Discrete-to-continuous transition in quantum phase estimation

Wojciech Rządkowski and Rafal Demkowicz-Dobrzański
Faculty of Physics, University of Warsaw, ul. Pasteura 5, PL-02-093 Warszawa, Poland

We analyze the problem of quantum phase estimation where the set of allowed phases forms a discrete $N$ element subset of the whole $[0, 2\pi]$ interval, $\varphi_n = 2\pi n/N$, $n = 0, \ldots, N - 1$ and study the discrete-to-continuous transition $N \to \infty$ for various cost functions as well as the mutual information. We also analyze the relation between the problems of phase discrimination and estimation by considering a step cost functions of a given width $\sigma$ around the true estimated value. We show that in general a direct application of the theory of covariant measurements for a discrete subgroup of the $U(1)$ group leads to suboptimal strategies due to an implicit requirement of estimating only the phases that appear in the prior distribution. We develop the theory of sub-covariant measurements to remedy this situation and demonstrate truly optimal estimation strategies when performing transition from a discrete to the continuous phase estimation regime.

I. INTRODUCTION

Quantum phase estimation is a paradigmatic model capturing the essence of all interferometric experiments irrespectively whether they are performed using atoms or light [1, 2]. This is at the same time the best studied model in the field of theoretical quantum estimation theory [3, 4] and lies at the very foundations of the whole field of quantum metrology [5–8]. It has been studied both in idealized scenarios [9–13] as well as in presence of various decoherence effects [14–20].

The problem has also been analyzed using two different conceptual perspectives: the frequentist approach and the Bayesian one. The first approach focuses on scenarios where the experiment is repeated many times, and provides useful bounds on the performance of optimal estimator in the form of the famous Cramér-Rao bound. The main tool here is the concept of Fisher information or its quantum generalization—the Quantum Fisher Information (QFI) [3, 21]. The second approach, requires specification of the prior distribution describing the knowledge on the parameter to be estimated but is capable of providing operationally meaningful results dealing directly with single-shot experiments without the need to go into the limit of many independent experiment repetitions [13, 22, 23].

Within the quantum estimation theory, both approaches are applied to models where the phase parameter to be estimated is treated as continuous. In the frequentist approach, this is manifested explicitly in the definition of the QFI, where derivatives with respect to the estimated parameter appear. In the Bayesian approach, one typically chooses a natural flat prior distribution for the phase $p(\varphi) = 1/2\pi$, representing our complete ignorance on the actual value of the phase.

In this paper we analyze the phase estimation situation in case the set of allowed phases is discrete: $\varphi \in \mathcal{S}_N$, $\mathcal{S}_N = \{n\theta, \theta = 2\pi/N, n = 0, \ldots, N - 1\}$, and analyze the transition to the continuous limit $N \to \infty$. Discretization of phase is often encountered in optical communication protocols, where discrete set of phases is encoded in states of light in the so-called phase-shift keying protocols [23, 24]. In metrological scenarios, such situations where the phase to be estimated is “quantized” are not commonly encountered, but can be of relevance for examples in models where atoms passing through an optical cavity acquire a phase proportional to the number of photons inside [28]. This said, we admit that our main motivation here is to understand the discrete-to-continuous transition from a purely conceptual perspective as we find this important aspect of quantum phase estimation surprisingly unexplored.

Clearly, the formulation of the problem, sets us immediately into the Bayesian framework, as the problem can be phrased in a Bayesian language by stating that the prior probability for phase estimation is simply $p(\varphi) = 1/\mathcal{S}_N = 1/N \sum_{n=0}^{N-1} \delta(\varphi - n\theta)$. One can argue that assuming a discrete set of allowed phases moves us from the problem of estimation of phase to the problem of phase discrimination [29]. Indeed, this can be viewed this way, but an important element to be specified here is the explicit form of the cost function that we assume in the problem. If we choose a simple delta cost function penalizing us equally strongly whenever we guess the wrong phase, we will indeed reduce our problem to the one studied in the quantum discrimination literature [29], and in a sense the discrete to continuous transition for such models will be trivial as discussed explicitly in Sec. IV. Still, for any other cost function, the transition will be nontrivial and one cannot directly utilize known results from quantum state discrimination theory.

This paper is organized as follows. In Sec. I we formulate the problem of discrete phase estimation considered throughout this work. Secs. II and IV provide details of discrete-to-continuous transition while focusing on two popular cost functions: the $4\sin^2 \frac{\varphi - \tilde{\varphi}}{2}$ cost function commonly used in continuous phase estimation problems and the fixed interval cost function more natural in discrimination-like problems, respectively. After discussing these two examples, in Sec. V we discuss behavior of the optimal estimation protocols without assuming any particular choice of a cost function and we
indicate conditions under which going beyond standard covariant measurements may lead to a reduced estimation cost. Sec. [V] provides a more abstract and formal consideration of the use of sub-covariant measurements in a general problem of sub-group element estimation, of which discrete phase estimation is a special case. In Sec. [VII], we describe a numerical framework within which one can numerically optimize a general sub-group estimation problem using the idea of sub-covariant measurements. The discrete-to-continuous transition is studied using a different figure of merit—mutual information in Sec. [VIII]. The last Sec. [IX] concludes the paper.

II. DISCRETE PHASE ESTIMATION

Let us consider a $D$-dimensional quantum system, where the phase parameter $\varphi$ is encoded on a general input probe state $|\varphi_0\rangle = \sum_{k=0}^{D-1} c_k |k\rangle$ as follows:

$$|\varphi\rangle = U_\varphi |\varphi_0\rangle = \sum_{k=0}^{D-1} c_k e^{ik\varphi} |k\rangle. \quad (1)$$

This model may represent the evolution of a single mode state of light traveling through a phase delaying element, in which case $|k\rangle$ represents a $k$-photon state. Equivalently, in a more physically meaningful scenario, one can think of a two arm interferometer where a $(D-1)$-photon state has been split into two arms of the interferometer and $\varphi$ represent the relative phase delay for the light traveling through the upper and lower arms. In this case $|k\rangle$ should be understood as $|k\rangle \otimes |D-1-k\rangle$, representing the state where $k$ photons go through the upper and $D-1-k$ photons go through the lower arm of the interferometer. Analogously in atomic Ramsey interferometry, we would think of $|k\rangle$ as representing the state with $k$ atoms in the exited and $D-1-k$ atoms in the ground state [30].

Let $C(\varphi, \hat{\varphi})$ be the cost function representing the cost of estimating the value $\hat{\varphi}$ while the true value of the phase is $\varphi$. In what follows, we will naturally assume $C$ to depend only on the difference of the phases $C(\varphi, \hat{\varphi}) = C(\varphi - \hat{\varphi})$. Since in discrete phase estimation, allowed phases are $\varphi_n = n\theta$, the average cost is given by:

$$\bar{C} = \frac{1}{N} \sum_{n=0}^{N-1} \sum_{m=0}^{D-1} \text{Tr} (|\varphi_n\rangle\langle \varphi_n| \Pi_m) C(\varphi_n - \hat{\varphi}(m)), \quad (2)$$

where $\{\Pi_n\}$ represent a generalized POVM measurement [31]. $\Pi_0 \geq 0$, $\sum_m \Pi_m = 1$, while $\hat{\varphi}$ is an estimator function assigning a given value of phase $\hat{\varphi}(m)$ to a given measurement outcome $m$. Determining the optimal discrete phase estimation protocol amounts to minimizing the above quantity over the input state $|\varphi_0\rangle$, measurement $\{\Pi_n\}$, and the estimator $\hat{\varphi}$.

Even though the above optimization appears extremely challenging, the symmetry of the problem helps to simplify it considerably. In case of continuous phase estimation, the problem has a natural symmetry with respect to phase shifts, or more formally $U(1)$ group. The prior distribution $p(\varphi) = 1/2\pi$ is invariant under phase shifts, as well as the cost function $C(\varphi + \xi, \varphi + \xi') = C(\varphi, \varphi')$, while the family of states $|\varphi\rangle$ is obtained by acting with a unitary representation $U_\varphi$ of the group on the probe state $|\varphi_0\rangle$. This is an example of a covariant estimation problem [4]. In this case, it has been proven that while looking for the optimal estimation scenarios one can restrict oneself to the so-called covariant measurements, where respective measurement operators are generated from a single seed POVM by the action of the representation $U_\varphi$: $\Pi_\varphi = U_\varphi \Pi_0 U_\varphi^\dagger$. Note that measurement operators are labeled by a continuous parameter $\varphi$, where it is implicitly assumed that the index represents also the value of the estimated phase, given a particular measurement result occurs. Thanks to the use of covariant measurements the formula for the average cost function simplifies to:

$$\bar{C} = \frac{1}{2\pi} \int d\varphi \text{Tr} (|\varphi\rangle\langle \varphi| \Pi_0) C(\varphi). \quad (3)$$

One of standard choices for the cost function in phase estimation literature is $C(\varphi) = 4\text{sin}^2 \frac{\varphi}{2}$. This function is the simplest (in the sense of expansion into a Fourier series in $\cos(m\varphi)$) non-trivial function that approximates squared error for small phase deviations. For the problem of phase estimation with the above chosen cost function, minimization of the above formula over $|\varphi_0\rangle$ and $\Pi_0$ can be done analytically [13], and results in $\Pi_0 = D|e\rangle\langle e|$, $|e\rangle = \sqrt{D} \sum_{k=0}^{D-1} |k\rangle$ and

$$|\varphi_0\rangle = \sqrt{\frac{2}{D+1}} \sum_{k=0}^{D-1} \sin \left( \frac{k + \frac{1}{2}}{D+1} \pi \right) |k\rangle, \quad (4)$$

yielding the minimal cost $\bar{C} = 2 \left[ 1 - \cos \left( \frac{\pi}{D+1} \right) \right]$. It would seem that when considering a discrete phase estimation problem, all one has to do is replace the whole $U(1)$ group with its discrete $N$-element subgroup $U_N$ and proceed as before. Indeed the problem is covariant with respect to the $U_N$ group, so the discrete variant of [4] reads

$$\bar{C} = \frac{1}{N} \sum_{n=0}^{N-1} \text{Tr} (|\varphi_n\rangle\langle \varphi_n| \Pi_0) C(\varphi_n). \quad (5)$$

Since the above formula can be rewritten as

$$\bar{C} = \langle \varphi_0 | A | \varphi_0 \rangle, \quad A = \frac{1}{N} \sum_{n=0}^{N-1} U_{\varphi_n} \Pi_0 U_{\varphi_n}^\dagger C(\varphi_n) \quad (6)$$

the optimal input probe state can be determined as the eigenvector corresponding to the minimal eigenvalue of the matrix $A$, while the optimal measurement can always be chosen as $\Pi_0 = D|e\rangle\langle e|$. [4].

As we discuss explicitly further on in the paper, this approach for determining the optimal discrimination strategy may sometimes lead to a counter-intuitive result that cost of discrete estimation for some finite $N$ appears to be larger than that of the continuous case. This
apparent paradox may be understood if we realize an implicit assumption we have additionally made while moving from the continuous to a discrete case, while keeping the structure of covariant measurements. Namely, we have assumed that estimated values \( \hat{\varphi}(m) \) are restricted to belong to the same set \( S_N \) of phases that are being encoded on the input state. Depending on the choice of the cost function, it might be the case that in order to minimize \( \bar{C} \) it is advantageous to estimate phases outside \( S_N \). In what follows we will refer to such strategies as sub-covariant measurements.

### III. STANDARD PHASE ESTIMATION COST FUNCTION

We start by considering the cost function \( C(\varphi) = 4 \sin^2(\varphi/2) = 2(1 - \cos \varphi) \) commonly encountered in continuous phase estimation problems. Let us first restrict to proper covariant measurements, so assume \( \hat{\varphi} \in S_N \), and \( \Pi_{\hat{\varphi}} = U_{\hat{\varphi}} \Pi_0 U_{\hat{\varphi}}^\dagger \). When substituting the explicit parametrization of the input state \( |\varphi_0\rangle = \sum_k c_k |k\rangle \) to the average cost reads:

\[
\bar{C} = 2 - \frac{1}{N} \sum_{n=0}^{N-1} 2 \cos(n \theta) \sum_{j,k=0}^{D-1} (\Pi_0)_{jk} c_j^* c_k e^{i(j-k)n \theta}. \tag{7}
\]

In case \( N = D \) a simple choice \( \Pi_0 = D|e\rangle \langle e|, |\varphi_0\rangle = |e\rangle \) leads to \( \bar{C} = 0 \), as \( \sum_{j,k=0}^{D-1} e^{ijn \theta} = \sum_{j=0}^{N-1} e^{2\pi ijN} = N \delta_{j,0} \) — the discrete Fourier transform property. Similarly \( \bar{C} \) can easily be made 0 for \( N < D \). One can simply choose the same measurement operator but consider an input state supported on a \( N \) dimensional subspace \( |\varphi_0\rangle = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} |k\rangle \) and the same argument follows. The above observation is in fact general. Irrespective of the cost function for \( N \leq D \) the cost can be made 0. This is simply a manifestation of the fact that the phase transformation is capable of generating up to \( D \) orthogonal states which can be perfectly discriminated.

The first non-trivial case is therefore \( N = D + 1 \). Due to \( \cos(n \theta) \) factor while performing the sum over \( n \) the following sums will appear \( \sum_{n=0}^{N-1} e^{i(j-k+1)n \theta} = N \delta_{j,k} \). This sums will be nonzero iff \( |j-k| \leq 1 \leq N \) (note that since \( D = N - 1 \) they \( |j-k| \pm 1 \leq N \)). Hence only first off-diagonal terms will contribute to the cost. Now, observe that had we considered continuous case \( N \to \infty \), the conclusions would remain the same. This implies that the cost, measurement as well as the optimal state will be the same as in the continuous case. This proves a rather surprising fact, that in this model estimating \( D + 1 \) phases is equally challenging as estimating a phase continuous parameter. This also implies that for \( N \leq D \) and \( N \geq D + 1 \) standard covariant measurement are optimal as they provide perfect discrimination for \( N \leq D \) and are known to be optimal in the continuous limit. Since all the sums we perform result in identical formulas as the ones in the continuous case, then if indeed a class of more general measurements allowed to further lower the cost then the same measurement could be used in the continuous case and would lead to a contradiction with the known optimal result for continuous phase estimation.

Let us also note that in order to solve this problem one could have also resorted to methods of where the problem of phase estimation with arbitrary prior distribution has been analyzed for the case of \( 4 \sin^2 \varphi/2 \) cost function—in particular a discrete prior distribution.

### IV. DISCRIMINATION-LIKE COST FUNCTION

Let us now consider the following step cost function: \( C(\varphi) = 0 \) for \( |\varphi| < \sigma/2 \) and 1 otherwise. If we take the limit \( \sigma \to 0 \) then we face phase discrimination problem, where unless we guess the correct phase we are penalized equally. From quantum state discrimination theory [29] it is known that for \( N \) arbitrary quantum states \( |\psi_n\rangle \) generated by any unitary \( |\psi_n\rangle = U_n |\psi\rangle \) such that \( U_n^\dagger U_n = 1 \) the optimal measurement minimizing the discrimination cost are the so-called square root measurements defined as follows: \( \Pi_n = \frac{1}{N} \rho^{-1/2} |\psi_n\rangle \langle \psi_n| \rho^{-1/2} \), where \( \rho = \frac{1}{N} \sum_{n=0}^{N-1} |\psi_n\rangle \langle \psi_n| \).

In our case \( U = U_\theta = \text{diag}(1,e^{i\theta},...,e^{i(D-1)\theta}) \), \( |\psi_n\rangle = |\varphi_n\rangle \) and for \( N \geq D \) we obtain that for input state with positive coefficients \( c_k > 0 \), the optimal square root measurement is just a standard covariant measurement \( \Pi_{\varphi_\theta} = U_{\varphi_\theta} \Pi_0 U_{\varphi_\theta}^\dagger \) with \( \Pi_0 = D|e\rangle \langle e| \) (for other states we just need to correct for the phases of coefficients \( c_k \)). The optimal state is the one that minimizes the mean cost, which in this case amounts to:

\[
\bar{C} = 1 - \frac{D}{N} |\langle \varphi_0|e\rangle|^2 = 1 - \frac{D}{N} \quad \text{for } N \geq D, \tag{8}
\]

and the above optimal value is obtained for \( |\varphi_0\rangle = |e\rangle \). For \( N \leq D \) as before the cost is zero.

The above formula is valid only in the limit \( \sigma \to 0 \). For any finite \( \sigma \) we expect the above formula to hold at least in the regime where width of the cost function \( \sigma \) is smaller than the separation between the phases \( \theta \), as in this case the width of the cost function should not play a role. At the point were more than one phase can fit into the width of the cost function the situation ceases to correspond to a discrimination problem and the more phases fit into the cost function width the more “estimation-like” the problem becomes.

We first assume we restrict ourselves to covariant measurements. We take \( \Pi_0 = D|e\rangle \langle e| \) while the optimal state we determine performing determining the vector corresponding to the minimal eigenvalue of matrix \( A \) as defined in Eq. (6). The results of the above procedure are depicted as a gray line in Fig. 1 for an exemplary case \( D = 10 \) and \( \sigma = \pi/10 \). We can observe an apparently paradoxical behavior of the cost for some \( N \) above the one achieved in the continuous limit, which could be phrased in a way that discriminating a subset of states is more difficult than the whole set, a clear contradiction.
In order to remedy for this apparent paradox one should go beyond the standard covariant measurement class and consider the possibility of estimating phases which are not inside the set of encoded phases. In the considered case it is enough to include shifted-covariant measurements where the estimated phases correspond not to \( \varphi_n = n\theta \) but to values exactly in between the actual phases encoded in the state, i.e., \( \varphi_n + \theta/2 \). In other words we should replace \( \Pi_0 \) with \( \Pi_{\varphi} = DU_{\varphi_0}\langle e|e\rangle U_{\varphi_0}^\dagger \). This is a special case of the class of measurements we refer to as sub-covariant measurements, which are described in a formal way in Sec. [V]. Depending on the value of \( N \) one should switch between the two strategies. The resulting minimal cost is depicted as black dots in Fig. 1.

The results are easy to understand. For \( N \leq 10 \), the \( \bar{C} = 0 \) as expected. Then for \( 11 \leq N \leq 20 \) the curve follows the \( 1 - \frac{\theta}{\sigma} \) formula corresponding to the optimal phase discrimination problem. At \( N = 21 \), we observe a change due to the fact that the separation between the phases \( 2\pi/21 \) is smaller than \( \sigma \) and hence two phases can fit into the width of the cost function. The optimal measurement in this case corresponds to sub-covariant measurement. Another transition happens at \( N = 41 \). This a point where \( \sigma > 2\theta \) and hence three phases can fit into one width of the cost function. At this point we recover the optimality of the standard covariant measurement. This situation repeats itself at every \( N = k \times 20 + 1 \), switching back and forth between the optimal covariant and shifted-covariant strategies.

In the continuous limit \( N \to \infty \) we may of course use the standard covariant measurements. Plugging in \( \Pi_0 = D|e\rangle\langle e| \) into (3) and integrating over \( \varphi \) for our step function cost we get:

\[
\bar{C} = \langle \varphi_0|B|\varphi_0 \rangle, \quad B_{jk} = \begin{cases} 
1 - \frac{\sigma}{\pi} & j = k \\
\frac{\sin[(j-k)/2]}{\sin[(j+k)/2]} & j \neq k.
\end{cases}
\]

The eigenproblem of matrix \( B \) is well known from Fourier analysis [22] and has also been studied in the context of quantum communication [33, 34]. The optimal input state is the eigenvector of matrix \( B \) corresponding to the smallest eigenvalue which is then the minimal cost. The coefficients of this optimal eigenvector form a Discrete Prolate Spheroidal Sequence (DPSS) [32]. The smallest eigenvalue (minimal cost) corresponding to the DPSS eigenvector is plotted as a dashed line in Fig. 1. Note that thanks to the use of sub-covariant measurement the optimal cost for finite \( N \) never exceeds the one corresponding to the continuous limit which should be the case as the continuous phase estimation should be never be less difficult than the discrete one.

### V. General Cost Function

Having studied this particular two examples, let us move on to a more general discussion where we will provide some general statements without specifying a particular form of the cost function. We consider a general cost function given by

\[
C(\varphi) = \sum_{m=0}^{M} \alpha_m \cos(m\varphi),
\]

where the Fourier series structure of the parametrization reflects the periodicity of the function and we will refer to \( M \) as order of the cost function. For the cost function to be meaningful, the coefficients \( \alpha_m \) should be such to make \( C(\varphi) \geq 0 \), and \( C(0) = 0 \). Let us first restrict to proper covariant measurements, so assume \( \varphi \in S_N \), and \( \Pi_\varphi = U_\varphi \Pi_0 U_\varphi^\dagger \). After inserting the cost function (10) and the explicit parametrization of the input state \( |\varphi_0 \rangle = \sum_k c_k |k \rangle \) to (3) the average cost reads:

\[
\bar{C} = \sum_{m=0}^{M} \alpha_m \frac{1}{N} \sum_{n=0}^{N-1} \cos(mn\theta) \sum_{j,k=0}^{D-1} (\Pi_0)_{jk} c_j^* c_k e^{i(j-k)n\theta}.
\]

As already discussed when analyzing the standard phase estimation cost function, in case \( N \leq D \), the cost can always be made zero.

Let us now consider \( N \geq D+1 \). Due to \( \cos(mn\theta) \) factor while performing the sum over \( n \), the following sums will appear \( \sum_{n=0}^{N-1} e^{i[(j-k)+am]n2\pi/N} \). This sums will be nonzero iff \( (j-k) \pm m = 0, \pm N, \pm 2N, \ldots \). If we now increase number of phases such that \( N = D + M \), then \( |(j-k) \pm m| < N \)

\[
C = \langle \varphi_0|B|\varphi_0 \rangle, \quad B_{jk} = \begin{cases} 
1 - \frac{\sigma}{\pi} & j = k \\
\frac{\sin[(j-k)/2]}{\sin[(j+k)/2]} & j \neq k.
\end{cases}
\]

The eigenproblem of matrix \( B \) is well known from Fourier analysis [22] and has also been studied in the context of quantum communication [33, 34]. The optimal input state is the eigenvector of matrix \( B \) corresponding to the smallest eigenvalue which is then the minimal cost. The coefficients of this optimal eigenvector form a Discrete Prolate Spheroidal Sequence (DPSS) [32]. The smallest eigenvalue (minimal cost) corresponding to the DPSS eigenvector is plotted as a dashed line in Fig. 1. Note that thanks to the use of sub-covariant measurement the optimal cost for finite \( N \) never exceeds the one corresponding to the continuous limit which should be the case as the continuous phase estimation should be never be less difficult than the discrete one.

### V. General Cost Function

Having studied this particular two examples, let us move on to a more general discussion where we will provide some general statements without specifying a particular form of the cost function. We consider a general cost function given by

\[
C(\varphi) = \sum_{m=0}^{M} \alpha_m \cos(m\varphi),
\]

where the Fourier series structure of the parametrization reflects the periodicity of the function and we will refer to \( M \) as order of the cost function. For the cost function to be meaningful, the coefficients \( \alpha_m \) should be such to make \( C(\varphi) \geq 0 \), and \( C(0) = 0 \). Let us first restrict to proper covariant measurements, so assume \( \varphi \in S_N \), and \( \Pi_\varphi = U_\varphi \Pi_0 U_\varphi^\dagger \). After inserting the cost function (10) and the explicit parametrization of the input state \( |\varphi_0 \rangle = \sum_k c_k |k \rangle \) to (3) the average cost reads:

\[
\bar{C} = \sum_{m=0}^{M} \alpha_m \frac{1}{N} \sum_{n=0}^{N-1} \cos(mn\theta) \sum_{j,k=0}^{D-1} (\Pi_0)_{jk} c_j^* c_k e^{i(j-k)n\theta}.
\]

As already discussed when analyzing the standard phase estimation cost function, in case \( N \leq D \), the cost can always be made zero.

Let us now consider \( N \geq D+1 \). Due to \( \cos(mn\theta) \) factor while performing the sum over \( n \), the following sums will appear \( \sum_{n=0}^{N-1} e^{i[(j-k)+am]n2\pi/N} \). This sums will be nonzero iff \( (j-k) \pm m = 0, \pm N, \pm 2N, \ldots \). If we now increase number of phases such that \( N = D + M \), then \( |(j-k) \pm m| < N \)

\[
C = \langle \varphi_0|B|\varphi_0 \rangle, \quad B_{jk} = \begin{cases} 
1 - \frac{\sigma}{\pi} & j = k \\
\frac{\sin[(j-k)/2]}{\sin[(j+k)/2]} & j \neq k.
\end{cases}
\]

The eigenproblem of matrix \( B \) is well known from Fourier analysis [22] and has also been studied in the context of quantum communication [33, 34]. The optimal input state is the eigenvector of matrix \( B \) corresponding to the smallest eigenvalue which is then the minimal cost. The coefficients of this optimal eigenvector form a Discrete Prolate Spheroidal Sequence (DPSS) [32]. The smallest eigenvalue (minimal cost) corresponding to the DPSS eigenvector is plotted as a dashed line in Fig. 1. Note that thanks to the use of sub-covariant measurement the optimal cost for finite \( N \) never exceeds the one corresponding to the continuous limit which should be the case as the continuous phase estimation should be never be less difficult than the discrete one.
Hence we can draw a general conclusion – after being given an arbitrary cost function with order \( M \), we predict the minimal cost to be zero for \( N \leq D \), and then increase up to \( N = D + M \), at which point it already reaches its continuous limit. This also implies that for \( N \leq D \) and \( N \geq N + M \) standard covariant measurement will suffice to reach the optimal performance. However, one cannot a priori assume that covariant measurements will be also optimal in the regime \( D < N < D + M \).

Let us now further analyze the structure of formula (11). We can write it in the following form:

\[
\bar{C} = \sum_{n=0}^{N} \sigma(\phi_n)C(\phi_n),
\]

(12)

where \( \sigma(\phi) = \text{Tr}(|\phi\rangle\langle\phi|\Pi_0)/N \). Intuitively, the continuous estimation corresponds to integrating the function \( \sigma(\phi)C(\phi) \) from 0 to 2\( \pi \), while discrete estimation corresponds to summing discretely probed values of this function. The simplest step beyond the covariant strategy is to introduce a non-zero offset \( \xi \):

\[
\bar{C}_\xi = \sum_{n=0}^{N} \sigma(\phi_n + \xi)C(\phi_n + \xi),
\]

(13)

i.e. to probe the function \( \sigma(\phi)C(\phi) \) at different values of \( \phi \). This physically is equivalent to providing estimated phases outside the set of true phases, but still keeping an equal number of true and estimated phases. This is what we referred to before as shifted-covariant measurement.

Due to natural symmetry of the problem considered, both \( C(\phi) \) and \( \sigma(\phi) \) are symmetric functions and hence also probability weighted cost function \( \sigma(\phi)C(\phi) \). This implies that \( \bar{C}_\xi \) has an extremum at \( \xi = 0 \). This also implies that another extremum should appear at \( \xi = \pi/N \) corresponding to a shift by half the interval between the encoded phases—increasing or decreasing the value of \( \xi \) around \( \xi = \pi/N \) has the same effect on \( \bar{C}_\xi \) as it can be seen as the same operation performed on original weighted cost function or on its reflection around \( \phi = 0 \) which is the same due to symmetry of the probability weighted cost function. We provide an illustrative example of this generic behavior in Fig. 2 for a simple qubit case \( D = 2 \) and \( N = 3 \) phases. In principle, there can be more extrema but we did not find them for any reasonable cost function studied. We did however observe additional extrema when considering mutual information as the figure of merit as outlined in Sec. VIII.

### VI. Sub-Covariant Measurements

In order to put our results in a more rigorous mathematical framework, we provide here a general theory of sub-covariant measurements, that help to remedy the apparent contradictions one may arrive at when applying the theory of standard covariant measurements to discrete parameter estimation problems, of which discrete
phase estimation is a special case. We would like to stress that this concept is completely general and may used in an arbitrary discrete parameter estimation models beyond phase estimation. For this reason we will formulate it in a general way, where an estimation model arises as a result of choosing a subgroup $H$ of a group $G$ that generates the original continuous estimation problem. Note that we do not even insist that the group $G$ that generates the original continuous estimation problem. Let $H$ be a subgroup of $G$. We associate elements of $H$ with possible true values of a quantum parameter encoded in a state and elements of $G$ with possible measurement outcomes which are also the possible estimated values of the parameter. By $c$ we denote the neutral element of $G$. Let $U_c$ be a unitary representation of group $G$ on the Hilbert space of interest.

**Definition.** A POVM measurement $\{\Pi_g\}$ is sub-covariant with respect to subgroup $H \leq G$ if and only if

$$\forall h \in H, \tilde{g} \in G U_h \Pi_g U_h^\dagger = \Pi_{h\tilde{g}}.$$  \hspace{1cm} (14)

Let us observe that according to the above definition every measurement sub-covariant with respect to subgroup $H$ is uniquely defined by a family of operators $\{\Pi_\gamma\}_{\gamma \in G/H}$ by the following relation:

$$\Pi_{\gamma} = U_h \Pi_{\gamma} U_h^\dagger,$$  \hspace{1cm} (15)

where $\tilde{g} = h \gamma$.

**Theorem.** If the estimation problem is sub-covariant with respect to $H \leq G$, i.e.: (i) $dh$—Haar measure on $H$, (ii) $\rho_h = U_h \rho_c U_h^\dagger$ for every $h \in H$ and (iii) $C(g''g', g''\tilde{g}) = C(g', g) \quad \text{for every} \quad g, g', g'' \in G$, then we can choose the optimal measurement from measurements sub-covariant with respect to $H \leq G$.

**Proof:** Let $\{\Pi_{\gamma}^{opt}\}$ be the optimal measurement which minimizes the average cost

$$\bar{C}(\Pi_{\gamma}^{opt}) = \int_H \int_G dh \int_G d\tilde{g} Tr(\Pi_{\gamma}^{opt} \rho_h) C(h, \tilde{g}).$$  \hspace{1cm} (16)

We define

$$\Pi_{\gamma}^{cov} = \int_H dh' U_{h'} \Pi_{\gamma}^{opt} U_{h'}^\dagger.$$  \hspace{1cm} (17)

Then $\Pi_{\gamma}^{cov}$ is sub-covariant with respect to $H$ since:

$$U_h \Pi_{\gamma}^{cov} U_h^\dagger = U_h \int_H dh' U_{h'} \Pi_{\gamma}^{opt} \Pi_{\gamma}^{opt} U_{h'}^\dagger = \int_H dh' U_{h'h^{-1}} \Pi_{\gamma}^{cov} U_{h'h^{-1}} = \int_H dh' U_{h'h^{-1}} \Pi_{\gamma}^{opt} U_{h'h^{-1}} = \Pi_{h\gamma}^{cov},$$  \hspace{1cm} (18)

where in the last equality we have substituted $h' \rightarrow h'h$, making use of the Haar measure property. Moreover, measurement $\Pi_{\gamma}^{cov}$ yields the same cost as $\Pi_{\gamma}^{opt}$, which can be seen as follows:

$$\bar{C}(\Pi_{\gamma}^{cov}) = \int_H \int_G dh \int_G d\tilde{g} Tr(\Pi_{\gamma}^{cov} \rho_h) C(h, \tilde{g}) = \int_H \int_G dh \int_G d\tilde{g} Tr(\Pi_{\gamma}^{opt} \rho_h U_h \Pi_{\gamma}^{opt} U_h^\dagger) C(h, \tilde{g}) = \int_H \int_G dh \int_G d\tilde{g} Tr(U_h \Pi_{\gamma}^{opt} U_h^\dagger \rho_c) C(h, \tilde{g}) = \int_H \int_G dh \int_G d\tilde{g} \int_G dh' Tr(U_{h'h} \Pi_{\gamma}^{opt} U_{h'h}^\dagger \rho_c) C(h'h^{-1}, h'^{-1} \tilde{g}) = \int_H \int_G dh \int_G d\tilde{g} \int_G dh' Tr(\Pi_{\gamma}^{opt} \rho_c) C(h, \tilde{g}) = \bar{C}(\Pi_{\gamma}^{opt}),$$

where in the third line we have substituted $h \rightarrow h'^{-1}h$, $\tilde{g} \rightarrow h'^{-1} \tilde{g}$. This ends the proof of the theorem.

Unlike in the full-covariant estimation problem, the search for the optimal measurement is not restricted here to identifying a single seed POVM $\Pi_0$, but rather a whole set $\Pi_{\gamma}$ where $\gamma \in G/H$. Still, this fact may significantly speed up the numerical search for the optimal estimation strategy as for sub-covariant measurements the average cost simplifies to:

$$\bar{C} = \int_{G/H} d\gamma \int_H dh Tr(U_h \Pi_{\gamma} U_h^\dagger \rho_c) C(0, h\gamma).$$  \hspace{1cm} (20)

Let us now discuss the relation of the above abstract and general theory with the phase estimation problem, especially with the examples presented in Secs. [XVI] and [XV]. In the continuous phase estimation problem $H = U(1)$ and sub-covariant measurements are not useful since there is no relevant group $G$ of which $H$ would be a nontrivial subgroup. In the discrete phase estimation problem $H = U_N$, and $\int_H dh \rightarrow \sum_{n=0}^{N-1}$, while in general $G = U(1)$. Hence the task is to find seed POVM set $\Pi_{\gamma}$, where $\gamma \in [0, \theta]$ = $U(1)/U_N$. However, is not always necessary to consider the full $G = U(1)$ group as it might be the case that the cost is minimal already for another discrete group $G = U_N$, $l \in Z$. An example is the shifted-covariant strategy optimal for some values of $N$ in Secs. [XVI] and [XV], which corresponds to $H = U_N$, $G = U_{2N}$ and the seed POVM set $\{0, DU_{\frac{\pi}{N}} \}$.

Let us also note that in the examples of discrete phase estimation discussed in the preceding sections we have not identified situations where more than one nonzero seed POVM would be necessary, and the optimal performance could always be reached by shifted-covariant measurements, which are the simplest non-trivial example of sub-covariant strategies. We expect, however, that for more involved discrete estimation problems more general class of measurements might be required. Therefore, in the next Sec. [VII] we provide a general algorithm to search for optimal sub-covariant strategies.
VII. ALGORITHM FOR IDENTIFYING THE OPTIMAL DISCRETE PHASE ESTIMATION STRATEGY

Here we show how to effectively use the concept of sub-covariant measurements in order to numerically find the optimal discrete parameter estimation problems. For concreteness, we will focus on discrete phase estimation, but the procedure can be directly generalized to any subgroup estimation problem using the framework described in Sec. VI.

First assume for the moment that the input state is fixed. We know that we can look for the optimal measurement/estimation strategy within the class of sub-covariant measurements. Hence we need to find the optimal seed POVM set \( \Pi_s \), where \( \gamma \in [0, \theta) = U(1)/U_N \). Clearly looking for continuous family of operators is not feasible numerically. Still one can proceed in steps. Let us define sets containing an increasing number of seed measurement operators: \( \mathcal{P}_1 = \{ \Pi_0 \} \), \( \mathcal{P}_2 = \{ \Pi_0, \Pi_\theta \} \), \( \mathcal{P}_3 = \{ \Pi_0, \Pi_{\theta \theta}, \Pi_{\theta \phi} \} \) etc. Considering a given number \( S \) of seed measurements, we minimize \( \bar{C} \) over operators \( \Pi_{\theta \theta} \), where \( s \in \{ 0, ..., S - 1 \} \), with the following constraints: \( \Pi_{\theta \theta} \geq 0, \frac{1}{N} \sum_{s=0}^{S-1} \sum_{n=0}^{N-1} U_{n0} \Pi_{\theta \theta} U_{n0}^\dagger = \mathbb{1} \). This is a standard semi-definite program and can be solved efficiently using e.g. the CVX package for Matlab [25]. We proceed by increasing \( S \) and once we observe that further increase in \( S \) does not reduce the cost we stop.

In order to find the optimal input state as well, one should adapt an iterative approach, as proposed in [22, 60], where the search for the optimal measurements is interleaved with the search for the optimal state. Applying the general form of Eq. (20) to our case of discrete quantum phase estimation, we obtain:

\[
\bar{C} = \langle \varphi_0 | A | \varphi_0 \rangle, \quad A = \frac{1}{N} \sum_{s=0}^{S-1} \sum_{n=0}^{N-1} U_{n0} \Pi_{\theta \theta} U_{n0}^\dagger C(n \theta + \frac{s}{S} \theta). \tag{21}
\]

We see that given a particular sub-covariant measurement, the search for the state minimizing the mean cost is equivalent to the search of an eigenvector corresponding to the minimal eigenvalue of the matrix \( A \). By iterating the above scheme of finding the measurement optimal for a particular input state, and the the state optimal for a particular measurement one arrives at the fully optimal solution.

In particular, we have followed this procedure to confirm that the discrimination strategy described in Sec. VI based on switching between covariant and shifted-covariant measurements was indeed optimal.

VIII. MUTUAL INFORMATION AS A FIGURE OF MERIT

In previous sections we have pursued an approach to quantum estimation/discrimination problems focused on minimization of Bayesian cost functions. This is a natural framework for determining optimal protocols in single-shot estimation/discrimination scenarios. One can, however, look at the discrete phase estimation setting as an example of a communication protocol, which is supposed to be repeated many times. In such situations a natural figure of merit is the mutual information between the encoded phases and measurement results which quantifies how many bits of noiseless information can be effectively transmitted per single use of the considered quantum channel. Given equiprobable phase encoded input states \( |\varphi_n\rangle = U_{n0}|\varphi_0\rangle \), \( \varphi_n = n \theta, n = 0, ..., N - 1 \), and POVM measurement \( \{ \Pi_m \} \), the resulting joined probability distribution of symbol \( n \) being sent and measurement outcome \( m \) being registers reads:

\[
p_{nm} = \frac{1}{N} \text{Tr}( |\varphi_n\rangle \langle \varphi_n | \Pi_m \rangle), \tag{22}
\]

while the corresponding mutual information reads:

\[
I = \sum_{nm} H(p_{nm}) - \sum_n H \left( \sum_m p_{nm} \right) - \sum_m H \left( \sum_n p_{nm} \right), \tag{23}
\]

where \( H(x) = -x \log_2 x \). The optimal protocol, from this point of view, is now the one that maximizes \( I \) over the choice of input state \( |\varphi_0\rangle \) and measurements \( \{ \Pi_m \} \).

Unlike the Bayesian cost, here \( I \) is nonlinear both in measurement and the input state which makes the maxi-
mization of $I$ much more challenging. For covariant problems, there is no general theorem on the optimality of covariant measurements in this case, except for the situation when the states used are generated by the action of an irreducible representation of a group. In our case the $U(1)$ representation is clearly reducible and hence Davies theorem cannot be invoked. Still, one can bound the maximal achievable mutual information using the famous Holevo bound \[4\.\] For our problem this bound implies:

$$I \leq \chi = S\left(\frac{1}{N} \sum_{n=0}^{N-1} |\varphi_n\rangle\langle\varphi_n|\right),$$

(24)

where $S(\rho) = -\text{Tr}\rho\log_2\rho$ is the von Neumann entropy. For $N \leq D$, maximum $\chi$ corresponds to situation where we have $N$ orthogonal states $|\varphi_n\rangle$ and $\chi = \log_2 N$. This bound can simply be achieved by using exactly the same strategy as described when minimizing Bayesian cost for $N \leq D$. Since $N$ orthogonal states are perfectly distinguishable by performing measurement in the basis containing the input states we can obtain $I = \log_2 N$, saturating the Holevo bound and proving that this strategy is indeed optimal. In general the Holevo bound can never be larger than $\log_2 D$ (maximum entropy is reached for $\rho = 1/D$), hence for $N > D$ the bound has the form $I \leq \log_2 D$ and does not further increase with $N$. Even though this bound cannot be in general achieved, it has been shown [39] that in the limit of continuous phase estimation $N \to \infty$, and large dimensions $D \to \infty$ one can reach almost optimal performance $I = \log_2 D - 1.2199$. This performance can be achieved using a strategy utilizing an equal superposition input states $|\varphi_0\rangle = |e\rangle = \frac{1}{\sqrt{D}} \sum_{k=0}^{D-1} |k\rangle$ and standard covariant continuous POVM, $\Pi_\xi = \frac{D}{\pi \xi} U_\xi^\dagger |e\rangle\langle e| U_\xi$. While studying discrete-to-continuous transition using mutual information figure of merit we will employ the above coding-decoding strategy. When going from the case $N = D$, where the mutual information can be made equal to the Holevo bound, to $N \to \infty$ (continuous regime) the accessible mutual information can only decrease, hence, we can sensibly expect that, since in the continuous limit the above described procedure performs almost optimally, it performs close to optimal also in the discrete case. Similarly as in the Bayesian approach, however, we have observed that it is essential to consider shifted-covariant measurements to obtain optimal performance. Note that when dealing with mutual information, there is no issue with the set of estimated values, as we deal solely with probability distribution. Still, when considering the specific class of POVMs as described above, we can naturally define a shifted-covariant POVM as $\Pi_\xi = \frac{D}{\pi \xi} U_{\xi+\tilde{\xi}}^\dagger |e\rangle\langle e| U_{\xi+\tilde{\xi}}$. In Fig. 3 we depict the results for the discrete-to-continuous transition for $D = 10$. Interestingly, in order to reach the optimal mutual information we occasionally have encountered situations where the optimum was reached for $0 < \xi < \pi/N$, unlike in Bayesian examples studied where either $\xi = 0$ and $\xi = \pi/N$ shifts were optimal. This is due to much more involved structure of the mutual information potential leading to more than two local extrema while changing the shift parameter $\xi$.

We leave it as an open question, whether the above described strategy is optimal, and in particular whether more general measurements, that go beyond the class shifted-covariant measurements, could be useful in increasing the mutual information which is known to be the case in some communication problems [40].

\section{IX. CONCLUSIONS}

In this paper, we have provided methods and results concerning the discrete phase estimation problem and its transition to the continuous limit. We have studied the transition for different cost functions and showed that in general one may need to go beyond the standard concept of covariant measurements and use a more general form—sub-covariant measurements introduced in this work. We believe that these results may find their applicability in quantum communication theory where phase shift keying protocols involve encoding a discrete set of phases on the transmitted states. We also see this work as a first step towards better understanding a general problem of digital-to-analog transition in decoding and encoding information using quantum states, e.g. in quantum reading protocols [41].

\section{ACKNOWLEDGMENTS}

We thank Michael J. W. Hall for pointing out the relationship between Sec. IV and Discrete Prolate Spheroidal Sequences. This work was supported by the Polish Ministry of Science and Higher Education Iuventus Plus program for years 2015-2017 No. 0088/IP3/2015/73.

[1] P. Hariharan, \textit{Optical interferometry} (Elsevier, Amsterdam, 2003).
[2] A. D. Cronin, J. Schmiedmayer, and D. E. Pritchard, \textit{Rev. Mod. Phys.} 81, 1051 (2009).
[3] C. W. Helstrom, \textit{Quantum detection and estimation theory} (Academic press, New York, 1976).
[4] A. S. Holevo, \textit{Probabilistic and Statistical Aspects of Quantum Theory} (North Holland, Amsterdam, 1982).
[5] V. Giovannetti, S. Lloyd, and L. Maccone, \textit{Nature Photonics} 5, 222 (2011).
[6] G. Toth and I. Apellaniz, \textit{J. Phys. A: Math. Theor.} 47, 424006 (2014).
[7] R. Demkowicz-Dobrzański, M. Jarzyna, and J. Kołodyński, in *Progress in Optics* Vol. 60, edited by E. Wolf (Elsevier, 2015) pp. 345–435.
[8] L. Pezzè, A. Smerzi, M. K. Oberthaler, R. Schmied, and P. Treutlein, ArXiv e-prints (2016), arXiv:1609.01609 [quant-ph].
[9] C. M. Caves, Phys. Rev. D 23, 1693 (1981).
[10] B. Yurke, S. L. McCail, and J. R. Klauder, Phys. Rev. A 33, 4033 (1986).
[11] M. J. Holland and K. Burnett, Phys. Rev. Lett. 71, 1355 (1993).
[12] J. P. Dowling, Phys. Rev. A 57, 4736 (1998).
[13] D. W. Berry and H. M. Wiseman, Phys. Rev. Lett. 85, 5098 (2000).
[14] S. F. Huelga, C. Macchiavello, T. Pellizzari, A. K. Ekert, M. B. Plenio, and J. I. Cirac, Phys. Rev. Lett. 79, 3865 (1997).
[15] A. Shaji and C. M. Caves, Phys. Rev. A 76, 032111 (2007).
[16] U. Dorner, R. Demkowicz-Dobrzański, B. J. Smith, J. S. Lundeen, W. Wasilewski, K. Banaszek, and I. A. Walmsley, Phys. Rev. Lett. 102, 040403 (2009).
[17] S. Knys, V. N. Smelyanskiy, and G. A. Durkin, Phys. Rev. A 83, 021804 (2011).
[18] M. G. Genoni, S. Olivares, and M. G. A. Paris, Phys. Rev. Lett. 106, 153603 (2011).
[19] B. M. Escher, R. L. de Matos Filho, and L. Davidovich, Nature Phys. 7, 406 (2011).
[20] R. Demkowicz-Dobrzański, M. Gută, and J. Kołodyński, Nat. Commun. 3, 1063 (2012).
[21] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. 72, 3439 (1994).
[22] R. Demkowicz-Dobrzański, Phys. Rev. A 83, 061802 (2011).
[23] M. J. W. Hall and H. M. Wiseman, New J. Phys. 14, 033040 (2012).
[24] M. Jarzyna and R. Demkowicz-Dobrzański, New Journal of Physics 17, 013010 (2015).
[25] K. Kitayama, *Optical Code Division Multiple Access. A practical Perspective* (Cambridge University Press, Cambridge, 2014).
[26] C. R. Miller and C. Marquardt, New Journal of Physics 17, 032003 (2015).
[27] M. Jarzyna, V. Lipińska, A. Klimek, K. Banaszek, and M. G. A. Paris, Opt. Express 24, 1693 (2016).
[28] M. Brune, J. Bernu, C. Guerlin, S. Deleglise, C. Sayrin, S. Gleyzes, S. Kuhr, I. Dotsenko, J. M. Raimond, and S. Haroche, Phys. Rev. Lett. 101, 240402 (2008).
[29] S. M. Barnett and S. Croke, Adv. Opt. Photon. 1, 238 (2009).
[30] N. F. Ramsey, Physics Today 33, 25 (1980).
[31] M. A. Nielsen and I. L. Chuang, *Quantum Computing and Quantum Information* (Cambridge University Press, Cambridge, 2000).
[32] D. Slepian, Bell System Technical Journal 57, 1371 (1978).
[33] M. J. W. Hall and I. G. Fuss, Quantum Optics: Journal of the European Optical Society Part B 3, 1 (2008).
[34] M. J. W. Hall, Journal of Modern Optics 40, 809 (1993).
[35] M. Grant and S. Boyd, “Cvx: Matlab software for disciplined convex programming,” (2012).
[36] K. Macieszczak, M. Fraas, and R. Demkowicz-Dobrzański, New Journal of Physics 16, 113002 (2014).
[37] E. Davies, IEEE Transactions on Information Theory 24, 596 (1978).
[38] M. Sasaki, S. M. Barnett, R. Jozsa, M. Osaki, and O. Hirota, Phys. Rev. A 59, 3325 (1999).
[39] L. Maccone, M. Hassani, and C. Macchiavello, ArXiv e-prints (2017), arXiv:1705.06666 [quant-ph].
[40] P. W. Shor, eprint arXiv:quant-ph/0009077 (2000), quant-ph/0009077.
[41] S. Pirandola, Phys. Rev. Lett. 106, 090504 (2011).