32] maximizing the mutual information in parallel-separable strategies.

On the other hand, the second distribution is the same as the optimal sequential strategy, whose typical application is the well-known QPEA [3,14,21,33–35]. The basic scheme of a sequential strategy is a single probe undergoing sequentially N times the transformation U_ϕ, in the presence of ancillary systems. The QPEA, as an optimal sequential strategy, has been extensively investigated also in the presence of noise, especially in terms of the mean-square error [14]. Note that in this case our protocol is radically different from the standard QPEA in terms of preparation and measurement. Indeed, the QPEA takes as the input state N copies of the equatorial state, which in the nondegenerate space is our state $|\Psi_0\rangle^{(1)}$ in Eq. (19), while our protocol realizes the same output distribution as the QPEA if we choose as an input the state $|\Psi_0\rangle^{(2)}$ in Eq. (20). The encoding unitary required by our protocol is simpler, since the QPEA needs a controlled unitary depending on the qubit position. As a last difference, in the QPEA the phase is measured through an inverse quantum Fourier transform whereas our protocol exploits the Holevo optimal measurement. Note that, differently from the QPEA, our protocol can be straightforwardly extended to the simultaneous estimation of an arbitrary number of phases. Despite the differences between them, we can again establish an operational equivalence between the two protocols based on Definition 1, i.e., we can say that running our protocol with the locally optimal state in Eq. (20) is equivalent to the optimal sequential strategy.

We conclude that the optimal parallel-separable strategy and the QPEA are particular applications of our protocol. In the next section we will briefly recap the main results of these strategies and provide some insights into the nonasymptotic scenario in terms of mutual information.

2. Performance

The performance of single-phase sequential and parallel protocols has been largely explored [3,19,21,33–35]. However, our digital approach can provide a deeper understanding of the topic and a basic reference to be compared to the multiphase case that we will investigate in the following sections. First, let us inspect the distribution $p^{(1)}(m|\phi)$. As far as we know, there is no closed form for that. However, we note that for large N a single period of $p^{(1)}(m|\phi)$ can be approximated as a Gaussian distribution, i.e.,

$$p^{(1)}(m|\phi)\xrightarrow[N\gg1]{}\frac{\sqrt{2\pi N}}{N+1}\exp\left[-2N\pi^2\left(\phi-\frac{m}{N+1}\right)^2\right]. \qquad (27)$$

In this case, the mutual information obtained from the tensor-product state $|\Psi_0\rangle^{(1)}$ is asymptotic to the standard quantum limit $I_{\text{SQL}}(1, N) \equiv (\log_2 N)/2$ [21], where we define $I_{\text{SQL}}(k, N)$ as the standard quantum limit in terms of mutual information when k phases are jointly estimated with N
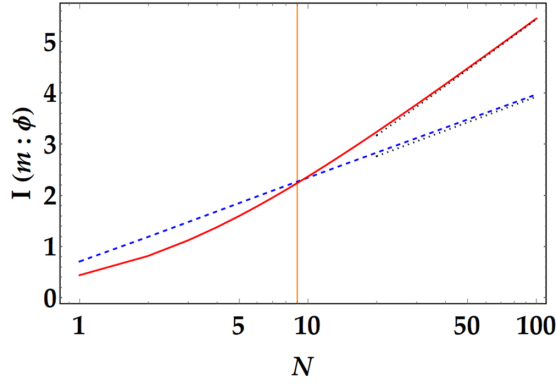
FIG. 1. Base-*e* log-linear plot of the mutual information $I(m : \phi)$ between the estimator m and parameter ϕ in the cases of sequential (red solid line) and parallel-separable (blue dashed line) strategies. The black dotted lines display the asymptotic behavior of the two cases.

resources. In particular, we find

$$
I^{(1)}(m : \phi) \xrightarrow[N \gg 1]{} I_{\text{SQL}}(1, N) + \frac{1}{2}\left(\log_2(2\pi) - \frac{1}{\ln 2}\right)
$$
$$
\sim I_{\text{SQL}}(1, N) + 0.6044. \tag{28}
$$

In contrast, the mutual information of the QPEA is found to be asymptotically nearly optimal, in the sense that it tends to the Heisenberg bound $I_{\text{HB}}(1, N) = \log_2(N + 1)$ [21], where the Heisenberg bound in terms of mutual information $I_{\text{HB}}(k, N)$ depends again on the number of jointly estimated phases and on the number of resources. More specifically, the mutual information provided by the QPEA or, equivalently, by our protocol with $|\Psi_0\rangle^{(2)}$ as a probe state, displays the behavior

$$
I^{(2)}(m : \phi) \xrightarrow[N \gg 1]{} I_{\text{HB}}(1, N) - 2\left(1 - \frac{\gamma + \ln 2 - 1}{\ln 2}\right)
$$
$$
\sim I_{\text{HB}}(1, N) - 1.2199, \tag{29}
$$

where γ is the Euler-Mascheroni constant.

These limits are displayed in Fig. 1 as black dotted and dashed lines, respectively. We note the well-known asymptotic optimality of the sequential strategies over parallel-separable ones [3,21], also apparent from the limits in Eqs. (28) and (29). Interestingly, we remark that for small N here we find a trade-off: If $N \leqslant 9$ the mutual information related to the parallel-separable protocol (blue line) is slightly larger (less than one bit) than the one corresponding to the sequential protocol (red line). This trade-off clearly outlines the limits of a local optimization such as the one proposed in Eq. (16).

V. DOUBLE-PHASE ESTIMATION

A. Theory

Holevo's POVM can be straightforwardly extended to account for more than one estimator [18]. Hence, we can readily generalize the single-phase analysis developed in the preceding section to the multiphase scenario. Here we focus on the simplest case, i.e., the estimation of two phases ϕ and θ

through the corresponding discrete estimators m_ϕ and m_θ. Differently from Ref. [18], we investigate the case of a discrete estimator, as mentioned above, and the figure of merit that best conveys the meaning of digital precision is the mutual information, rather than the fidelity of the estimated state with respect to the probe, which is exploited there. In addition, we do not restrict the analysis to the case of a tensor-product probe state. Due to the optimization procedure performed when applying our protocol, we can focus on the input state that locally optimizes our figure of merit.

1. Preparation

Consider the qutrit nondegenerate space $\mathcal{H}_\parallel$ spanned by the normalized vectors

$$
|n_0, n_1, n_2\rangle_\parallel = |N - n_1 - n_2, n_1, n_2\rangle_\parallel \equiv |n_1, n_2\rangle_\parallel
$$
$$
\equiv \frac{1}{\sqrt{\Lambda_n}} \sum_{\nu_1=0}^{2} \cdots \sum_{\nu_N=0}^{2} \delta\left(\sum_k \nu_k - n_1 - 2n_2\right)
$$
$$
\bigotimes_{k=1}^{N} |\nu_k\rangle, \tag{30}
$$

where, in analogy with the projection method devised in the qubit case [18], n_1 and n_2 are the numbers of $|1\rangle$ and $|2\rangle$ states in the N-tensor-product state, while Λ_n is the multinomial coefficient $\binom{N}{n_1, n_2}$.

Again, we take as inputs, on the one hand, a tensor-product state and, on the other hand, a uniform superposition of the basis elements $|n_1, n_2\rangle_\parallel$. Explicitly, the former reads

$$
|\Psi_0\rangle^{(1)} = \frac{1}{\sqrt{3^N}} \sum_{n_1=0}^{N} \sum_{n_2=0}^{N-n_1} \sqrt{\binom{N}{n_1, n_2}} |n_1, n_2\rangle_\parallel \tag{31}
$$

and the latter

$$
|\Psi_0\rangle^{(2)} = \frac{1}{\sqrt{M}} \sum_{n_1=0}^{N} \sum_{n_2=0}^{N-n_1} |n_1, n_2\rangle_\parallel, \tag{32}
$$

where $\binom{N}{n_1, n_2}$ is again the multinomial coefficient and $M = \sum_{n_1=0}^{N} \sum_{n_2=0}^{N-n_1} 1 = (N + 1)(N/2 + 1)$. The entanglement of the Holland-Burnett qutrit states $|\Psi_0\rangle^{(2)}$ is discussed in Appendix B. The unitary encoding ϕ and θ can be represented in the form $U_{\phi,\theta} = e^{i(\phi H_1 + \theta H_2)}$, with generators $H_1 = |1\rangle\langle 1|$ and $H_2 = |2\rangle\langle 2|$. Given this representation, one applies $U_{\phi,\theta}^{\otimes N}$ to the probe state.

As for the qubit case, we show that the N copies of the equatorial qutrit state $|\Psi_0\rangle^{(1)}$ used as probes for our protocol achieve the standard quantum limits, whereas the Holland-Burnett state $|\Psi_0\rangle^{(2)}$ provides a precision asymptotic to the Heisenberg bound. In other words, our protocol run with the former probe generalizes the single-phase optimal parallel-separable strategy, while, if it is started with the latter probe, it generalizes the single-phase optimal sequential strategy. As for the qubit case, this shows that, to achieve the Heisenberg bound, we need the N probes to be entangled. In this simple example, the generalization of the single-phase case is limited to two phases, but the protocol is viable for the application to any number of phases. We will define the ultimate limits

attainable if a generic number k of phases is jointly estimated with N resources in the next section.

2. Measurement

Holevo's POVM in the case of two estimators $\tilde{\phi}$ and $\tilde{\theta}$ is a projection over Susskind-Glogower vectors developed on the elements of the nondegenerate qutrit Hilbert space, i.e.,

$$d\mu_\parallel(\tilde{\phi}, \tilde{\theta}) = \frac{d\tilde{\phi}}{2\pi}\frac{d\tilde{\theta}}{2\pi}|E(\tilde{\phi}, \tilde{\theta})\rangle\langle E(\tilde{\phi}, \tilde{\theta})|, \qquad (33)$$

with

$$|E(\tilde{\phi}, \tilde{\theta})\rangle \equiv \sum_{n_1=0}^{N}\sum_{n_2=0}^{N-n_1} e^{i(n_1\tilde{\phi}+n_2\tilde{\theta})}|n_1, n_2\rangle. \qquad (34)$$

The discretization procedure works similarly to the single-parameter case, through the prescription

$$p(m_\phi, m_\theta|, \phi, \theta)$$
$$= \frac{1}{(N+1)^2}\Pr\left(\tilde{\phi} = \frac{2\pi m_\phi}{N+1}, \tilde{\theta} = \frac{2\pi m_\theta}{N+1}\Big|\theta, \phi\right). \qquad (35)$$

B. Results

1. Performance

The input states $|\Psi_0\rangle^{(1)}$ and $|\Psi_0\rangle^{(2)}$ in our protocol provide, respectively, the conditional densities

$$p^{(1)}(m_\phi, m_\theta|\phi, \theta) = \frac{1}{(N+1)^2}\left|\frac{1}{\sqrt{3^N}}\sum_{n_1=0}^{N}\sum_{n_2=0}^{N-n_1}\sqrt{\binom{N}{n_1, n_2}}\exp\left\{2\pi i\left[n_1\left(\frac{m_\phi}{N+1}-\phi\right)+n_2\left(\frac{m_\theta}{N+1}-\theta\right)\right]\right\}\right|^2 \qquad (36)$$

and

$$p^{(2)}(m_\phi, m_\theta|\phi, \theta) = \frac{1}{(N+1)^3(N/2+1)}\left|\sum_{n_1=0}^{N}\sum_{n_2=0}^{N-n_1}\exp\left\{2\pi i\left[n_1\left(\frac{m_\phi}{N+1}-\phi\right)+n_2\left(\frac{m_\theta}{N+1}-\theta\right)\right]\right\}\right|^2. \qquad (37)$$

The distribution $p^{(1)}$ can be again approximated within a single period by a Gaussian for $N \gg 1$, i.e.,

$$p^{(1)}(m_\phi, m_\theta|\phi, \theta) \xrightarrow[N\gg 1]{} \frac{8\sqrt{3}}{9}\pi\frac{N}{(N+1)^2}\exp\left\{-\frac{16}{9}N\pi^2\left[\left(\frac{m_\phi}{N+1}-\phi\right)^2+\left(\frac{m_\theta}{N+1}-\theta\right)^2-\left(\frac{m_\phi}{N+1}-\phi\right)\left(\frac{m_\theta}{N+1}-\theta\right)\right]\right\}. \qquad (38)$$

The sum defining distribution $p^{(2)}$ can be analytically calculated and provides

$$p^{(2)}(m_\phi, m_\theta|\phi, \theta)$$
$$= \frac{1}{M}\big[S_{\Delta\phi}^2 + S_{\Delta\theta}^2 + S_{\Delta\phi-\Delta\theta}^2 - 2S_{\Delta\phi}S_{\Delta\theta}C_{(3+2N)(\Delta\phi-\Delta\theta)}$$
$$+ 2S_{\Delta\phi-\Delta\theta}(S_{\Delta\theta}S_{(3+2N)\Delta\phi} - S_{\Delta\phi}C_{(3+2N)\Delta\theta})\big], \qquad (39)$$

where

$$M = 8(N+2)(N+1)^3 S_{\Delta\phi}^2 S_{\Delta\theta}^2 S_{\Delta\phi-\Delta\theta}^2, \qquad (40)$$

$$S_x \equiv \sin(\pi x), \qquad (41)$$

$$C_x \equiv \cos(\pi x), \qquad (42)$$

$$\Delta\alpha \equiv \alpha - \frac{m_\alpha}{N+1}, \quad \alpha = \phi, \theta. \qquad (43)$$

Note that the distributions above are obtained by directly generalizing the qubit probe states employed for the single-phase estimation to the qutrit case suited for double-phase estimation. Therefore, we will refer to the estimation strategies related to $p^{(1)}$ and $p^{(2)}$ as the double-phase optimal parallel-separable and sequential strategies, respectively.

We start by comparing the mutual information $I(m_\phi, m_\theta : \phi, \theta)$ provided by $p^{(1)}$ and $p^{(2)}$ for small N (<20). The numeric evaluation shown in Fig. 2 reveals that we find again a trade-off between the protocol started with $|\Psi^{(1)}\rangle$ and the one

started with $|\Psi^{(2)}\rangle$. In particular, the first slightly outperforms the second up to a threshold on N, which in the double-phase case is nearly twice the single-phase threshold.

If we focus on the asymptotics, we find that our protocol run with the tensor-product state $|\Psi^{(1)}\rangle$ is limited by $2I_{\text{SQL}}(1, N)$, whereas if we start it with the Holland-Burnett state $|\Psi^{(2)}\rangle$ we achieve $2I_{\text{HB}}(1, N)$. We recall that $I_{\text{SQL}}(1, N)$ and $I_{\text{HB}}(1, N)$ are, respectively, the standard quantum limit
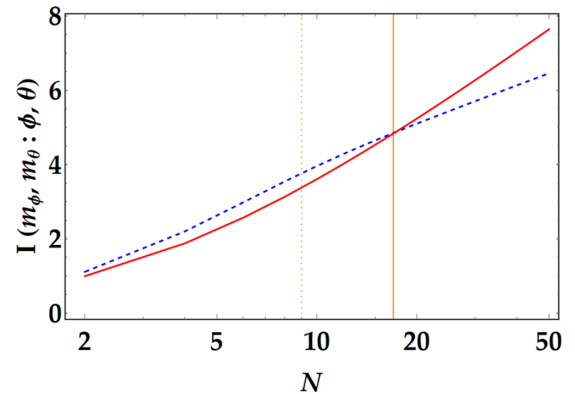

FIG. 2. Base-e log-linear plot of the mutual information in the double-parameter estimation case. The blue dashed line shows the double-phase optimal parallel-separable strategy and the red line the double-phase optimal sequential strategy.

and the Heisenberg bound in the single-phase estimation case. Explicitly,

$$I^{(1)}(m_\phi, m_\theta : \phi, \theta) \xrightarrow[N \gg 1]{} 2I_{\text{SQL}}(1, N) + \log_2 \left(\frac{8\sqrt{3}\pi}{9} \right)$$

$$- (\ln 2)^{-1} \sim 2I_{\text{SQL}}(1, N) + 0.8314 \tag{44}$$

and

$$I^{(2)}(m_\phi, m_\theta : \phi, \theta) \xrightarrow[N \gg 1]{} 2I_{\text{HB}}(1, N) - 3.7899, \tag{45}$$

where we derived the limit of $I^{(1)}$ by means of an analytic approximation while the limit on $I^{(2)}$ is a numeric result obtained by taking $N = 500$.

2. Comparison with single-phase estimation

Now we investigate if we can find any advantage in estimating two phases simultaneously rather than performing two independent single-phase estimations. For the sake of clarity, herein we use for the mutual information between k parameters and estimators the shorthand notation $I_k \equiv I(\boldsymbol{m} : \boldsymbol{\phi})$.

We have to compare double-phase estimation with two single-phase estimations sharing the same total number of encoding operations N to see if there is an advantage in terms of mutual information in estimating two phases simultaneously by encoding them on the same state or in estimating them one by one. Note that if the phases are estimated independently through two distinct single-phase estimation protocols, the mutual information between estimators and phases is simply the sum of the mutual information between each estimator and the corresponding phase, i.e.,

$$I(m_\phi, m_\theta : \phi, \theta) = I(m_\phi : \phi) + I(m_\theta : \theta) = 2I(m_\phi : \phi), \tag{46}$$

where the second equality holds only if the same number of states $N/2$ is employed for the estimation. Shortly, using our new notation, we are comparing $I_2(N)$ with $2I_1(N/2)$.

Figures 3 and 4 show that for small N the simultaneous encoding on qutrits provides a slight advantage over independent single-phase estimations, even if the latter is supported with one, for $N > 3$, or two, for $N > 6$, more resources for each phase. This result holds for both generalized optimal parallel-separable and sequential strategies. The asymptotics shown above prove that exploiting a large number of resources does not improve much the performance of double-phase estimation with respect to two separate single-phase estimations. Indeed, we remark that the advantage does not scale with N and approaches a constant factor. Specifically, comparing the asymptotics in Eqs. (28) and (29) with the ones in Eqs. (44) and (45), we note that the difference in terms of mutual information amounts to approximately one bit.

One may still wonder if this advantage, which is small in the case of double-phase estimation, depends at least on the number of simultaneously estimated phases. This is the problem of multiphase estimation that we mentioned in the Introduction. In the next section we show that the ultimate bound on the achievable precision in terms of mutual information does depend on the number of jointly estimated phases, but the actual gain with respect to repeated single-phase estimation still amounts to a constant factor at most.
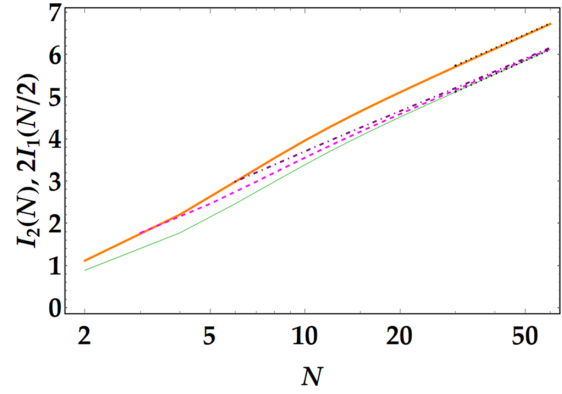

FIG. 3. Double-phase optimal parallel-separable strategy. The base-e log-linear plot of the mutual information in the double-phase estimation case $I_2(N)$ (orange thick solid line) is compared with the mutual information of two independent single-phase estimations $2I_1(N/2)$ where the same number of resources for each phase is employed (green thin solid line). The black dotted lines display the asymptotic behavior of the two cases. We also show $I_1[(N + 1)/2]$ (magenta dashed line) and $I_1[(N + 2)/2]$ (purple dash-dotted line) and note that, for $N > 3$ and $N > 6$, respectively, the generalized parallel-separable strategy with double-phase estimation can outperform the corresponding single-parameter estimation even if the latter is supported with one or two probes more for each phase.

VI. DIGITAL HEISENBERG BOUND ON MULTIPHASE ESTIMATION

Finally, we derive the Heisenberg bound on the mutual information $I_k(N)$ between estimators and phases in the case where N resources are employed and k phases simultaneously estimated. Then we find the number of bits shared between phases and estimators when the k phases are
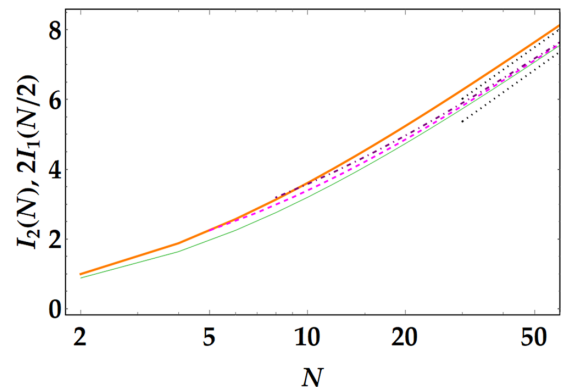

FIG. 4. Double-phase optimal sequential strategy. The base-e log-linear plot of the mutual information in the double-parameter estimation case $I_2(N)$ (orange thick solid line) is compared with the mutual information of two independent single-parameter estimations $2I_1(N/2)$ where the same number of probes per parameter were employed (green thin solid line). The black dotted lines display the asymptotic behavior of the two cases. We also show $I_1[(N + 1)/2]$ (magenta dashed line) and $I_1[(N + 2)/2]$ (purple dash-dotted line) and note that, for small N, the generalized sequential strategy with double-parameter estimation can outperform the corresponding single-parameter estimation even if the latter is supported with one or two probes more.

estimated separately, again for a fixed number of resources. Hence, we discuss if there actually is an advantage by exploiting multiphase estimation with respect to the straightforward application of a single-phase estimation algorithm for each phase.

We start by stating a proposition that generalizes for a generic number of simultaneously estimated phases the Heisenberg bound on the mutual information which was derived in Ref. [21] in the case of single-phase estimation. The result is a generalization of the Holevo bound.

Proposition 1. Given a set of k phases $\boldsymbol{\phi}$ to be estimated with N resources, the ultimate bound on the mutual information $I_k(N)$ between the phases and the corresponding set of estimators $\tilde{\boldsymbol{\phi}}$ reads

$$I_k(N) \leqslant \log_2 \binom{N+k}{N} \equiv I_{\text{HB}}(k, N). \tag{47}$$

The proof can be found in Appendix C. As mentioned above, $I_{\text{HB}}(k, N)$ identifies the Heisenberg limit in terms of mutual information for a generic number k of phases jointly estimated with N resources.

From the Heisenberg bound in Eq. (47) we find that the performance of a multiphase estimation protocol in this context strongly depends on the choice of the parameters N and k. Depending on the ratio N/k, we can devise three specific asymptotic scenarios. We can evaluate the largest mutual information attainable for each jointly estimated phase $I_{\text{HB}}(k, N)/k$ for each one.

First, if we fix the number of phases and take $N/k \gg 1$, we immediately find for each jointly estimated phase the Heisenberg bound of the single-phase estimation, namely,

$$I_{\text{HB}}(k, N)/k \xrightarrow[N \gg k]{} \log_2 N - \frac{1}{k} \log_2 k! \sim I_{\text{HB}}(1, N). \tag{48}$$

Second, we take $N/k \sim 1$. We observe that this case is quite a typical choice since one may want to use every encoding operation for a different phase and take $k = N$. In such a case, we can also explore the asymptotic limit of the Heisenberg bound through Stirling's approximation and find

$$I_{\text{HB}}(k, N)/k \sim \frac{1}{N}[\log_2 (2N)! - 2 \log_2 N!] \xrightarrow[N \gg 1]{} 2. \tag{49}$$

Finally, we can fix N and take the limit $N/k \ll 1$. We get

$$I_{\text{HB}}(k, N)/k \xrightarrow[k/N \gg 1]{} \frac{N}{k} \log_2 \left[e \left(1 + \frac{k}{N} \right) \right]. \tag{50}$$

The collapse of the ratio between the upper bound on the mutual information and the number of phases in Eq. (50) is expected since in this case the number of phases is larger than the number of encoding operations, so there are $k - N$ phases that cannot be distinguished from the others.

These three different regimes are displayed in Fig. 5. We conclude that an efficient multiphase estimation needs to set the number of resources much larger than the number of simultaneously estimated phases. In this optimal regime, we compare the ultimate bound for multiphase estimation of k phases with N resources in Eq. (47) with the expected optimal performance of k independent single-phase estimation with N/k resources each. The mutual information in this latter case
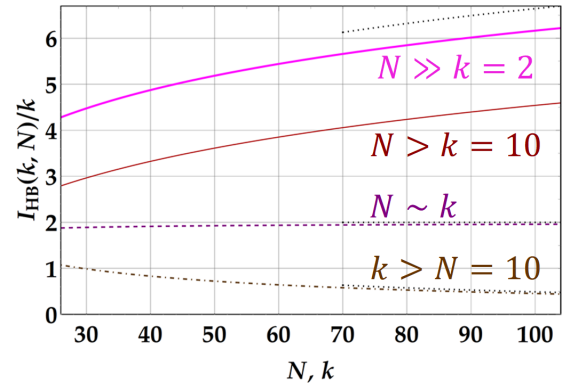

FIG. 5. Heisenberg bound in Eq. (47) for the simultaneous estimation of k phases with N resources. The bound displays three different scalings according with the ratio N/k. The first case is $N/k \gg 1$, shown by thick magenta ($k = 2$) and thin red ($k = 10$) solid lines. The asymptotic behavior of the bound is expressed by Eq. (48). The corresponding black dotted line displays the ultimate limit $I_{\text{HB}}(1, N)$. The second case is $N/k \sim 1$, shown by a purple dashed line. The asymptotic behavior of the bound is expressed by Eq. (49), i.e., it approaches a plateau emphasized by the black dotted line. The third case is $N/k < 1$, shown by the brown dash-dotted line ($N = 10$). The asymptotic behavior of the bound is expressed by Eq. (50).

generalizes the one shown in Eq. (46) and the ultimate bound on that reads

$$k I_1(N/k) \leqslant k(\log_2 N - \log_2 k). \tag{51}$$

We readily find that the multiphase advantage $\Delta I_k \equiv I_{\text{HB}}(k, N) - k I_{\text{HB}}(1, N/k)$ depends on the number of simultaneously encoded phases only, namely,

$$\Delta I_k = k \log_2 k - \log_2(k!), \tag{52}$$

implying that the information gain for each phase reads $\Delta I_k/k = \log_2 k - \log_2(k!)/k$. Asymptotically, through Stirling's approximation, we find that the gain approaches a small constant, i.e.,

$$\Delta I_k/k \xrightarrow[k \gg 1]{} \log_2(e). \tag{53}$$

Therefore, we can conclude that we do get an advantage from multiphase estimation, but asymptotically it is reduced to a small constant gain, as found also in Ref. [20].

VII. CONCLUSION

Exploiting the generalization of Holevo's POVM to the multiphase scenario devised in Ref. [18], we developed a protocol for global multiphase estimation. The protocol generalizes the analysis in Ref. [18] to arbitrary cost functions by optimizing over the set of probe states. In particular, differently from Ref. [18], here we considered the case of discrete estimators and took the mutual information between phases and estimators as a figure of merit, thus extending the single-phase digital estimation in Ref. [21]. Remarkably, the protocol performs a global estimation; therefore it can be used to test new estimation strategies in terms of the desired cost function without any *a priori* notion of the phase to be estimated.

We showed that in the single-phase case we retrieve known optimal estimation strategies. Specifically, optimal parallel-separable and sequential (QPEA) strategies were investigated, also in the nonasymptotic case.

More importantly, our protocol provides a multiphase estimation, without assuming any specific encoding strategy. We analyzed here the most basic nontrivial case, i.e., the case of two unknown phases, and pointed out the performance of generalized sequential and parallel-separable strategies in terms of mutual information.

We also retrieved the ultimate bound on precision in terms of mutual information when a generic number of phases is jointly estimated given a fixed number of resources. Our bound contributes to clarify the controversy about the advantage of multiphase estimation [6–8,19,20].

APPENDIX A: DISCRETIZATION

1. Conditions for the invariance of cost functions under discretization

In parameter estimation, the mutual information is the number of bits in common between the parameter to be estimated and a suitable estimator. In many applications, the parameter is a continuous quantity but the estimator is a map from the parameter to a discrete set, which intrinsically limits the precision of the estimation. This is a realistic occurrence: Achieving the required number of decimal digits for the estimation of a parameter is one of the most important issues in metrology. Discrete phase estimation applied to figures of merit different from the mutual information was considered in Refs. [14,35].

In this context, the discretization is determined by the number of qudits N that we can exploit for our estimation. The estimator is then defined by the partition in N subsets of the parameter set $[0, 2\pi)$, i.e., $\tilde{\phi} = 2\pi m/(N+1)$ with $m \in [0, N]$.

First, we need to map the continuous probability distributions $\text{Pr}(\tilde{\phi}|\phi)$ into discrete quantities $p(m|\phi)$. This can be done by taking the limit

$$\lim_{N \to \infty} \text{Pr}\left(\frac{2\pi m}{N+1}\bigg|\phi\right) = \text{Pr}(\tilde{\phi}|\phi) \equiv \alpha \lim_{N \to \infty} p(m|\phi), \quad \text{(A1)}$$

where α is a normalization constant. Indeed, we also need the discrete quantities $p(m|\phi)$ to be probabilities. Then, for every fixed phase $\phi = \bar{\phi}$ we require $0 \leqslant p(m|\bar{\phi}) \leqslant 1 \, \forall \, m \in [0, N]$ and $\sum_{m=0}^{N} p(m|\bar{\phi}) = 1$. The normalization over the discrete

set of the outcomes implies

$$1 = \int_0^{2\pi} \text{Pr}(\tilde{\phi}|\bar{\phi})\frac{d\tilde{\phi}}{2\pi} = \lim_{N \to \infty} \frac{1}{N+1} \sum_{m=0}^{N} \text{Pr}\left(\frac{2\pi m}{N+1}\bigg|\bar{\phi}\right)$$

$$= \lim_{N \to \infty} \frac{\alpha}{N+1} \sum_{m=0}^{N} p(m|\bar{\phi}),$$

(A2)

implying $\alpha = N + 1$ and the prescription

$$p(m|\phi) = \frac{1}{N+1}\text{Pr}\left(\tilde{\phi} = \frac{2\pi m}{N+1}\bigg|\phi\right). \quad \text{(A3)}$$

The same distribution can be derived by equivalently discretizing Holevo's POVM $d\mu(\tilde{\phi}) \to \mu(m)$ through the same procedure, i.e.,

$$d\mu(\tilde{\phi}) = \frac{d\tilde{\phi}}{2\pi}|e(\tilde{\phi})\rangle\langle e(\tilde{\phi})| = \lim_{N \to \infty} \frac{1}{N+1}|e(m)\rangle\langle e(m)|$$

$$\equiv \lim_{N \to \infty} \mu(m). \quad \text{(A4)}$$

Note that in general the distribution in Eq. (A3) is normalized just in the limit of large N. However, one can check that for the distributions $p^{(1)}$ and $p^{(2)}$ we consider here $\sum_{m=0}^{N} p^{(1)}(m|\phi) = \sum_{m=0}^{N} p^{(1)}(m|\phi) \sim 1 \, \forall N$ because the width of their peaks is limited to the interval $m \in (\phi(N+1) - 1, \phi(N+1) + 1)$ and is the effective support of the distributions, meaning that $p^{(x)}(m \neq \phi(N+1)) = o((N+1)^{-2})$ while $p^{(x)}(m = \phi(N+1)) \sim 1$, with $x = 1, 2$.

However, we still need to check how discretization affects the cost functional in Eq. (2), which is now expressed as

$$\bar{C} = \int_0^{2\pi} \frac{d\phi}{2\pi} \int_0^{2\pi} \frac{d\tilde{\phi}}{2\pi}\text{Pr}(\tilde{\phi}|\phi)c(\phi, \tilde{\phi})$$

$$= \lim_{N \to \infty} \frac{1}{N+1} \int_0^{2\pi} \frac{d\phi}{2\pi} \sum_{m=0}^{N} \text{Pr}(m|\phi)c(\phi, m)$$

$$= \lim_{N \to \infty} \int_0^{2\pi} \frac{d\phi}{2\pi} \sum_{m=0}^{N} p(m|\phi)c(\phi, m). \quad \text{(A5)}$$

The form of the cost function $c(\phi, \tilde{\phi})$ is crucial to the difference between the continuous and discrete cases. Through the following observations, we will define the cases where the discretization works for every N.

Proposition 2. Consider a two-variable function $f(\phi, \tilde{\phi}) \equiv \text{Pr}(\tilde{\phi}|\phi)c(\phi, \tilde{\phi})$, defined as the product of a conditional density and a generic function c. Let the two variables share a continuous domain. If (a) f depends only on the difference between the two variables, i.e., $f(\phi, \tilde{\phi}) = f(\phi - \tilde{\phi})$, (b) f is periodic, (c) $c(\phi, \tilde{\phi})$ does not depend on $\text{Pr}(\tilde{\phi}|\phi)$, and (d) the discrete distribution implied by Eq. (A3) is normalized for all N, then

$$\int_0^{2\pi} \frac{d\tilde{\phi}}{2\pi} \int_0^{2\pi} \frac{d\phi}{2\pi}f(\phi - \tilde{\phi})$$

$$= \frac{1}{N+1} \int_0^{2\pi} \frac{d\phi}{2\pi} \sum_{m=0}^{N} f\left(\phi - \frac{2\pi m}{N+1}\right) \forall N. \quad \text{(A6)}$$

This result in our case implies

$$\bar{C} = \int_0^{2\pi} \frac{d\phi}{2\pi} \int_0^{2\pi} \frac{d\tilde{\phi}}{2\pi} \Pr(\tilde{\phi}|\phi) c(\phi, \tilde{\phi})$$

$$= \int_0^{2\pi} \frac{d\phi}{2\pi} \sum_{m=0}^{N} p(m|\phi) c(\phi, m),$$

(A7)

i.e., we can ignore the limit $N \to \infty$ in Eq. (A5).

Proof. Let $\tilde{\phi}$ be a continuous variable, periodic with period 2π. Then

$$\int_0^{2\pi} \frac{d\tilde{\phi}}{2\pi} \int_0^{2\pi} \frac{d\phi}{2\pi} f(\phi - \tilde{\phi}) = \int_0^{2\pi} \frac{d\gamma}{2\pi} \int_0^{2\pi} \frac{d\phi}{2\pi} f(\gamma)$$

$$= \int_0^{2\pi} \frac{d\gamma}{2\pi} f(\gamma).$$

(A8)

On the other hand, if $\tilde{\phi}$ is discrete and, in particular, $\tilde{\phi} = 2\pi m/(N+1)$, we take the sum

$$\sum_{m=0}^{N} \int_0^{2\pi} \frac{d\phi}{2\pi} f\left(\phi - \frac{2\pi m}{N+1}\right)$$

$$= \sum_{m=0}^{N} \int_{-2\pi m/(N+1)}^{2\pi[1-m/(N+1)]} \frac{d\gamma}{2\pi} f(\gamma)$$

$$= (N+1) \int_0^{2\pi} \frac{d\gamma}{2\pi} f(\gamma)$$

$$= (N+1) \int_0^{2\pi} \frac{d\tilde{\phi}}{2\pi} \int_0^{2\pi} \frac{d\phi}{2\pi} f(\phi - \tilde{\phi}),$$

(A9)

where we use the periodicity of f in the second equality and the result of Eq. (A8) in the third one. The third and fourth conditions guarantee that the ratio over $N+1$ normalizes the resulting discrete probability according to the prescription in Eq. (A3). ∎

Cost functions belonging to the Holevo class, like $c(\phi, \tilde{\phi}) = 4\sin(\phi - \tilde{\phi})^2/2$, all satisfy the hypotheses of Proposition 2. In such a case, for any number of qubits N employed for estimation, continuous and discrete estimators are equivalent to the cost function.

2. Effects on the mutual information

Here we are interested in the mutual information, whose related cost functional is the conditional entropy $H(\tilde{\phi}|\phi) = -\Pr(\tilde{\phi}|\phi) \log_2 \Pr(\tilde{\phi}|\phi)$. Note that, as long as $\Pr(\tilde{\phi}|\phi)$ is a periodic function, this functional satisfies the second of the conditions required in Proposition 2. However, the cost function related to $H(\tilde{\phi}|\phi)$ is the so-called surprise $c(\phi, \tilde{\phi}) = -\log_2 \Pr(\tilde{\phi}|\phi)$, which depends on the conditional density $\Pr(\tilde{\phi}|\phi)$, against the third condition, implying that the discretized distribution defining the surprise is not normalized. Indeed, continuous and discrete estimators provide different conditional entropies. Notwithstanding, the mutual information is invariant under discretization, as we show in the following.

Proposition 3. Given a periodic and covariant conditional density $P(\tilde{\phi}|\phi)$, the mutual information between an estimator $\tilde{\phi}$ and a real parameter $\phi \in [0, 1]$, defined as

$I(\tilde{\phi} : \phi) \equiv H(\tilde{\phi}) - H(\tilde{\phi}|\phi)$, is invariant under discretization of the estimator.

Proof. Let $\tilde{\phi}$ be a continuous variable. Then the conditional entropy $H(\tilde{\phi}|\phi)$ reads

$$H(\tilde{\phi}|\phi) = -\int_0^1 d\phi \int_0^1 d\tilde{\phi} \Pr(\tilde{\phi}|\phi) \log_2 \Pr(\tilde{\phi}|\phi). \quad (A10)$$

We assume to have no prior information on the estimator. Thus the related probability distribution is uniform and the differential entropy $H(\tilde{\phi})$ depends on the integration domain. Here we have

$$H(\tilde{\phi}) = -\int_0^1 \Pr(\tilde{\phi}) \log_2 \Pr(\tilde{\phi}) = 0. \quad (A11)$$

Therefore, we find for the mutual information

$$I(\tilde{\phi} : \phi) = -H(\tilde{\phi}|\phi) = \int_0^1 d\phi \int_0^1 d\tilde{\phi} \Pr(\tilde{\phi}|\phi) \log_2 \Pr(\tilde{\phi}|\phi)$$

$$= \int_0^1 d\gamma \Pr(\gamma) \log_2 \Pr(\gamma),$$

(A12)

where in the last equality we exploit covariance. Now take a discrete estimator $\tilde{\phi} = m/(N+1)$. Without assuming any convergence to Eq. (A10), we write the conditional entropy as

$$H(m|\phi) = -\sum_{m=0}^{N} \int_0^1 d\phi \, p(m|\phi) \log_2 p(m|\phi)$$

$$= -\sum_{m=0}^{N} \int_0^1 d\phi \frac{\Pr(m|\phi)}{N+1} \log_2 \frac{\Pr(m|\phi)}{N+1}$$

$$= -\sum_{m=0}^{N} \int_{-m/(N+1)}^{1-m/(N+1)} d\gamma \frac{P(\gamma)}{N+1} \log_2 \frac{P(\gamma)}{N+1}$$

$$= \log_2(N+1) - \int_0^1 d\gamma \Pr(\gamma) \log_2 \Pr(\gamma).$$

(A13)

On the other hand, the assumption of a uniform prior distribution for the estimator in the discrete case provides

$$H(m) = -\sum_{m=0}^{N} p(m) \log_2 p(m) = \log_2(N+1), \quad (A14)$$

implying

$$I(m : \phi) = H(m) - H(m|\phi)$$

$$= \int_0^1 d\gamma \Pr(\gamma) \log_2 \Pr(\gamma) = I(\tilde{\phi} : \phi) \, \forall N,$$

(A15)

which completes the proof. ∎

APPENDIX B: ENTANGLEMENT OF THE LOCALLY OPTIMAL INPUT STATE

In this Appendix we discuss the entanglement content of the states introduced in the main text. In particular, we analytically evaluate the N-partite entanglement among the N

probes of the states that are locally optimal for the mutual information, i.e., the Holland-Burnett states. We show that, as expected, the qutrit states are more entangled than the qubit ones. We use as a multipartite entanglement quantifier the geometric measure of entanglement [36], which can be easily computed in this case due to the symmetry of the states. The geometric measure of entanglement of a pure state $|\psi\rangle$ is

$$E_G(|\psi\rangle) = 1 - \max_{|\Phi\rangle \in \text{prod}} |\langle\Phi|\psi\rangle|^2, \tag{B1}$$

where we denote by prod the set of pure product states. The geometric measure of the state $|\Psi_0\rangle^{(1)}$ in Eq. (19) is zero since the state is separable. The state $|\Psi_0\rangle^{(2)}$ in Eq. (20) is pure and symmetric and hence the closest pure product state is symmetric [37]. We have [38]

$$E_G(|\Psi_0\rangle^{(2)}) \simeq 1 - \frac{\sqrt{2\pi N}}{N+1}. \tag{B2}$$

In the qutrit case, we have have similarly that the geometric measure of the state $|\Psi_0\rangle^{(1)}$ in Eq. (31) is zero since the state is separable. We now calculate the geometric measure of entanglement of the state $|\Psi_0\rangle^{(2)}$ of Eq. (32). We first note that we can write the closest symmetric product state as $|\Phi\rangle = (\sqrt{p}|0\rangle + \sqrt{q}e^{i\phi}|1\rangle + \sqrt{1-p-q}e^{i\theta}|2\rangle)^{\otimes n}$. We then obtain

$$E_G(|\Psi_0\rangle^{(2)}) = 1 - \frac{1}{M}\left|\sum_{n_1=0}^{N}\sum_{n_2=0}^{N-n_1}\sqrt{M(N, n_1, n_2)}\sqrt{p}^{N-n_1-n_2}\right.$$
$$\left.\times \sqrt{q}^{n_1}\sqrt{1-p-q}^{n_2}e^{-in_1\phi}e^{-in_1\theta}\right|^2. \tag{B3}$$

It is easy to see that this expression is minimized for $\phi = \theta = 0$ and $p = q = 1/3$. We can approximate the sum by an integral and the multinomial distribution by a two-dimensional Gaussian with mean $\vec{\mu} = N(1/3, 1/3)$ and variance $\Sigma = N((2/9, -1/9), (-1/9, 2/9))$. Straightforwardly, we get

$$E_G(|\Psi_0\rangle^{(2)}) \simeq 1 - \frac{8\pi N}{3\sqrt{3}M}. \tag{B4}$$

The above result implies that the qutrit states are more entangled than the qubit ones. Note that the relative entropy of entanglement is divergent in both cases since $E(\rho) \geqslant -\log_2[1 - E_G(\rho)]$ (see Fig. 6).

APPENDIX C: HEISENBERG BOUND ON THE MUTUAL INFORMATION FOR MULTIPHASE ESTIMATION

Here we provide the proof of Proposition 1. We follow the approach given in the Supplemental Material of Ref. [21].

Proof. Let $\boldsymbol{\phi} = (\phi_1, \phi_2, \dots, \phi_k)$ be the vector of phases we wish to estimate. We have the unitary representation $U_{\boldsymbol{\phi}} \equiv \exp(i\sum_{j=1}^{k} H_j\phi_j)$, with $\{H_j\}_{j=1}^{k}$ a set of commuting generators $H_j = |j\rangle\langle j|$. Take a generic probe state $\rho_0 \in \mathcal{S}(\mathcal{H}_{\parallel})$, where $\mathcal{S}(\cdot)$ is a set of states, here defined on the Hilbert space $\mathcal{H}_{\parallel}$ spanned by the nondegenerate vectors $|\boldsymbol{n}\rangle \equiv \{|n_j\rangle_{\parallel}\}_{j=0}^{k}$ with the constraint on the sum of the corresponding

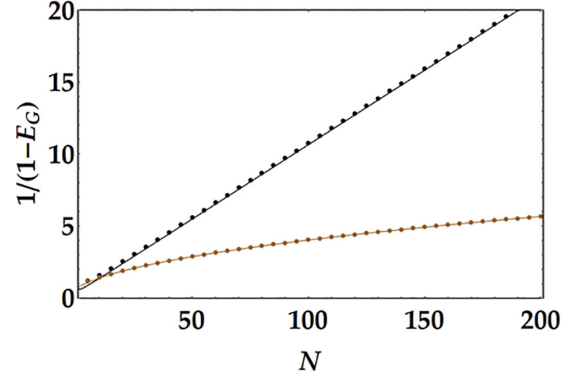

FIG. 6. Entanglement of the state $|\Psi_0\rangle^{(2)}$ in the qubit (brown) and qutrit (black) cases as a function of N. The black and brown solid lines are the asymptotic behaviors in Eqs. (B2) and (B4), respectively, while the dots displays the corresponding exact values. For the qubit case the inverse of the fidelity of separability, i.e., $1/(1 - E_G)$, scales as $\sqrt{N}$. In the qutrit case it is linear in N.

eigenvalues given by $\sum_{j=0}^{k} n_j = N$, i.e., $n_0 = N - \sum_{j=1}^{k} n_j$. Then ρ_0 can be expressed as

$$\rho_0 = \sum_{\boldsymbol{n},\boldsymbol{m}} \lambda_{\boldsymbol{n},\boldsymbol{m}}|\boldsymbol{n}\rangle\langle\boldsymbol{m}|, \tag{C1}$$

with $\sum_{\boldsymbol{n}} \equiv \sum_{n_1=0}^{N}\sum_{n_2=0}^{N-n_1}\cdots\sum_{n_k=0}^{N-\sum_{j=1}^{k-1}n_j}$ and $\sum_{\boldsymbol{n}} \lambda_{\boldsymbol{n},\boldsymbol{n}} = 1$. Note that we set the number of vectors $|n_j\rangle_{\parallel}$ to $k+1$ in order to exploit the whole basis of the nondegenerate subspace, except for a reference state, to encode the phases to be estimated.

The application of the unitary on the probe state yields

$$\rho_{\boldsymbol{\phi}} = U_{\boldsymbol{\phi}}^{\otimes N}\rho_0(U_{\boldsymbol{\phi}}^{\dagger})^{\otimes N}. \tag{C2}$$

Here $p_{\boldsymbol{\phi}} = p(\phi_1, \phi_2, \dots, \phi_k)$ is the corresponding joint probability distribution. Finally, let $S(\rho)$ be the von Neumann entropy related to state ρ. Then the maximum information on $\boldsymbol{\phi}$ extractable through a corresponding set of estimators $\tilde{\boldsymbol{\phi}}$ from $\rho_{\boldsymbol{\phi}}$ is given by the Holevo bound, i.e.,

$$I(\tilde{\boldsymbol{\phi}} : \boldsymbol{\phi}) \leqslant S\left(\int d\boldsymbol{\phi}\, p_{\boldsymbol{\phi}}\rho_{\boldsymbol{\phi}}\right) - \int d\boldsymbol{\phi}\, p_{\boldsymbol{\phi}}S(\rho_{\boldsymbol{\phi}}). \tag{C3}$$

We remark that the bound is maximized for pure states, since, for pure ρ, we have $S(\rho) = 0$. Therefore, henceforth we take the probe in a pure state. In this case, the Holevo bound reads $I(\tilde{\boldsymbol{\phi}} : \boldsymbol{\phi}) \leqslant S(\int d\boldsymbol{\phi}\, p_{\boldsymbol{\phi}}\rho_{\boldsymbol{\phi}})$. Note that the mutual information $I(\tilde{\boldsymbol{\phi}} : \boldsymbol{\phi})$ is the same one that in the main text we denoted by $I_k(N)$ to emphasize the main parameters, i.e., the number of phases to be estimated k and the number of resources N.

Now consider the projectors over the eigenbases of the generators defined as $P_n^{(k)} \equiv |n\rangle\langle n|$. Since the sum over the eigenvalues $\{n_j\}_{j=0}^{k}$ is constrained to be equal to the number of resources N, the number of projectors $P_n^{(k)}$ is

$$M \equiv \sum_{n_1=0}^{N}\sum_{n_2=0}^{N-n_1}\cdots\sum_{n_k=0}^{N-\sum_{j=1}^{k-1}n_j} 1 = \binom{N+k}{N}, \tag{C4}$$

which is the dimension of $\mathcal{H}_{\parallel}$ [39].

The entropy bounding the mutual information can be bounded through the data-processing inequality for projectors as

$$S\left(\int d\boldsymbol{\phi}\, p_{\boldsymbol{\phi}}\rho_{\boldsymbol{\phi}}\right)$$

$$= S\left(\int d\boldsymbol{\phi}\, p_{\boldsymbol{\phi}} U_{\boldsymbol{\phi}}^{\otimes N}\rho_0 (U_{\boldsymbol{\phi}}^{\dagger})^{\otimes N}\right)$$

$$\leqslant S\left(\int d\boldsymbol{\phi}\, p_{\boldsymbol{\phi}} \sum_{\boldsymbol{n}} P_{\boldsymbol{n}}^{(k)} U_{\boldsymbol{\phi}}^{\otimes N}\rho_0 (U_{\boldsymbol{\phi}}^{\dagger})^{\otimes N} P_{\boldsymbol{n}}^{(k)}\right) \qquad \text{(C5)}$$

and, by applying the unitaries to the projectors, we have

$$S\left(\int d\boldsymbol{\phi}\, p_{\boldsymbol{\phi}} \sum_{\boldsymbol{n}} P_{\boldsymbol{n}}^{(k)} U_{\boldsymbol{\phi}}^{\otimes N}\rho_0 (U_{\boldsymbol{\phi}}^{\dagger})^{\otimes N} P_{\boldsymbol{n}}^{(k)}\right)$$

$$= S\left(\int d\boldsymbol{\phi}\, p_{\boldsymbol{\phi}} \sum_{\boldsymbol{n}} P_{\boldsymbol{n}}^{(k)} \rho_0 P_{\boldsymbol{n}}^{(k)}\right). \qquad \text{(C6)}$$

Then, by exploiting the normalization of the joint probability distribution $\int d\boldsymbol{\phi}\, p_{\boldsymbol{\phi}} = 1$, we find

$$S\left(\int d\boldsymbol{\phi}\, p_{\boldsymbol{\phi}} \sum_{\boldsymbol{n}} P_{\boldsymbol{n}}^{(k)}\rho_0 P_{\boldsymbol{n}}^{(k)}\right) = S\left(\sum_{\boldsymbol{n}} P_{\boldsymbol{n}}^{(k)}\rho_0 P_{\boldsymbol{n}}^{(k)}\right)$$

$$= S\left(\sum_{\boldsymbol{n}} \lambda_{\boldsymbol{n},\boldsymbol{n}} |\boldsymbol{n}\rangle\langle\boldsymbol{n}|\right) = H(\lambda_{\boldsymbol{n},\boldsymbol{n}}) + \sum_{\boldsymbol{n}} \lambda_{\boldsymbol{n},\boldsymbol{n}} S(|\boldsymbol{n}\rangle\langle\boldsymbol{n}|)$$

$$= H(\lambda_{\boldsymbol{n},\boldsymbol{n}}) \leqslant \log_2(M). \qquad \text{(C7)}$$

The quantity $H(\lambda)$ appearing in the last two lines is the Shannon entropy $H(\lambda) \equiv -\sum_k \lambda_k \log_2 \lambda_k$.

Note that the number of projectors included in the second line is strictly connected to the number of encoded phases k, accordingly with the definition of M in Eq. (C4), which eventually provides the bound

$$I(\tilde{\boldsymbol{\phi}} : \boldsymbol{\phi}) \leqslant \log_2 \binom{N+k}{N} \qquad \text{(C8)}$$

and completes the proof. ∎

---

[1] S. L. Braunstein, C. M. Caves, and G. J. Milburn, Ann. Phys. (NY) **247**, 135 (1996).

[2] Z. Y. Ou, Phys. Rev. A **55**, 2598 (1997).

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

[4] M. G. A. Paris, Int. J. Quantum Inf. **07**, 125 (2009).

[5] M. Barbieri, PRX Quantum **3**, 010202 (2022).

[6] M. Szczykulska, T. Baumgratz, and A. Datta, Adv. Phys.-X **1**, 621 (2016).

[7] P. C. Humphreys, M. Barbieri, A. Datta, and I. A. Walmsley, Phys. Rev. Lett. **111**, 070403 (2013).

[8] T. Baumgratz and A. Datta, Phys. Rev. Lett. **116**, 030801 (2016).

[9] R. Demkowicz-Dobrzański, W. Górecki, and M. Guţă, J. Phys. A: Math. Theor. **53**, 363001 (2020).

[10] H. L. V. Trees, *Detection, Estimation, and Modulation Theory* (Wiley, New York, 2001).

[11] J. Ziv and M. Zakai, IEEE Trans. Inf. Theory **15**, 386 (1969).

[12] M. Hayashi, Phys. Lett. A **354**, 183 (2006).

[13] H. Imai and M. Hayashi, New J. Phys. **11**, 043034 (2009).

[14] M. Hayashi, Z.-W. Liu, and H. Yuan, Quantum Sci. Technol. **7**, 025030 (2022).

[15] M. Hayashi, S. Vinjanampathy, and L. C. Kwek, J. Phys. B **52**, 015503 (2019).

[16] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).

[17] G. M. D'Ariano, C. Macchiavello, and M. F. Sacchi, Phys. Lett. A **248**, 103 (1998).

[18] C. Macchiavello, Phys. Rev. A **67**, 062302 (2003).

[19] V. Gebhart, A. Smerzi, and L. Pezzè, Phys. Rev. Appl. **16**, 014035 (2021).

[20] W. Górecki and R. Demkowicz-Dobrzański, Phys. Rev. Lett. **128**, 040504 (2022).

[21] M. Hassani, C. Macchiavello, and L. Maccone, Phys. Rev. Lett. **119**, 200502 (2017).

[22] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 2006).

[23] M. Hayashi, *A Group Theoretic Approach to Quantum Information* (Springer, Berlin, 2017).

[24] H. P. Yuen, in *Quantum Squeezing*, edited by P. D. Drummond and Z. Ficek, Springer Series on Atomic, Optical, and Plasma Physics Vol. 27 (Springer, Berlin, 2004), pp. 227–261.

[25] R. Nair, J. Phys. A: Math. Theor. **51**, 434001 (2018).

[26] M. J. W. Hall, J. Phys. A: Math. Theor. **51**, 364001 (2018).

[27] M. J. W. Hall and H. M. Wiseman, Phys. Rev. X **2**, 041006 (2012).

[28] M. J. Holland and K. Burnett, Phys. Rev. Lett. **71**, 1355 (1993).

[29] A. Luis and J. Peřina, Phys. Rev. A **54**, 4564 (1996).

[30] M. Hayashi, Commun. Math. Phys. **347**, 3 (2016).

[31] V. Bužek, R. Derka, and S. Massar, Phys. Rev. Lett. **82**, 2207 (1999).

[32] E. Davies, IEEE Trans. Inf. Theory **24**, 596 (1978).

[33] W. van Dam, G. M. D'Ariano, A. Ekert, C. Macchiavello, and M. Mosca, Phys. Rev. Lett. **98**, 090501 (2007).

[34] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[35] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca, Proc. R. Soc. London A **454**, 339 (1998).

[36] T.-C. Wei and P. M. Goldbart, Phys. Rev. A **68**, 042307 (2003).

[37] H. Zhu, L. Chen, and M. Hayashi, New J. Phys. **12**, 083002 (2010).

[38] J. Martin, O. Giraud, P. A. Braun, D. Braun, and T. Bastin, Phys. Rev. A **81**, 062347 (2010).

[39] R. F. Werner, Phys. Rev. A **58**, 1827 (1998).