Beyond quantum Fisher information: optimal phase estimation with arbitrary a priori knowledge

Rafał Demkowicz-Dobrzański[1]

[1] *Institute of Theoretical Physics, University of Warsaw, ul. Hoża 69, PL-00-681 Warszawa, Poland*

The optimal phase estimation strategy is derived when partial a priori knowledge on the estimated phase is available. The structure of the optimal measurements, estimators and the optimal probe states is analyzed. The results fill the gap in the literature on the subject which until now dealt almost exclusively with two extreme cases: almost perfect knowledge (local approach based on Fisher information) and no a priori knowledge (global approach based on covariant measurements). Special attention is paid to a natural a priori probability distribution arising from a diffusion process.

PACS numbers: 03.65.Ta, 06.20.Dk, 42.50.St

Quantum states may be employed to improve the precision of measurements without committing additional resources such as energy or time [1–4]. A paradigmatic model is the estimation of relative phase delay between two arms of a Mach-Zehnder interferometer [5–8]. When the interferometer is fed with a light pulse prepared in a coherent state, the precision of phase estimation scales as $\delta_c\varphi \propto 1/\sqrt{N}$, where N is the mean number of photons in a pulse. On the other hand, by an appropriate preparation of N photons, entering the two input ports of the interferometer, a quadratic precision enhancement may be achieved leading to $\delta_q\varphi \propto 1/N$ referred to as the Heisenberg limit [9, 10]. The enhancement is due to the presence of entanglement between photons while they travel through the interferometer.

These idealized results need to be contrasted with a more realistic ones when environmental noise and experimental imperfections are taken into account [11–16]. For optical implementations the most relevant disruptive factor is the photon loss. If the overall power transmission of the interferometer (including scattering, reflections, detector efficiencies etc.) is denoted by η, then, for large N, precision of phase estimation approaches [17, 18]

$$\delta_c\varphi \approx 1/\sqrt{\eta N}, \quad \delta_q\varphi \approx \sqrt{1-\eta}/\sqrt{\eta N} \qquad (1)$$

for coherent state and the optimally entangled state respectively. The Heisenberg scaling is asymptotically lost and the quantum enhancement amounts to $\sqrt{1-\eta}$ factor. Although asymptotically the prospect of quantum enhanced metrology may look bleak, still for moderate N the advantage of using entangled states may be very significant [15].

Interestingly, the formulas (1) were obtained independently with two different approaches. The first, *local approach*, aims at maximizing the estimation sensitivity to small phase variations around an a priori known value $\varphi = \varphi_0$. The main tool is the *quantum Fisher information* (QFI), F_Q, which defines an asymptotically achievable quantum Cramér-Rao (CR) bound [19, 20] on estimation precision $\delta\varphi \geq 1/\sqrt{F_Q}$. The second, *global approach*, assumes no a priori knowledge on the phase — a priori probability distribution is uniform over the $[-\pi, \pi)$ region — and makes use of the phase shift symmetry to restrict the class of measurements to the covariant ones [21, 22]. It is quite intuitive that in asymptotic regime the role of a priori knowledge should be negligible, since in principle a minute fraction of resources might be committed to preestimate the phase and compensate for the lack of a priori knowledge.

For finite N, however, the amount of a priori knowledge may strongly influence both the optimal precision and the estimation strategy itself. Since the main prospects for applications of quantum enhanced metrology lay in the regime of moderate values of N, the amount of a priori knowledge may play a key role in designing the optimal phase estimation schemes. This paper provides a way to find the optimal estimation schemes for an arbitrary form of a priori knowledge. As a result, one can also easily study the transition from *global* to *local* approaches by tuning the a priori probability distribution from uniform, $p(\varphi) = 1/2\pi$, to peaked $p(\varphi) \approx \delta(\varphi - \varphi_0)$ (see [23] which is one of very few papers dealing with phase estimation in the intermediate regime).

Let $|n\rangle \in \mathcal{H}$, $n = 0 \dots N$, denote an eigenbasis of the phase shift operator U_φ, such that $U_\varphi|n\rangle = \exp(-in\varphi)|n\rangle$. Thinking in terms of an N photon state inside a Mach-Zehnder interferometer, $|n\rangle$ denotes a state in which n photons travel through the upper and $N-n$ photons through the lower arm, while the upper arm is delayed by φ with respect to the lower one. Let $\rho = \sum_{ij} c_{ij}|i\rangle\langle j|$ be the probe state, which under the phase shift becomes $\rho_\varphi = U_\varphi \rho U_\varphi^\dagger$. Note, that ρ may represent a decohered pure probe state, provided the decoherence process commutes with the phase shift operation, which is the case for the most relevant decoherence mechanisms such as the photon loss in optical [24] and dephasing or damping in atomic systems [11].

The estimation strategy is defined by a POVM measurement $\Pi_{\tilde{\varphi}}$, $\int_{-\pi}^{\pi} \mathrm{d}\tilde{\varphi}\, \Pi_{\tilde{\varphi}} = \mathbb{1}$, $\Pi_{\tilde{\varphi}} \geq 0$, where the measurement outcome $\tilde{\varphi}$ is at the same time the estimator of the phase. Hence, $\mathrm{Tr}(\rho_\varphi \Pi_{\tilde{\varphi}})$ is the conditional probability of estimating $\tilde{\varphi}$ provided the true phase is φ. Taking into account the a priori probability distribution $p(\varphi)$, the average cost of the estimation strategy reads:

$$\overline{C} = \int_{-\pi}^{\pi} \mathrm{d}\varphi \mathrm{d}\tilde{\varphi}\, p(\varphi) \mathrm{Tr}(\rho_\varphi \Pi_{\tilde{\varphi}}) C_{\varphi,\tilde{\varphi}}, \tag{2}$$

where $C_{\varphi,\tilde{\varphi}}$ is the cost for guessing $\tilde{\varphi}$, while the true value was φ. The $C_{\varphi,\tilde{\varphi}} = (\varphi - \tilde{\varphi})^2$ cost function is only appropriate for narrow distribution around $\varphi = 0$ since it does not respect periodic conditions $C_{\varphi+2\pi,\tilde{\varphi}}$. In what follows we choose the cost function $C_{\varphi,\tilde{\varphi}} = 4\sin^2 \frac{\varphi - \tilde{\varphi}}{2}$ and denote its average by $\widetilde{\delta^2\varphi}$, as it is the simplest cost function approximating the variance for narrow distributions [25].

Consider an auxiliary two dimensional Hilbert space $\mathcal{H}_A$, where we define $|\varphi\rangle = (|0\rangle + \exp(-i\varphi)|1\rangle)/\sqrt{2}$. Notice that we can rewrite the cost function as

$$C_{\varphi,\tilde{\varphi}} = 4\left[1 - \mathrm{Tr}\left(|\varphi\rangle\langle\varphi||\tilde{\varphi}\rangle\langle\tilde{\varphi}|\right)\right]. \tag{3}$$

Substituting the above formula to Eq. (2), making use of the fact that $\mathrm{Tr}(AB)\mathrm{Tr}(CD) = \mathrm{Tr}(A \otimes C\; B \otimes D)$ and inserting both integrals under the trace we finally arrive at:

$$\widetilde{\delta^2\varphi} = \overline{C} = 4\left(1 - F\right), \quad F = \mathrm{Tr}\left(RM\right) \tag{4}$$

where

$$R = \int_{-\pi}^{\pi} \mathrm{d}\varphi\, p(\varphi) |\varphi\rangle\langle\varphi| \otimes \rho_\varphi, \tag{5}$$

$$M = \int_{-\pi}^{\pi} \mathrm{d}\tilde{\varphi}\, |\tilde{\varphi}\rangle\langle\tilde{\varphi}| \otimes \Pi_{\tilde{\varphi}}. \tag{6}$$

Assuming the probe state ρ is given, the problem of finding the optimal estimation strategy amounts to maximizing the fidelity F, over operators M which are of the form (6). The structure of M is analogous to the structure of a separable state, with an additional constraint resulting from the POVM completeness condition: $\mathrm{Tr}_A M = \int_{-\pi}^{\pi} \mathrm{d}\varphi \Pi_\varphi = \mathbb{1}$. This observation has been employed in [26] to cast the problem of optimization of state estimation into a maximization of linear functional over separable states, which in principle may be solved numerically using semi-definite programming. Surprisingly, as demonstrated below, the solution to the problem considered in this paper is found analytically, without resorting to semi-definite programming, and just with the help of simple positive partial transpose (PPT) necessary criterion for separability [27].

Let us write the a priori distribution as $p(\varphi) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} p_k \exp(ik\varphi)$ where $p_i^* = p_{-i}$ guarantees $p(\varphi) \in \mathbb{R}$ and $p_0 = 1$ assures normalization. Both R and M may be written using a block form respecting the structure of the tensor product $\mathcal{H}_A \otimes \mathcal{H}$:

$$R = \begin{pmatrix} R_0^0 & R_1^0 \\ R_0^1 & R_1^1 \end{pmatrix}, \; M = \begin{pmatrix} M_0^0 & M_1^0 \\ M_0^1 & M_1^1 \end{pmatrix} \tag{7}$$

where $R_j^i = {}_A\langle i|R|j\rangle_A$, $M_j^i = {}_A\langle i|M|j\rangle_A$. Hermicity implies $R_j^i = R_i^{j\dagger}$, $M_j^i = M_i^{j\dagger}$. Straightforward calculation of R_j^i using Eq. (5) yields $R_j^i = \frac{1}{2}\rho \odot P_j^i$, where $\odot$ denotes the entry-wise matrix product (Hadamard product) with

$$(P_0^0)_l^k = (P_1^1)_l^k = p_{k-l}, \; (P_1^0)_l^k = (P_0^1)_k^{l*} = p_{k-l+1}. \tag{8}$$

The completeness constraint $\mathrm{Tr}_A M = \mathbb{1}$, implies $M_0^0 + M_1^1 = \mathbb{1}$. Using the fact that $R_0^0 = R_1^1$ and $\mathrm{Tr}\rho \odot P_k^k = 1$, we get $F = \frac{1}{2} + \Re\left[\mathrm{Tr}(R_0^1 M_1^0)\right]$.

Main result. The cost of the optimal phase estimation strategy reads

$$\widetilde{\delta^2\varphi} = 4\left(\tfrac{1}{2} - \|R_0^1\|_1\right), \tag{9}$$

where $\|A\|_1$ denotes the trace norm of a matrix. The optimal estimation strategy itself is given by

$$M = \sum_{k=0}^{N} |\varphi_k\rangle\langle\varphi_k| \otimes |\psi_k\rangle\langle\psi_k| \tag{10}$$

where $\exp(-i\varphi_k)$, $|\psi_k\rangle$ are eigenvalues and eigenvectors of $U = V_R U_R^\dagger$, while V_R, U_R are unitaries appearing in the singular value decomposition (SVD) $R_0^1 = U_R \Lambda_R V_R^\dagger$. Notice that Eq. (9) provides us with an explicit recipe for the practical implementation of the optimal estimation strategy. The optimal measurement is a projective measurement on basis $|\psi_k\rangle$,

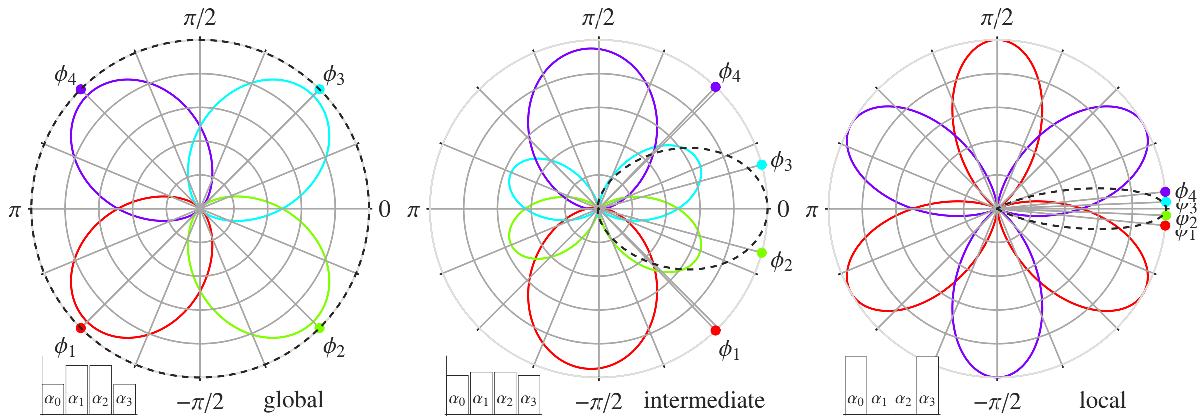

Figure 1. (color online) Optimal phase estimation strategies for $N = 3$ photon probe states $|\psi\rangle$ and different degrees of a priori knowledge $p_t(\varphi)$ (black, dashed) [see Eq. (13)] for $t = 20$ (global regime), $t = 0.2$ (intermediate regime), $t = 0.02$ (local regime) respectively. ϕ_k is a phase that is estimated once a measurement result k is obtained, while the corresponding curve depicts $p(k|\phi) = |\langle\psi_k|U_\varphi|\psi\rangle|^2$ — conditional probability that measurement outcome k is obtained if the true phase is φ. Insets in the bottom left corners illustrate parameters α_k of the optimal probe state. In the local regime only two measurement outcomes are relevant and the optimal state is the N00N state — a result known from the quantum Fisher information approach.

with φ_k being the estimated phase given a measurement result k.

Proof. Let M' represent the optimal measurement strategy, and define M'' such that off-diagonal blocks remain the same while the diagonal blocks are interchanged: $M''^0_0 = M'^1_1$, $M''^1_1 = M'^0_0$. M' is of course positive semi-definite, but this is not guaranteed for M''. Notice, however, that $M'' = WM'^{T_A}W^\dagger$, where $W = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}$ is a unitary operation and T_A is the partial transposition with respect to the A subsystem. Since M' is separable, then by the PPT criterion $M'^{T_A} \geq 0$, and hence also $M'' \geq 0$.

F depends only on the off-diagonal blocks via $\Re\left[\mathrm{Tr}(R^1_0 M^0_1)\right]$. Therefore, M'' provides the same fidelity as M'. Keeping fidelity the same we construct

$$M = (M' + M'')/2 = \begin{pmatrix} \mathbb{1}/2 & M'^0_1 \\ M'^{0\dagger}_1 & \mathbb{1}/2 \end{pmatrix}. \qquad (11)$$

Positive semi-definiteness of M implies that $M'^{0\dagger}_1 M'^0_1 \leq \mathbb{1}/4$, hence all singular values of M'^0_1 are no greater than 1/2.

Let $R^1_0 = U_R\Lambda_R V^\dagger_R$, $M^0_1 = U_M\Lambda_M V^\dagger_M$ be SVDs. Thanks to $\Lambda_M \leq \mathbb{1}/2$, and the fact that $|\mathrm{Tr}\,(U_1\Lambda_1U_2\Lambda_2)| \leq \mathrm{Tr}(\Lambda_1\Lambda_2)$ the following chain of inequalities holds:

$$\Re\left[\mathrm{Tr}\left(R^1_0 M^0_1\right)\right] \leq \left|\mathrm{Tr}\left(U_R\Lambda_R V^\dagger_R U_M\Lambda_M V^\dagger_M\right)\right| \leq$$
$$\leq \mathrm{Tr}\left(\Lambda_R\Lambda_M\right) \leq \frac{1}{2}\mathrm{Tr}\left(\Lambda_R\right) = \frac{1}{2}\|R^1_0\|_1. \qquad (12)$$

All inequalities are saturated for $M^0_1 = U/2$, where $U = V_R U^\dagger_R$. Using the eigen decomposition $U = \sum_{k=0}^{N}\exp(-i\varphi_k)|\psi_k\rangle\langle\psi_k|$ we arrive at Eq. (10). ■

A natural prior $p(\varphi)$ that can be tuned from uniform to $\delta(\varphi)$ distribution is a solution of the diffusive evolution on the $[-\pi,\pi)$ interval with periodic boundary conditions and initial $\delta(\varphi)$ distribution:

$$p_t(\varphi) = \frac{1}{2\pi}\left(1 + 2\sum_{n=1}^{\infty}\cos(n\varphi)e^{-n^2t}\right). \qquad (13)$$

With this choice $(P^1_0)^k_l = \exp[-(l-k+1)^2t]$.

For the sake of clarity, in the following examples we consider only the decoherence-free phase estimation, where the probe state is pure $\rho = |\psi\rangle\langle\psi|$, $|\psi\rangle = \sum_{n=0}^{N}\alpha_n|n\rangle$ (as stressed before the method can be applied to more general scenarios). In this case, $2(R^1_0)^k_l = \alpha_k\alpha^*_l\exp[-(l-k+1)^2t]$, and the optimal probe state corresponds to the choice of α_k such that the trace norm of the above matrix is maximized. For $t = \infty$ (global approach), $2(R^1_0)^k_l = \alpha_k\alpha_{k-1}\delta_{k-1,l}$, $2\|R^1_0\|_1 = \sum_{k=0}^{N-1}|\alpha_k||\alpha_{k+1}|$

which is maximized by $\alpha_k = \sqrt{2/(N+2)}\sin[(k+1)\pi/(N+2)]$ which are the Berry-Wiseman (BW) states [25]. For general t there is no analytical formula for the trace norm as a function of α_k and one needs to maximize the trace norm numerically [28].

The structure of the optimal estimation strategy and the optimal probe state is depicted in Fig. 1 for $N = 3$ and three different a priori distributions. Notice, how, with the increasing a priori knowledge, the optimal probe state evolves to the N00N state [1], which is the optimal solution of the QFI approach. The periodic $(2\pi/N)$ structure of the conditional probabilities $p(k|\varphi)$, visible in the local regime, clearly reminds of the fact that this estimation strategy is useless unless the prior is highly peaked since otherwise there is strong ambiguity in using the measurement result to estimate the phase.

It is worth mentioning, that although both the local approach discussed in this paper (corresponding to the limit $t \to 0$), and the QFI approach yield the same optimal probe states and the same optimal measurements, they in general yield different estimation precisions. While F_Q is an extremely useful tool, it just provides a CR bound on the achievable estimation precision $\delta\varphi \geq 1/\sqrt{F_Q}$. This bound, in general, cannot be saturated by an estimator based on the results of a single measurement. Only in the asymptotic limit of infinitely many repetitions of the experiment one may construct a max-likelihood estimator which saturates the bound. The approach presented in this paper, on the other hand, gives an operationally meaningful answer for single shot estimation procedure. Moreover, when $t \to 0$, then $\delta_\varphi \to 0$ by the obvious fact that that the phase is known perfectly. Hence, for t small enough the CR bound is violated (this is no contradiction since our estimator is not locally unbiased [29]).

Performance of the optimal estimation strategy for $N = 10$ and different probe states is depicted in Fig. 2 as a function of a priori uncertainty $\delta\varphi_{\mathrm{prior}} = \sqrt{\int \mathrm{d}\varphi\, 4\sin^2(\varphi/2)p_t(\varphi)}$. N00N states are optimal up to a threshold (which scales as $1/N$) above which they become useless due to $2\pi/N$ phase estimation ambiguity. The optimal states clearly demonstrate their superiority over the BW states for moderate $\delta\varphi_{\mathrm{prior}}$ and lose their advantage with the increasing prior ignorance on the phase value.

In summary, the problem of optimal phase estimation with arbitrary a priori knowledge has been solved analytically, allowing to investigate the regime of estimation not accessible via neither the QFI approach nor covariant measurements. Optimal measurements and estimators have been explicitly constructed so that it is immediate to apply the results to any practical phase estimation problem.

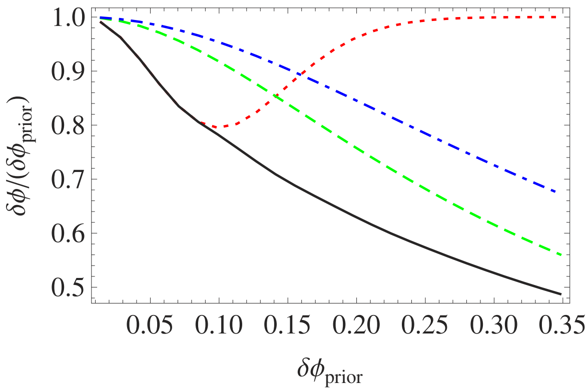

Figure 2. (color online) Relative reduction of the uncertainty $\delta\varphi$ with respect to $\delta\varphi_{\mathrm{prior}}$ after the optimal estimation procedure is applied to the $N = 10$ optimal probe state (black, solid); N00N state (optimal in the local regime) (red, dotted); BW state (optimal in the global regime) (green, dashed); „classical" state with $\alpha_k^2 = \binom{N}{k}$ (blue, dash-dotted).

This research was supported by the European Commission under the Integrating Project Q-ESSENCE and the Foundation for Polish Science under the TEAM programme.

[1] J. J. . Bollinger, W. M. Itano, D. J. Wineland, and D. J. Heinzen, Phys. Rev. A, **54**, R4649 (1996).
[2] J. P. Dowling, Phys, Rev. A, **57**, 4736 (1998).
[3] V. Giovannetti, S. Lloyd, and L. Maccone, Physical Review Letters, **96**, 010401 (2006).
[4] K. Banaszek, R. Demkowicz-Dobrzanski, and I. A. Walmsley, Nature Photonics, **3**, 673 (2009).
[5] C. M. Caves, Phys. Rev. D, **23**, 1693 (1981).
[6] M. Xiao, L.-A. Wu, and H. J. Kimble, Phys. Rev. Lett., **59**, 278 (1987).
[7] M. J. Holland and K. Burnett, Phys. Rev. Lett., **71**, 1355 (1993).
[8] L. Pezzé and A. Smerzi, Phys. Rev. Lett., **100**, 073601 (2008).
[9] B. Yurke, S. L. McCall, and J. R. Klauder, Phys. Rev. A, **33**, 4033 (1986).
[10] V. Giovannetti, S. Lloyd, and L. Maccone, Science, **306**, 1330 (2004).
[11] S. F. Huelga, C. Macchiavello, T. Pellizzari, A. K. Ekert, M. B. Plenio, and J. I. Cirac, Phys. Rev. Lett., **79**, 3865 (1997).

[12] M. A. Rubin and S. Kaushik, Phys. Rev. A, **75**, 053805 (2007).
[13] G. Gilbert, M. Hamrick, and Y. S. Weinstein, J. Opt. Soc. Am B, **25**, 1336 (2008).
[14] S. D. Huver, C. F. Wildfeuer, and J. P. Dowling, Phys. Rev. A, **78**, 063828 (2008).
[15] U. Dorner, R. Demkowicz-Dobrzanski, B. J. Smith, J. S. Lundeen, W. Wasilewski, K. Banaszek, and I. A. Walmsley, Phys. Rev. Lett., **102**, 040403 (2009).
[16] J. Chwedeńczuk, F. Piazza, and A. Smerzi, Phys. Rev. A, **82**, 051601 (2010).
[17] S. Knysh, V. N. Smelyanskiy, and G. A. Durkin, arXiv:1006.1645 (2010).
[18] J. Kolodynski and R. Demkowicz-Dobrzanski, Phys. Rev. A, **82**, 053804 (2010).
[19] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett., **72**, 3439 (1994).
[20] O. E. Barndorff-Nielsen and R. D. Gill, J.Phys.A, **30**, 4481 (2000).
[21] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North Holland, Amsterdam, 1982).
[22] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, Rev. Mod. Phys., **79**, 555 (2007).
[23] G. A. Durkin and J. P. Dowling, Phys. Rev. Lett., **99**, 070801 (2007).
[24] R. Demkowicz-Dobrzanski, U. Dorner, B. J. Smith, J. S. Lundeen, W. Wasilewski, K. Banaszek, and I. A. Walmsley, Phys. Rev. A, **80**, 013825 (2009).
[25] D. W. Berry and H. M. Wiseman, Phys. Rev. Lett., **85**, 5098 (2000).
[26] M. Navascués, Phys. Rev. Lett., **100**, 070503 (2008).
[27] A. Peres, Phys. Rev. Lett., **77**, 1413 (1996).
[28] To facilitate the numerical search for the optimal probe state, one may implement an iterative procedure. In the i-th step starting with a probe state $\psi^{(i)}$, one calculates the optimal estimation strategy $M^{(i)}$ — as shown in the paper this requires only a single run of SVD algorithm. Having found $M^{(i)}$, one looks for the optimal probe state for this particular measurement. Be rewriting Eq. (4) it can be shown that this problem is equivalent to performing eigenvalue decomposition of a certain matrix and choosing eigenvector corresponding to the maximal eigenvalue, which should be used as $\psi^{(i+1)}$ in the next iteration.
[29] C. W. Helstrom, *Quantum detection and estimation theory* (Academic press, 1976).