PHYSICAL REVIEW A **94**, 022334 (2016)

Optimizing qubit phase estimation

François Chapeau-Blondeau
Laboratoire Angevin de Recherche en Ingénierie des Systèmes (LARIS), Université d'Angers,
62 avenue Notre Dame du Lac, 49000 Angers, France

(Received 5 June 2016; revised manuscript received 2 August 2016; published 24 August 2016)

The theory of quantum state estimation is exploited here to investigate the most efficient strategies for this task, especially targeting a complete picture identifying optimal conditions in terms of Fisher information, quantum measurement, and associated estimator. The approach is specified to estimation of the phase of a qubit in a rotation around an arbitrary given axis, equivalent to estimating the phase of an arbitrary single-qubit quantum gate, both in noise-free and then in noisy conditions. In noise-free conditions, we establish the possibility of defining an optimal quantum probe, optimal quantum measurement, and optimal estimator together capable of achieving the ultimate best performance uniformly for any unknown phase. With arbitrary quantum noise, we show that in general the optimal solutions are phase dependent and require adaptive techniques for practical implementation. However, for the important case of the depolarizing noise, we again establish the possibility of a quantum probe, quantum measurement, and estimator uniformly optimal for any unknown phase. In this way, for qubit phase estimation, without and then with quantum noise, we characterize the phase-independent optimal solutions when they generally exist, and also identify the complementary conditions where the optimal solutions are phase dependent and only adaptively implementable.

DOI: 10.1103/PhysRevA.94.022334

I. INTRODUCTION

To efficiently infer information from measurements, the theory of quantum estimation enables one to characterize the overall best performance in estimating parameters defining a quantum state from quantum measurements on this state. Since its introduction in Refs. [1–3], this theory of quantum estimation has received a significant number of applications, extensions, experimental implementations, with many possible variations according to the conditions and quantum systems undergoing estimation [4–6]. Several related aspects of an estimation task can be addressed through this theory. The theory, by means of the quantum Fisher information, places an upper bound on the classical Fisher information, and in turn the classical Fisher information determines a lower bound for the mean-squared error of any conceivable estimator. Once the bounds are established, their conditions of attainability can also be investigated. It turns out that, depending on the conditions, the best performances dictated by the bounds are not always attainable [5–7]. This is especially true if attention is paid to practical realizability which restricts one to parameter-independent optimal solutions: since the parameters are unknown, if they enter in the formulation of the optimal solutions, then these solutions are inaccessible for practical implementation. Optimal quantum measurement protocols and associated optimal estimators can be investigated to reach or best approach these bounds. Some studies have concentrated on expressing the ultimate best performance in definite conditions of estimation [7–11]; others concentrated more on devising efficient quantum measurement protocols and associated estimators [12–16]. Here, we will consistently address these successive stages in conjunction, for optimizing an estimation task.

Most of the studies so far have considered estimation on a generic quantum state generally represented by a density operator carrying the dependence with the unknown parameters of interest. A more realistic approach, especially natural in a quantum signal or information processing perspective, is to consider that estimation has to be performed from a noisy quantum state altered by decoherence. The initial quantum state carrying the dependence with the unknown parameters is subsequently affected by some specified quantum noise, before it becomes accessible for estimation. Such a more realistic scenario allows one to explicitly examine the impact of the noise or decoherence on the optimal estimation strategies and their performance. Such an approach of estimation from a noisy quantum state has recently been considered in [17–20], with the quantum Fisher information to assess the performance which is maximized through selection of the probe state best resistant to the noise, for several noise models.

We will especially consider such conditions with noise here. We address estimation of the phase of a qubit in a rotation around an arbitrary axis, equivalent to estimating the phase of an arbitrary single-qubit quantum gate, and generalizing the rotations of Refs. [17,18]. The ultimate best performance for phase estimation is characterized through explicit evaluation of the quantum Fisher information, in the presence of arbitrary quantum noise on the qubit. We examine the successive stages to be implemented for phase estimation, and their optimization towards realizing the ultimate best performance. In this way, optimization of the quantum probe carrying the phase dependence of the quantum measurement protocol and of the optimal estimator, are successively addressed. We especially examine the possibility of determining phase-independent optimal solutions at each of these stages, so as to enable direct practical implementations. According to the noise conditions, we investigate, consistently for the quantum probe, the quantum measurement and the estimator, whether there exist phase-independent optimal solutions, and when they are feasible we exhibit the forms of such phase-independent solutions that best approach the ultimate maximal performance in qubit phase estimation.

In the following, in Sec. II we briefly review elements of quantum estimation theory providing a general characterization of the ultimate best performance via the quantum Fisher

 ©2016 American Physical Society

information. Section III specializes to phase estimation on a qubit, first with no noise, where are successively established the optimal probe state, quantum measurement, and estimator together achieving the ultimate best estimation performance, with a phase-independent solution. Section IV addresses the same phase estimation task, yet from a qubit affected by an arbitrary quantum noise, and where also the possibility of phase-independent optimal solutions are investigated while targeting maximal performance in estimation.

II. GENERAL PERFORMANCE IN QUANTUM ESTIMATION

A. Quantum score and Fisher information

A quantum system with D-dimensional Hilbert space $\mathcal{H}_D$ has its state represented by a density operator ρ_ξ dependent upon an unknown scalar parameter ξ to be estimated. A quantum measurement protocol is implemented on the state ρ_ξ, and the measurement outcomes are processed through an estimator $\widehat{\xi}$, in order to infer a value for the unknown parameter ξ. To assess the performance in such an estimation task, a meaningful metric is the mean-squared estimation error $\langle(\widehat{\xi}-\xi)^2\rangle$. From classical estimation theory [21,22], it is known that any conceivable estimator $\widehat{\xi}$ for ξ is endowed with a mean-squared error $\langle(\widehat{\xi}-\xi)^2\rangle$ which is lower bounded by the Cramér-Rao bound involving the reciprocal of the classical Fisher information $F_c(\xi)$. Estimators are known, such as the maximum likelihood estimator, that allow us to reach the Cramér-Rao bound in definite (usually asymptotic) conditions [22]. Higher Fisher information $F_c(\xi)$ generally entails better performance in estimation, and one has then the faculty to devise the quantum measurement protocol so as to maximize $F_c(\xi)$. In this respect, there is a fundamental upper bound [1,23] provided by the quantum Fisher information $F_q(\xi)$ which sets a limit to the classical Fisher information $F_c(\xi)$, i.e., fixing $F_c(\xi)\leqslant F_q(\xi)$.

The quantum Fisher information $F_q(\xi)$ for estimating the scalar parameter ξ from the quantum state ρ_ξ is defined as the mean-squared quantum score, i.e. [2,3,6,23],

$$F_q(\xi)=\langle L_\xi^2\rangle=\mathrm{tr}\left(\rho_\xi L_\xi^2\right). \tag{1}$$

The quantum score L_ξ, or symmetrized logarithmic derivative, is a Hermitian operator defined by the equation [1–3,23]

$$\partial_\xi\rho_\xi\equiv\frac{\partial\rho_\xi}{\partial\xi}=\frac{1}{2}(L_\xi\rho_\xi+\rho_\xi L_\xi). \tag{2}$$

By referring to the spectral decomposition of ρ_ξ in its orthonormal eigenbasis $\rho_\xi=\sum_{n=1}^D\lambda_n|\lambda_n\rangle\langle\lambda_n|$, it is possible to obtain a more explicit expression for the quantum score as [6,20]

$$L_\xi=2\sum_{m,n}\frac{1}{\lambda_m+\lambda_n}|\lambda_m\rangle\langle\lambda_m|\partial_\xi\rho_\xi|\lambda_n\rangle\langle\lambda_n|, \tag{3}$$

where the sums as in Eq. (3) include all terms corresponding to eigenvalues $\lambda_m+\lambda_n\neq 0$. This leads, from Eq. (1), to the quantum Fisher information [3,6,20]

$$F_q(\xi)=2\sum_{m,n}\frac{|\langle\lambda_m|\partial_\xi\rho_\xi|\lambda_n\rangle|^2}{\lambda_m+\lambda_n}. \tag{4}$$

For the special case of a pure state $\rho_\xi=|\lambda\rangle\langle\lambda|$ then $\partial_\xi\rho_\xi=|\partial_\xi\lambda\rangle\langle\lambda|+|\lambda\rangle\langle\partial_\xi\lambda|$, and also the expressions of Eqs. (3) and (4) are replaced [6,20] by

$$L_\xi=2(|\partial_\xi\lambda\rangle\langle\lambda|+|\lambda\rangle\langle\partial_\xi\lambda|), \tag{5}$$

and

$$F_q(\xi)=4(\langle\partial_\xi\lambda|\partial_\xi\lambda\rangle+\langle\partial_\xi\lambda|\lambda\rangle^2). \tag{6}$$

For further determination of the quantum score L_ξ and quantum Fisher information $F_q(\xi)$, we now examine a specific and useful form for the parametric dependence with ξ determining the derivative ∂_ξ.

B. Estimation in a unitary family

An important parametric family [2,6] of quantum states ρ_ξ arises when a quantum probe state ρ_0 experiences an arbitrary unitary transformation U_ξ to yield the quantum state $\rho_\xi=U_\xi\rho_0U_\xi^\dagger$, with the parametrization

$$U_\xi=\exp(-i\xi G), \tag{7}$$

where G is an arbitrary Hermitian operator forming the generator of the unitary U_ξ, with G and U_ξ which commute since they are diagonal in the same orthonormal basis. For this type of parametric dependence of ρ_ξ, one obtains $\partial_\xi\rho_\xi=i[\rho_\xi,G]=iU_\xi[\rho_0,G]U_\xi^\dagger$, leading for the quantum score of Eq. (3) to $L_\xi=U_\xi L_0U_\xi^\dagger$, with

$$L_0=2i\sum_{m,n}\frac{1}{\lambda_m+\lambda_n}\left|\lambda_m^0\right\rangle\left\langle\lambda_m^0\right|[\rho_0,G]\left|\lambda_n^0\right\rangle\left\langle\lambda_n^0\right|. \tag{8}$$

For Eq. (8), when the transformed state $\rho_\xi=U_\xi\rho_0U_\xi^\dagger$ has eigenvalues λ_n and eigenvectors $|\lambda_n\rangle$, the initial probe state ρ_0 has same eigenvalues λ_n and eigenvectors $|\lambda_n^0\rangle$ with $|\lambda_n\rangle=U_\xi|\lambda_n^0\rangle$. The quantum Fisher information of Eq. (1) is equally $F_q(\xi)=\mathrm{tr}(\rho_0 L_0^2)$ expressible as

$$F_q(\xi)=2\sum_{m,n}\frac{(\lambda_m-\lambda_n)^2}{\lambda_m+\lambda_n}\left|\left\langle\lambda_m^0\right|G\left|\lambda_n^0\right\rangle\right|^2. \tag{9}$$

For the special case of a pure state $|\lambda\rangle=U_\xi|\lambda^0\rangle$, Eqs. (5) and (6) lead, with $\rho_0=|\lambda^0\rangle\langle\lambda^0|$, to

$$L_0=2i[\rho_0,G], \tag{10}$$

and

$$F_q(\xi)=4\left(\langle\lambda^0|G^2|\lambda^0\rangle-\langle\lambda^0|G|\lambda^0\rangle^2\right)=4\langle\lambda^0|\Delta G^2|\lambda^0\rangle. \tag{11}$$

These elements of quantum estimation theory are now applied to estimation from a qubit state in the two-dimensional Hilbert space $\mathcal{H}_2$.

III. OPTIMAL QUBIT PHASE ESTIMATION

For parametric estimation on a qubit, it is possible to come up with explicit expressions for the quantum score and quantum Fisher information fixing the ultimate maximal performance, and also to devise optimal strategies achieving this maximal performance, as we now address.

A. Optimizing the quantum Fisher information

For qubit states in $\mathcal{H}_2$, a convenient representation allowing an insightful geometric picture is the Bloch representation [24], where the density operators are expressed in the basis of the four Pauli operators $\{I_2,\sigma_x,\sigma_y,\sigma_z\}$. In Bloch representation the quantum state ρ_ξ is generally expressed as

$$\rho_\xi = \tfrac{1}{2}(I_2 + \vec{r}_\xi \cdot \vec{\sigma}), \tag{12}$$

with I_2 the identity of $\mathcal{H}_2$, and $\vec{\sigma}$ a formal vector assembling the three 2×2 (traceless Hermitian unitary) Pauli matrices $[\sigma_x,\sigma_y,\sigma_z] = \vec{\sigma}$. The coordinates of ρ_ξ are specified by the Bloch vector $\vec{r}_\xi$ in $\mathbb{R}^3$, with norm $\|\vec{r}_\xi\| = 1$ for a pure state and $\|\vec{r}_\xi\| < 1$ for a mixed state. The qubit state ρ_ξ of Eq. (12) has the two eigenvalues $\lambda_\pm = (1 \pm \|\vec{r}_\xi\|)/2$ and normalized eigenvectors $|\lambda_\pm\rangle$, the two projectors on these eigenvectors having the Bloch representation $|\lambda_\pm\rangle\langle\lambda_\pm| = (I_2 \pm \vec{r}_\xi \cdot \vec{\sigma}/\|\vec{r}_\xi\|)/2$.

When estimation is performed from the qubit state ρ_ξ characterized by the Bloch vector $\vec{r}_\xi$, for any arbitrary dependence on ξ, as derived in [20], the quantum score of Eq. (3) follows as

$$L_\xi = -\frac{\vec{r}_\xi \partial_\xi \vec{r}_\xi}{1-\vec{r}_\xi^{\,2}} I_2 + \left(\frac{\vec{r}_\xi \partial_\xi \vec{r}_\xi}{1-\vec{r}_\xi^{\,2}}\vec{r}_\xi + \partial_\xi \vec{r}_\xi\right)\cdot\vec{\sigma}, \tag{13}$$

and the quantum Fisher information of Eq. (4) as

$$F_q(\xi) = \frac{(\vec{r}_\xi\, \partial_\xi \vec{r}_\xi)^2}{1-\vec{r}_\xi^{\,2}} + (\partial_\xi \vec{r}_\xi)^2, \tag{14}$$

for the general case of a mixed state ρ_ξ. Meanwhile, for the special case of a pure state ρ_ξ, Eq. (5) becomes

$$L_\xi = \partial_\xi \vec{r}_\xi \cdot \vec{\sigma}, \tag{15}$$

and Eq. (6),

$$F_q(\xi) = (\partial_\xi \vec{r}_\xi)^2. \tag{16}$$

For estimation in a unitary family, as in Sec. II B, we consider a qubit, representing the probe, which is prepared in a quantum state ρ_0 characterized by the Bloch vector $\vec{r}_0$. An arbitrary unitary transformation on the qubit (an arbitrary single-qubit quantum gate) can be expressed [24] in the form $U = \exp(i\gamma)\exp(-i\xi\vec{n}\cdot\vec{\sigma}/2)$; and since the overall scalar phase γ is unimportant here, we consider the general unitary transformation on the qubit as

$$U_\xi = \exp\left(-i\frac{\xi}{2}\vec{n}\cdot\vec{\sigma}\right), \tag{17}$$

where $\vec{n} = [n_x,n_y,n_z]^\top$ is a real unit vector of $\mathbb{R}^3$, equivalent in Eq. (7) with a Hermitian operator $G = \vec{n}\cdot\vec{\sigma}/2$. The transformation of Eq. (17) acts on the input probe state ρ_0 so as to yield the qubit state $\rho_\xi = U_\xi \rho_0 U_\xi^\dagger$ (especially the presence of a scalar phase γ in U would have had no effect on ρ_ξ). As a result, the transformed qubit state ρ_ξ is characterized by the Bloch vector $\vec{r}_\xi$ which geometrically in $\mathbb{R}^3$ represents the input Bloch vector $\vec{r}_0$ rotated by the angle ξ around the axis $\vec{n}$.

In Bloch representation, one can deduce the commutator $[\rho_0,G] = -i(\vec{n}\times\vec{r}_0)\cdot\vec{\sigma}/2$, and then evaluate Eq. (8) as the compact expression

$$L_0 = (\vec{n}\times\vec{r}_0)\cdot\vec{\sigma}, \tag{18}$$

which also holds when ρ_0 is a pure state ruled by Eq. (10). And next, the quantum Fisher information of Eq. (9) evaluates to

$$F_q(\xi) = (\vec{n}\times\vec{r}_0)^2, \tag{19}$$

which also holds when ρ_0 is a pure state ruled by Eq. (11).

Alternatively, for $\vec{r}_0$ in $\mathbb{R}^3$, it can always be written $\vec{r}_0 = r_{0\parallel}\vec{n} + r_{0\perp}\vec{n}_\perp$ defining the unit vector $\vec{n}_\perp$ orthogonal to $\vec{n}$. The rotated Bloch vector is then $\vec{r}_\xi = r_{0\parallel}\vec{n} + r_{0\perp}\cos(\xi)\vec{n}_\perp + r_{0\perp}\sin(\xi)\vec{n}\,'_\perp$ with a third orthogonal unit vector $\vec{n}\,'_\perp = \vec{n}\times\vec{n}_\perp$. The derivative follows as $\partial_\xi\vec{r}_\xi = -r_{0\perp}\sin(\xi)\vec{n}_\perp + r_{0\perp}\cos(\xi)\vec{n}\,'_\perp = \vec{n}\times\vec{r}_\xi$, offering an intrinsic geometric characterization of the derivative

$$\partial_\xi\vec{r}_\xi = \vec{n}\times\vec{r}_\xi, \tag{20}$$

which entails the vanishing inner product $\vec{r}_\xi\,\partial_\xi\vec{r}_\xi = 0$. It then follows in Eqs. (13) and (15) equally that the quantum score is

$$L_\xi = (\vec{n}\times\vec{r}_\xi)\cdot\vec{\sigma}, \tag{21}$$

and it also follows in Eqs. (14) and (16) equally that $F_q(\xi) = (\vec{n}\times\vec{r}_\xi)^2$ which is similar to Eq. (19) due to the geometry of $\vec{r}_0$ and $\vec{r}_\xi$.

For estimating the phase ξ of Eq. (17) on the qubit, it is remarkable that Eqs. (18), (19), and (21) provide very concise expressions for the quantum score and Fisher information, with insightful geometric formulations, under a form that we did not find previously in the literature. The expressions for the quantum score and Fisher information of Eqs. (13)–(16) are general for parameter estimation on a qubit and were derived in detail in [20]. For phase estimation, these equations are transformed here to obtain the simple geometric forms of Eqs. (18), (19), and (21) which are not in [20]. It is for instance interesting to observe from Eq. (19) that the quantum Fisher information $F_q(\xi)$ depends on the properties of the quantum states only through the probe ρ_0 via its Bloch vector $\vec{r}_0$, and is independent of the phase ξ. Equation (19) renders limpid the issue of optimizing the input probe ρ_0 in order to maximize the quantum Fisher information $F_q(\xi)$. From Eq. (19), to maximize $F_q(\xi)$, the optimal input probe ρ_0 has to be chosen as a pure state (i.e., satisfying $\|\vec{r}_0\| = 1$), with a unit Bloch vector $\vec{r}_0$ orthogonal to the rotation axis $\vec{n}$, to achieve the overall maximum of $F_q(\xi)$ which is $F_q^{\max} = 1$, uniformly for any ξ. The next step then, is to devise a quantum measurement protocol capable of reaching this maximal Fisher information $F_q^{\max} = 1$.

B. Optimizing the quantum measurement

We envisage a generalized quantum measurement on the qubit, by means of a positive operator valued measure (POVM) [24]. For the qubit, a measurement operator indexed by k can be expressed in Bloch representation as $M_k = b_k I_2 + \vec{a}_k\cdot\vec{\sigma}$. The determinant is $\det(M_k) = b_k^2 - \vec{a}_k^{\,2}$, and M_k has the two eigenvalues $b_k \pm \sqrt{\vec{a}_k^{\,2}}$ with the two projectors in $\mathcal{H}_2$ on the two associated eigenvectors reading $(I_2 \pm \vec{a}_k \cdot \vec{\sigma}/\sqrt{\vec{a}_k^{\,2}})/2$. Such an M_k is Hermitian if and only if $(b_k,\vec{a}_k)$ are real, giving $\sqrt{\vec{a}_k^{\,2}} = \|\vec{a}_k\|$ real; moreover it satisfies $0 \leqslant M_k$ if and only if $0 \leqslant b_k - \|\vec{a}_k\|$, and $M_k \leqslant I_2$ if and only if $b_k + \|\vec{a}_k\| \leqslant 1$. Under these conditions, assembling a set of K positive $\{M_k, k = 1,\dots K\}$ forms a valid POVM if and only if $\sum_{k=1}^K \vec{a}_k = \vec{0}$ and $\sum_{k=1}^K b_k = 1$, so as to realize $\sum_{k=1}^K M_k = I_2$. A qubit in state ρ_ξ as in Eq. (12), when measured by such a POVM, leads in Bloch representation to

$$\rho_\xi M_k = \tfrac{1}{2}[(b_k + \vec{r}_\xi \vec{a}_k)I_2 + (b_k\vec{r}_\xi + \vec{a}_k + i\vec{r}_\xi \times \vec{a}_k)\cdot\vec{\sigma}], \tag{22}$$

yielding the probability of measuring M_k as

$$P(k;\xi) = \mathrm{tr}(\rho_\xi M_k) = b_k + \vec{r}_\xi \vec{a}_k. \tag{23}$$

When the measurement outcomes are then processed for estimating ξ, the best achievable performance is controlled by the classical Fisher information defined as [21,22]

$$F_c(\xi) = \sum_{k=1}^{K} \frac{1}{P(k;\xi)} (\partial_\xi P(k;\xi))^2. \tag{24}$$

From Eq. (23), one has the derivative $\partial_\xi P(k;\xi) = \vec{a}_k \partial_\xi \vec{r}_\xi$, giving access to the classical Fisher information associated with the POVM $\{M_k, k = 1,\dots K\}$, as

$$F_c(\xi) = \sum_{k=1}^{K} \frac{(\vec{a}_k \partial_\xi \vec{r}_\xi)^2}{b_k + \vec{a}_k \vec{r}_\xi}. \tag{25}$$

For phase estimation, with the geometric characterization of the derivative in Eq. (20), the classical Fisher information becomes

$$F_c(\xi) = \sum_{k=1}^{K} \frac{[\vec{a}_k(\vec{n} \times \vec{r}_\xi)]^2}{b_k + \vec{a}_k \vec{r}_\xi}. \tag{26}$$

Since any POVM is constrained by the quantum Cramér-Rao inequality $F_c(\xi) \leqslant F_q(\xi)$, the task to maximize the performance is to seek a POVM achieving $F_c(\xi) = F_q(\xi)$, and this especially when $F_q(\xi)$ is at its maximum $F_q^{\max} = 1$ realized by employing an optimal pure probe ρ_0 with a unit Bloch vector $\vec{r}_0$ orthogonal to the rotation axis $\vec{n}$, as seen at the end of Sec. III A. We now show that this can be accomplished by a POVM with $K = 2$ elements.

A valid POVM with $K = 2$ elements is defined by the two measurement operators $M_\pm = (I_2 \pm \vec{a} \cdot \vec{\sigma})/2$, with $\|\vec{a}\,\| = 1$, forming two projectors on two orthogonal directions in $\mathcal{H}_2$, i.e., a von Neumann projective measurement. This is equivalent to measuring on the qubit the spin observable $\Omega = \vec{a} \cdot \vec{\sigma}$ with eigenvalues $\pm\|\vec{a}\,\| = \pm 1$, and we are going to show that there always exists an optimal spin observable $\Omega = \vec{a} \cdot \vec{\sigma}$ to achieve $F_c(\xi) = F_q^{\max} = 1$ for phase estimation on the qubit. The classical Fisher information of Eq. (26) follows as

$$F_c(\xi) = \frac{[\vec{a}\,(\vec{n} \times \vec{r}_\xi)]^2}{1 - (\vec{a}\,\vec{r}_\xi)^2}. \tag{27}$$

With the optimal pure probe of unit $\vec{r}_0$ orthogonal to $\vec{n}$, the rotated Bloch vector $\vec{r}_\xi$ is also orthogonal to $\vec{n}$ and with unit norm. In another orthonormal basis $\{\vec{n}, \vec{n}_\perp = \vec{r}_\xi, \vec{n}\,'_\perp = \vec{n} \times \vec{n}_\perp\}$ of $\mathbb{R}^3$, we introduce for the real unit vector $\vec{a}$ the three coordinates $\vec{a} = [a_1,a_2,a_3]^\top$, with $a_1^2 + a_2^2 + a_3^2 = 1$, leading for Eq. (27) to

$$F_c(\xi) = \frac{a_3^2}{a_1^2 + a_3^2}. \tag{28}$$

By taking $a_1 = 0$, the classical Fisher information $F_c(\xi)$ of Eq. (28) is thus maximized at $F_c^{\max} = F_q^{\max} = 1$, for any (a_2,a_3). Therefore, any pair of orthogonal projectors, or any spin observable $\Omega = \vec{a} \cdot \vec{\sigma}$, defined by a unit vector $\vec{a}$ chosen in the plane orthogonal to the rotation axis $\vec{n}$, realizes an optimal measurement capable of reaching the maximal performance $F_c(\xi) = F_c^{\max} = F_q^{\max} = 1$ uniformly for any unknown phase ξ.

It can be noted that [23] also proposes a general characterization of an optimal measurement, under the form of one-dimensional projectors onto a complete set of orthonormal eigenstates of the quantum score L_ξ, i.e., the score L_ξ as an optimal observable to measure. However, the quantum score L_ξ is defined, via Eq. (2), from the quantum state ρ_ξ carrying the dependence with the unknown parameter ξ to be estimated. As a result, in general, the score L_ξ, and its eigenstates, are expected to depend on the unknown parameter ξ. This is remarked for instance in [7], that in general a projector on the eigenstates of the score L_ξ represents a ξ-dependent measurement and therefore is unaccessible as a realizable solution since ξ in unknown. Moreover, as also remarked in [20], its derivation shows that the inequality $F_c(\xi) \leqslant F_q(\xi)$ applies only to ξ-independent measurements. Here, the analytical expression of Eq. (21) allows us to explicitly perform the eigendecomposition of the quantum score L_ξ. From Eq. (21) we deduce that L_ξ has generally two real eigenvalues $\pm\|\vec{n} \times \vec{r}_\xi\| = \pm\|\vec{n} \times \vec{r}_0\|$, reducing to ± 1 for the optimal probe with a unit $\vec{r}_0$ orthogonal to the rotation axis $\vec{n}$. Also from Eq. (21), the two projectors on the two orthogonal eigenvectors of L_ξ read $(I_2 \pm \vec{a}_\xi \cdot \vec{\sigma})/2$, forming two projectors defined by the unit Bloch vector $\vec{a}_\xi = \vec{n} \times \vec{r}_\xi / \|\vec{n} \times \vec{r}_\xi\|$ which generally depend on the unknown angle ξ via the direction of $\vec{r}_\xi$ in $\mathbb{R}^3$. For the optimal probe with unit $\vec{r}_0 \perp \vec{n}$, then $\vec{a}_\xi = \vec{n} \times \vec{r}_\xi$ is also in the plane orthogonal to $\vec{n}$ and in this plane $\vec{a}_\xi$ makes an angle $\xi + \pi/2$ with $\vec{r}_0$. The measurement formed by the projectors on the eigenstates of L_ξ, equivalent to the spin observable $\Omega = \vec{a}_\xi \cdot \vec{\sigma}$, thus represents a ξ-dependent measurement, therefore not practically realizable. Meanwhile, our optimization above has shown that *any* spin observable $\Omega = \vec{a} \cdot \vec{\sigma}$ with unit $\vec{a} \perp \vec{n}$ is optimal, not necessarily with the ξ-dependent $\vec{a} = \vec{a}_\xi = \vec{n} \times \vec{r}_\xi$, and it offers in this way a ξ-independent optimal measurement protocol. It remains now to find an optimal estimator with a mean-squared error saturating the classical Cramér-Rao inequality controlled by $F_c(\xi)$.

C. Optimizing the estimator

For constructing an efficacious estimator for the phase ξ, we turn to the maximum likelihood method [22]. A measurement performed with the optimal protocol or optimal spin observable $\Omega = \vec{a} \cdot \vec{\sigma}$, defined at the end of Sec. III B, has two outcomes that we denote by ± 1, of probabilities expressible as $P(\pm 1) = (1 \pm s)/2$ with the auxiliary scalar parameter $s = s(\xi) = \vec{a}\,\vec{r}_\xi = \langle \Omega \rangle$. In practice, a sequence of N independent measurements is performed from N identical copies of the qubit in state ρ_ξ, leading to N_+ measurement outcomes at $+1$ and $N - N_+$ measurement outcomes at -1. Estimating s is similar to estimating the parameter of a binomial distribution, and with the likelihood $\mathcal{L}(s) = [(1+s)/2]^{N_+}[(1-s)/2]^{N-N_+}$ the maximum likelihood estimator for s is known to be

$$\widehat{s}_{\rm ML} = 2\frac{N_+}{N} - 1. \tag{29}$$

Based on the expectation $\langle N_+\rangle = N(1+s)/2$ and variance $\mathrm{var}(N_+) = N(1-s^2)/4$, it is also known that $\langle \widehat{s}_{\rm ML}\rangle = s$ so that $\widehat{s}_{\rm ML}$ is an unbiased estimator of s for any N, with the mean-squared error $\langle(\widehat{s}_{\rm ML} - s)^2\rangle = \mathrm{var}(\widehat{s}_{\rm ML}) = 4\mathrm{var}(N_+)/N^2 = (1-s^2)/N$. From Eq. (24), the classical Fisher information for estimating s is known to be $F_c(s) = [\partial_s P(+1)]^2/P(+1) + [\partial_s P(-1)]^2/P(-1) = 1/(1-s^2)$ establishing the known result that $\langle(\widehat{s}_{\rm ML} - s)^2\rangle = \mathrm{var}(\widehat{s}_{\rm ML}) = 1/[NF_c(s)] = (1-s^2)/N$ expressing that $\widehat{s}_{\rm ML}$ of Eq. (29) is a (maximally) efficient estimator of s, for any N.

Next, from these known standard results for estimating the auxiliary parameter s from a binomial distribution, we must go back to our specific problem of constructing an efficacious estimator for the quantum phase ξ. Based on the behavior of the likelihood for a transformed parameter [22], determining the maximum likelihood estimator $\widehat{\xi}_{\rm ML}$ for ξ can be accomplished by inverting the relation $\widehat{s}_{\rm ML} = s(\widehat{\xi}_{\rm ML})$ with $s(\xi) = \vec{a}\,\vec{r}_\xi$ specific to our quantum phase estimation problem. For an angular parameter ξ, the relation $s(\xi) = \vec{a}\,\vec{r}_\xi$ is usually nonlinear. This usually prevents $\widehat{\xi}_{\rm ML}$ from being an unbiased estimator of ξ for any N, since $\widehat{s}_{\rm ML}$ is always an unbiased estimator for s, and the expectation $\langle\cdot\rangle$ and nonlinear $s(\cdot)$ in general do not commute. However, as N increases, commutation is restored with $\langle\widehat{s}_{\rm ML}\rangle = s(\langle\widehat{\xi}_{\rm ML}\rangle)$ entailing $\langle\widehat{\xi}_{\rm ML}\rangle = \xi$, as expected for a maximum likelihood estimator, which is always guaranteed asymptotically unbiased [22].

For the variance of $\widehat{\xi}_{\rm ML}$, we use the theory of error propagation as in [25] to obtain $\mathrm{var}(\widehat{s}_{\rm ML}) = (\partial s/\partial\xi)^2\mathrm{var}(\widehat{\xi}_{\rm ML})$. We have $\partial s/\partial\xi = \vec{a}\,\partial_\xi\vec{r}_\xi = \vec{a}(\vec{n}\times\vec{r}_\xi)$, and since $\mathrm{var}(\widehat{s}_{\rm ML}) = (1-s^2)/N$ we obtain

$$\mathrm{var}(\widehat{\xi}_{\rm ML}) = \langle(\widehat{\xi}_{\rm ML} - \xi)^2\rangle = \frac{1-(\vec{a}\,\vec{r}_\xi)^2}{[\vec{a}\,(\vec{n}\times\vec{r}_\xi)]^2}\frac{1}{N}. \tag{30}$$

Comparing with Eq. (27), we observe that, in the regime of large N, the mean-squared error $\langle(\widehat{\xi}_{\rm ML} - \xi)^2\rangle = 1/[NF_c(\xi)]$, i.e., $\widehat{\xi}_{\rm ML}$, saturates the classical Cramér-Rao inequality, and is thus an efficient estimator as expected from a maximum likelihood estimator [22]. Since $F_c(s) = 1/(1-s^2)$, by confronting with Eq. (27) we also verify that $F_c(\xi) = (\partial s/\partial\xi)^2F_c(s)$ consistently with the definition of the classical Fisher information from Eq. (24). Also, since $\Omega^2 = \mathrm{I}_2$, for Eq. (30) one has $\mathrm{var}(\widehat{\xi}_{\rm ML}) = N^{-1}(\langle\Omega^2\rangle - \langle\Omega\rangle^2)/|\partial\langle\Omega\rangle/\partial\xi|^2$.

It now remains to write an operative expression for the phase estimator $\widehat{\xi}_{\rm ML}$ by explicitly inverting $\widehat{s}_{\rm ML} = s(\widehat{\xi}_{\rm ML})$ with $s(\xi) = \vec{a}\,\vec{r}_\xi$. To simplify this inversion, we have the faculty, as seen at the end of Sec. III B, to choose the measurement vector $\vec{a}$ anywhere in the plane orthogonal to the rotation axis $\vec{n}$, in the presence of a pure probe with unit $\vec{r}_0 \perp \vec{n}$ so as to ensure the conditions of maximal estimation performance. A convenient choice is to fix the unit vector $\vec{a}$ identical to the probe vector $\vec{r}_0$. In such configuration one has $s(\xi) = \vec{a}\,\vec{r}_\xi = \cos(\xi)$, providing via Eq. (29) the explicit phase estimator

$$\widehat{\xi}_{\rm ML} = \arccos(\widehat{s}_{\rm ML}) = \arccos\left(2\frac{N_+}{N} - 1\right), \tag{31}$$

achieving the ultimate best performance $\langle(\widehat{\xi}_{\rm ML} - \xi)^2\rangle = 1/[NF_q^{\max}] = 1/N$.

Strictly speaking, the optimal estimator of Eq. (31) returns a value for the phase ξ in $[0,\pi]$. This is due to the two-element POVM or spin observable preceding this estimator, which conveys information on ξ only through the inner product $s(\xi) = \vec{a}\,\vec{r}_\xi = \cos(\xi)$ according to Eq. (23). If one wants a broader determination of ξ in $[0,2\pi)$, then no spin observable suffices, and one needs to resort to a generalized measurement through a POVM with at least $K = 3$ elements. Nevertheless, as we shall demonstrate in Secs. IV C and IV D, at any $K \geqslant 3$ there always exist an optimal POVM and optimal estimator achieving the ultimate best performance uniformly for any $\xi \in [0,2\pi)$, just as the solution of Eq. (31) does for $\xi \in [0,\pi]$.

IV. ESTIMATION WITH NOISE

A. General quantum noise

We now consider the situation where the quantum state ρ_ξ, before it becomes accessible to measurement for estimation, is affected by quantum noise. The general derivations of Sec. II A still apply for any noisy quantum state ρ_ξ. With quantum noise however, the conditions of Sec. II B defining the dependence of ρ_ξ on ξ need to be modified. As in Sec. II B, we consider an input probe state ρ_0 experiencing the unitary transformation U_ξ of Eq. (7), but this time to yield an intermediate quantum state $\rho_1(\xi) = U_\xi\rho_0U_\xi^\dagger$. And then $\rho_1(\xi)$ is subjected to quantum noise to deliver the state ρ_ξ accessible to measurement for estimation, and ruled by Sec. II A. The action of the noise can generally be represented by a quantum operation under the Kraus form [24]

$$\rho_\xi = \sum_\ell \Lambda_\ell\rho_1(\xi)\Lambda_\ell^\dagger, \tag{32}$$

with the Kraus operators Λ_ℓ satisfying $\sum_\ell \Lambda_\ell^\dagger\Lambda_\ell = I_D$, the identity operator of $\mathcal{H}_D$, in order for Eq. (32) to realize a completely positive trace-preserving linear map. The operator-sum representation of Eq. (32) is a modeling choice for the noise which is often adopted in quantum information [24]. It provides an end-to-end modeling of a change of state, and requires a limited amount of modeling resources since any state transformation can always be described with a maximum of D^2 Kraus operators for quantum states in D dimension, whatever the complexity of the underlying detailed processes. Such an end-to-end transformation is well suited for our estimation task where essentially a description of the noisy state at the time and stage of the quantum measurement is needed. Alternatively, the noise could be modeled by a Lindblad equation describing the evolution of a quantum state in continuous time with a differential equation [24]. This would however require to specify Lindblad operators for each of the underlying physical process to be taken into account, possibly allowing more detailed control, although this is not necessary for obtaining a general characterization of the performance in qubit phase estimation as will follow.

The noise model of Eq. (32) provides direct access to the derivative $\partial_\xi \rho_\xi$ intervening in the quantum score and quantum Fisher information of Sec. II A, as, for any ξ-independent noise,

$$\partial_\xi \rho_\xi = \sum_\ell \Lambda_\ell \partial_\xi \rho_1(\xi) \Lambda_\ell^\dagger, \tag{33}$$

with $\partial_\xi \rho_1 = i[\rho_1, G] = i U_\xi [\rho_0, G] U_\xi^\dagger$ as in Sec. II B. Next, in the noisy case, the noise on the quantum system should be further specified before one can obtain useful expressions extending those of Sec. II B.

In the line of Sec. III, we now address the case of the qubit, with noise. When the general expressions of Sec. II A are specified to a qubit state ρ_ξ characterized by the Bloch vector $\vec{r}_\xi$ as in Eq. (12), they give rise to Eqs. (13)–(16) which are general for any (noisy) qubit state with any dependence on ξ. On a qubit state $\rho_1(\xi)$ with Bloch vector $\vec{r}_1(\xi)$, the action of the quantum noise expressed by Eq. (32) can always be described as an affine transformation on the Bloch vectors reading [20,24,26]

$$\vec{r}_\xi = A\vec{r}_1(\xi) + \vec{c}, \tag{34}$$

with A a 3×3 real matrix and $\vec{c}$ a real vector in $\mathbb{R}^3$ together characterizing the noise, and realizing a mapping of the Bloch ball of $\mathbb{R}^3$ onto itself. We then obtain $\partial_\xi \vec{r}_\xi = A \partial_\xi \vec{r}_1(\xi)$.

For phase estimation now, as in Sec. III, the probe state ρ_0 is characterized by the Bloch vector $\vec{r}_0$, the intermediate qubit state $\rho_1(\xi)$ by the Bloch vector $\vec{r}_1(\xi)$ which is $\vec{r}_0$ rotated by the angle ξ around the axis $\vec{n}$. We still have as in Eq. (20) the useful geometric characterization $\partial_\xi \vec{r}_1(\xi) = \vec{n} \times \vec{r}_1(\xi)$. With this specification of the dependence on ξ as a phase on a noisy qubit, the expressions of Eqs. (13) and (14) can be further specified, leading in particular to the quantum Fisher information

$$F_q(\xi) = \frac{[(A\vec{r}_1 + \vec{c})A(\vec{n} \times \vec{r}_1)]^2}{1 - (A\vec{r}_1 + \vec{c})^2} + [A(\vec{n} \times \vec{r}_1)]^2, \tag{35}$$

especially matching Eq. (19) when there is no noise, i.e., $A = I_3$ and $\vec{c} = \vec{0}$, since always $\|\vec{n} \times \vec{r}_1\| = \|\vec{n} \times \vec{r}_0\|$.

As for qubit phase estimation in the noise-free case of Sec. III A, it is possible in the noisy case to address the issue of optimizing the input probe ρ_0 to maximize the quantum Fisher information $F_q(\xi)$ of Eq. (35). When seen as a function of $\vec{r}_1$, the Fisher information $F_q(\xi)$ of Eq. (35) takes its maximum at $\|\vec{r}_1\| = 1$, and the vectors $A\vec{r}_1$ and $\vec{c}$ with an acute angle (otherwise by reversing $\vec{r}_0$ into $-\vec{r}_0$ reverses $\vec{r}_1$ into $-\vec{r}_1$ and would increase $F_q(\xi)$, which is not feasible at the maximum). The condition $\|\vec{r}_1\| = 1$ is obtainable only through $\|\vec{r}_0\| = 1$, i.e., again with a pure input probe ρ_0. We conclude in the noisy case as in the noise-free case of Sec. III A, that the maximum of the quantum Fisher information $F_q(\xi)$ of Eq. (35) is realized necessarily by a pure input probe ρ_0.

In general it is known that the quantum Fisher information $F_q(\xi) \equiv F_q(\xi; \rho_\xi)$ when seen as a functional of the density operator ρ_ξ is a convex ($\cup$) functional of ρ_ξ [27,28]. This convexity property implies that, when there is no constraint on ρ_ξ, the Fisher information $F_q(\xi)$ is maximized by a pure state ρ_ξ [27]. Here ρ_ξ is constrained by the constitutive transformation $\rho_0 \to \rho_\xi$ which includes the action of the noise that usually results in ρ_ξ being forced to be a mixed state. Yet, the linearity of the map $\rho_0 \to \rho_\xi$ ensures that the quantum Fisher information $F_q(\xi)$ when seen as a functional of the input probe ρ_0 is also a convex functional of ρ_0. This convexity in ρ_0 also implies that $F_q(\xi)$ is maximized by a pure input probe ρ_0. Here, our theoretical expressions of Eqs. (35) and (19) for the Fisher information $F_q(\xi)$ based on a geometric picture of qubit Bloch vectors in $\mathbb{R}^3$ offers an alternative proof of the optimality of a pure input probe ρ_0 not resorting to a convexity argument. Furthermore, this geometric picture will allow us here to characterize, through its Bloch vector $\vec{r}_0$, which pure state ρ_0 precisely is optimal.

Beyond, in the noisy case, with a pure probe ρ_0, it does not suit in general to take a unit $\vec{r}_0$ orthogonal to $\vec{n}$ to maximize $F_q(\xi)$ of Eq. (35). On the contrary, in general, the direction of the optimal $\vec{r}_0$ has to specifically match the geometric properties of $(A,\vec{c})$ characterizing the noise in $\mathbb{R}^3$. The optimal unit vector $\vec{r}_1$ maximizing $F_q(\xi)$ of Eq. (35) is conveniently referred to an orthonormal basis $\{\vec{n}, \vec{n}_\perp, \vec{n}\,'_\perp = \vec{n} \times \vec{n}_\perp\}$ of $\mathbb{R}^3$ as in Sec. III, and in this basis $\vec{r}_1$ is specified by a coelevation angle θ_1 with $\vec{n}$ and an azimuth angle φ_1 in the plane $(\vec{n}_\perp, \vec{n}\,'_\perp)$ orthogonal to $\vec{n}$. Then the maximum of $F_q(\xi)$ in Eq. (35) will occur at some optimal point $\vec{r}_1 = \vec{r}_1^{\,\text{opt}}$ characterized by $(\|\vec{r}_1\| = 1, \theta_1^{\text{opt}}, \varphi_1^{\text{opt}})$ determined by the geometry of $(A,\vec{c})$ and $\vec{n}$. Such optimal conditions have been explicitly worked out in [20] for different noise models $(A,\vec{c})$ relevant to the qubit. Moreover, the optimal $\vec{r}_1^{\,\text{opt}} = (\|\vec{r}_1\| = 1, \theta_1^{\text{opt}}, \varphi_1^{\text{opt}})$ determines the optimal input probe ρ_0^{opt} before the rotation by ξ via $\vec{r}_0^{\,\text{opt}} = (\|\vec{r}_0\| = 1, \theta_0^{\text{opt}} = \theta_1^{\text{opt}}, \varphi_0^{\text{opt}} = \varphi_1^{\text{opt}} - \xi)$, by virtue of referring to the basis $\{\vec{n}, \vec{n}_\perp, \vec{n}\,'_\perp\}$. This demonstrates that in general the optimal pure input probe ρ_0^{opt} maximizing $F_q(\xi)$ of Eq. (35) is specifically determined by the noise via $(A,\vec{c})$ and by the rotation axis $\vec{n}$, but also that ρ_0^{opt} can usually be expected to depend on the unknown phase ξ via its azimuth $\varphi_0^{\text{opt}} = \varphi_1^{\text{opt}} - \xi$, when φ_1^{opt} is determined by $(A,\vec{c})$ and $\vec{n}$. This is especially what is verified in [20], where feedback adaptive methods or Bayesian approaches are envisaged to handle such situations of a ξ-dependent solution and obtain an optimized input probe. Only in the special conditions where $F_q(\xi)$ of Eq. (35) could remain at its maximum in θ_1^{opt} for any φ_1, could one obtain a ξ-independent optimal probe ρ_0^{opt}; but this can only occur for special noises with specific geometry for $(A,\vec{c})$ in $\mathbb{R}^3$ and matched to special orientations of the rotation axis $\vec{n}$.

In the presence of an arbitrary noise, the program of Sec. III is thus no longer feasible in general, which would allow us to determine a ξ-independent optimal strategy combining an input probe, quantum measurement, and estimator capable of reaching uniformly for any unknown phase ξ the ultimate best performance for estimation fixed by the maximum of the quantum Fisher information in Eq. (35). Instead, adaptive techniques with feedback have to be employed to iteratively construct an efficient ξ-dependent estimation strategy, as performed for instance in [7,17,18,29–31]. There is however a specific noise, highly relevant for the qubit, where the maximization of the quantum Fisher information $F_q(\xi)$ of Eq. (35) is not limited by a ξ-dependent solution. This is the case of the depolarizing noise, which we now address.

B. Optimizing the quantum Fisher information with depolarizing noise

The depolarizing noise [24] implements a quantum operation of Eq. (32) taking the form

$$\rho_\xi = (1-p)\rho_1 + \frac{p}{3}(\sigma_x\rho_1\sigma_x^\dagger + \sigma_y\rho_1\sigma_y^\dagger + \sigma_z\rho_1\sigma_z^\dagger), \quad (36)$$

where the action of the noise is to leave the qubit state unchanged with the probability $1-p$ and to apply any one of the three Pauli operators with equal probability $p/3$. This is equivalent in Eq. (34) for the Bloch vectors with a matrix $A=\alpha I_3$ proportional to the identity matrix I_3 of $\mathbb{R}^3$ with $\alpha = 1-4p/3$, and $\vec{c}=\vec{0}$. The depolarizing noise is an important noise model often considered in quantum information [24,32]; it also represents in some sense a worse-case situation of quantum noise [32], and in this respect its analysis provides a picture interpretable as a conservative reference. Here we assume that the noise level, quantified by p or α, is perfectly known, so that we address the reference scenario of estimation of only the parameter of interest ξ with no nuisance parameters. However, we show in the Appendix that our approach can be applied as well to derive the optimal strategy and performance for a separate estimation of the noise level, simply by considering that the unitary transformation U_ξ is absent and only the depolarizing noise channel is present to act on the input probe ρ_0.

In Bloch representation, the action of the depolarizing noise resulting from Eq. (36) is to isotropically compress the Bloch vector $\vec{r}_1(\xi)$ by the factor α to produce $\vec{r}_\xi = \alpha\vec{r}_1(\xi)$. Geometrically in $\mathbb{R}^3$, this action preserves the orientation of $\vec{r}_1(\xi)$ unchanged in $\vec{r}_\xi$, and also preserves other interesting geometric properties useful to the analysis. Since from Eq. (20) we still have $\partial_\xi\vec{r}_1(\xi) = \vec{n}\times\vec{r}_1(\xi)$, we therefore also have $\partial_\xi\vec{r}_\xi = \vec{n}\times\vec{r}_\xi$, entailing $\vec{r}_\xi\,\partial_\xi\vec{r}_\xi = 0$ as in the noise-free case. The quantum score of Eq. (13) then follows as

$$L_\xi = (\vec{n}\times\vec{r}_\xi)\cdot\vec{\sigma} = \alpha(\vec{n}\times\vec{r}_1)\cdot\vec{\sigma}, \quad (37)$$

and the quantum Fisher information of Eq. (14) as

$$F_q(\xi) = (\vec{n}\times\vec{r}_\xi)^2 = \alpha^2(\vec{n}\times\vec{r}_0)^2, \quad (38)$$

with Eq. (38) especially matching Eq. (19) when the depolarizing noise vanishes at $p=0$ and $\alpha=1$. Equations (37) and (38), like Eqs. (18), (19), and (21), with their concise geometric forms, are not contained in [20]. Also, from its action in Bloch representation it is visible that the depolarizing noise commutes with the unitary U_ξ, since the compression by α and rotation by ξ commute in $\mathbb{R}^3$. So we model a noise which can act equally before or after the unitary U_ξ. The depolarizing noise can even be distributed, with one part acting before U_ξ with a factor α_1, and another part acting after U_ξ with a factor α_2, with a net effect which can be lumped into a single noise channel with a factor $\alpha=\alpha_1\alpha_2$ as considered here. So the modeling (and estimation in the Appendix) of a noise channel with a single lumped parameter α can in practice represent situations of several distributed noises.

Based on the simple form obtained for Eq. (38), the issue of optimizing the input probe ρ_0 is again readily solved, so as to maximize the quantum Fisher information $F_q(\xi)$ for qubit phase estimation in the presence of depolarizing noise. As in the noise-free case of Sec. III A, the optimal input probe ρ_0 has to be chosen as a pure state with a unit Bloch vector $\vec{r}_0$ orthogonal to the rotation axis $\vec{n}$; this to achieve the overall maximum of $F_q(\xi)$ which is $F_q^{\max}=\alpha^2$. This especially demonstrates the possibility of obtaining an optimal input probe independent of the unknown phase ξ, that represents as explained above a rather exceptional property in the presence of noise, and that occurs here because of the high symmetry (isotropy) of the depolarizing noise. The next step then, as in the noise-free case of Sec. III B, is to examine the possibility of a quantum measurement protocol capable of reaching this maximal Fisher information $F_q^{\max}=\alpha^2$.

C. Optimizing the quantum measurement with depolarizing noise

For implementing a quantum measurement on the noisy qubit, we envisage a general POVM as described in the beginning of Sec. III B. Such a general POVM is associated with the classical Fisher information $F_c(\xi)$ defined by Eq. (25), with a Bloch vector $\vec{r}_\xi$ when measuring a noisy qubit which is given by Eq. (34). For phase estimation $\partial_\xi\vec{r}_1(\xi)=\vec{n}\times\vec{r}_1(\xi)$, and with the depolarizing noise, we finally obtain the classical Fisher information

$$F_c(\xi) = \alpha^2\sum_{k=1}^{K}\frac{[\vec{a}_k(\vec{n}\times\vec{r}_1)]^2}{b_k+\alpha\vec{a}_k\vec{r}_1}. \quad (39)$$

In the line of Sec. III B, we seek an optimal POVM reaching the overall maximal performance $F_c(\xi)=F_q(\xi)$ when $F_q(\xi)=F_q^{\max}=\alpha^2$ achieved by an optimal input probe ρ_0 with a unit $\vec{r}_0$ orthogonal to the rotation axis $\vec{n}$. For this goal, as in Sec. III B, we envisage a POVM with $K=2$ elements, defined by the two measurement operators $M_\pm = (I_2\pm\vec{a}\cdot\vec{\sigma})/2$, with $\|\vec{a}\|=1$, equivalent to measuring the spin observable $\Omega=\vec{a}\cdot\vec{\sigma}$. The classical Fisher information of Eq. (39) follows as

$$F_c(\xi) = \alpha^2\frac{[\vec{a}(\vec{n}\times\vec{r}_1)]^2}{1-\alpha^2(\vec{a}\,\vec{r}_1)^2}, \quad (40)$$

matching the noise-free case of Eq. (27) when $\alpha=1$.

With the optimal pure probe of unit $\vec{r}_0$ orthogonal to $\vec{n}$, the rotated Bloch vector $\vec{r}_1(\xi)$ is also orthogonal to $\vec{n}$ and with unit norm. In the orthonormal basis $\{\vec{n},\vec{n}_\perp=\vec{r}_1,\vec{n}\,'_\perp=\vec{n}\times\vec{n}_\perp\}$ of $\mathbb{R}^3$, we introduce again for the real unit vector $\vec{a}$ the three coordinates $\vec{a}=[a_1,a_2,a_3]^\top$, with $a_1^2+a_2^2+a_3^2=1$, leading for Eq. (40) to

$$F_c(\xi) = \alpha^2\frac{a_3^2}{1-\alpha^2a_2^2}. \quad (41)$$

To maximize Eq. (41) it is necessary to set $a_1=0$, then $1/F_c(\xi)=1+(1-\alpha^2)/(\alpha^2a_3^2)$ is clearly minimized by $a_3=1$, therefore $F_c(\xi)$ is maximized by $(a_1=0,a_2=0,a_3=1)$ at $F_c^{\max}=F_q^{\max}=\alpha^2$. This is a necessary and sufficient condition, in order to maximize $F_c(\xi)$ of Eq. (40) at $F_c^{\max}=F_q^{\max}=\alpha^2$, to select $\vec{a}$ in the plane orthogonal to $\vec{n}$ and also orthogonal to the rotated Bloch vector $\vec{r}_1(\xi)$. This is a ξ-dependent optimal solution. This means that, contrary to the noise-free case of Sec. III B, with the depolarizing noise there is no ξ-independent optimal POVM with two elements, or optimal spin observable $\Omega=\vec{a}\cdot\vec{\sigma}$, that would be able to

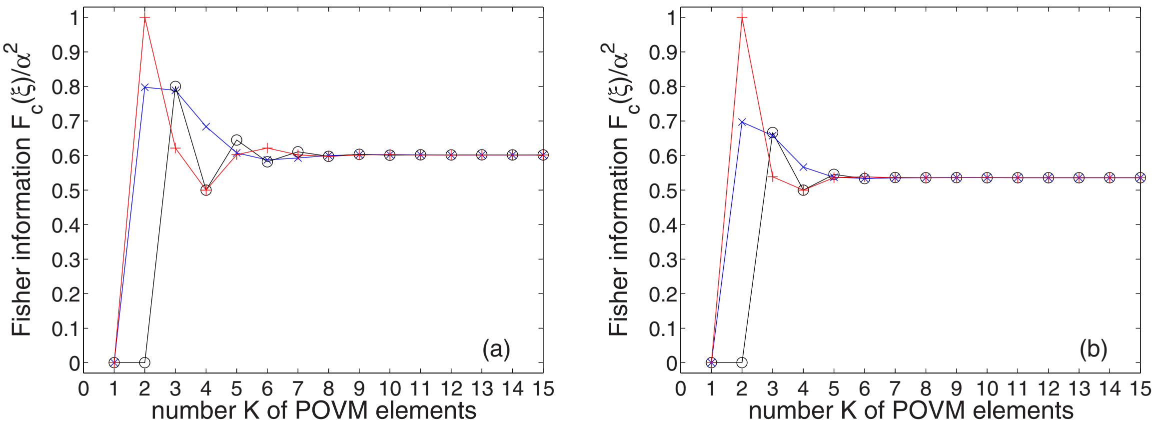

FIG. 1. Classical Fisher information $F_c(\xi)/\alpha^2$ from Eq. (42) normalized by the overall maximum dictated by the quantum Fisher information $F_q^{\max} = \alpha^2$, as a function of the number K of POVM elements for the quantum measurement, for different values of the phase angle ξ to be estimated: ($\circ$) $\xi = 0$, ($\times$) $\xi = \pi/\sqrt{2}$, (+) $\xi = 3\pi/2$. The depolarizing noise factor is $\alpha = 0.75$ in (a), and $\alpha = 0.5$ in (b).

realize the overall maximal performance $F_c(\xi) = F_q^{\max} = \alpha^2$ uniformly for any ξ.

We can also explicitly exhibit, as done for the noise-free case in the last paragraph of Sec. III B, the measurement formed by the projectors on the eigenstates of the quantum score L_ξ. From Eq. (37), the score L_ξ has generally two real eigenvalues $\pm\|\vec{n} \times \vec{r}_\xi\| = \pm\alpha\|\vec{n} \times \vec{r}_0\|$, reducing to $\pm\alpha$ for the optimal probe with a unit $\vec{r}_0 \perp \vec{n}$. Also from Eq. (37), the two projectors on the two orthogonal eigenvectors of L_ξ are $(I_2 \pm \vec{a}_\xi \cdot \vec{\sigma})/2$ with the unit Bloch vector $\vec{a}_\xi = \vec{n} \times \vec{r}_\xi / \|\vec{n} \times \vec{r}_\xi\|$ and they generally depend on the unknown angle ξ via the direction of $\vec{r}_\xi$, defining a ξ-dependent measurement through the spin observable $\Omega = \vec{a}_\xi \cdot \vec{\sigma}$. For the optimal probe with unit $\vec{r}_0 \perp \vec{n}$, then $\vec{a}_\xi = \vec{n} \times \vec{r}_1(\xi) = \vec{n} \times \vec{r}_\xi/\alpha$ is also in the plane orthogonal to $\vec{n}$ and is orthogonal to $\vec{r}_1(\xi)$ or equivalently making an angle $\xi + \pi/2$ with $\vec{r}_0$. Such a ξ-dependent observable $\Omega = \vec{a}_\xi \cdot \vec{\sigma}$ coincides with our previous optimal solution found above, and from our optimization above is now known to be the only solution for an optimal measurement, when measurements are restricted to spin observables. By contrast, it was not the only solution in the noise-free case of Sec. III B, where ξ-independent optimal observables $\Omega = \vec{a} \cdot \vec{\sigma}$ were characterized.

The analysis of Eq. (41) also shows that a spin observable $\Omega = \vec{a} \cdot \vec{\sigma}$ with an $\vec{a}$ approaching the configuration ($a_1 = 0, a_2 = 1, a_3 = 0$), i.e., an $\vec{a}$ approaching $\vec{r}_1(\xi)$, would lead to a classical Fisher information $F_c(\xi)$ approaching zero, representing a quantum measurement completely inoperative for estimating ξ. The blind selection of a fixed $\vec{a}$ in the plane orthogonal to $\vec{n}$ is thus a rather hazardous issue, because this plane contains both the best and the worse configurations feasible for $\vec{a}$, both depending on the unknown value of the phase ξ to be estimated. Selecting an $\vec{a}$ outside this plane (with $a_1 \neq 0$) does not solve the difficulty, which remains the same for the selection of the projection of $\vec{a}$ in the plane, and only reduces the maximum feasible for $F_c(\xi)$ without avoiding the risk of a zero $F_c(\xi)$.

Since we just found that there is no ξ-independent optimal observable, when seeking a ξ-independent optimal measurement we still have the possibility to turn to POVM with a higher number $K > 2$ of elements. We consider a POVM with K measurement operators $M_k = (I_2 + \vec{a}_k \cdot \vec{\sigma})/K$, with $\|\vec{a}_k\| = 1$, each proportional to a projector in $\mathcal{H}_2$. To maximize $F_c(\xi)$, based on Eq. (39) we know that each $\vec{a}_k$ is better placed in the plane orthogonal to $\vec{n}$. The optimal pure input probe $\rho_0^{\rm opt}$ has also its unit Bloch vector $\vec{r}_0^{\,\rm opt}$ in this plane. In the orthonormal basis $\{\vec{r}_0^{\,\rm opt}, \vec{n} \times \vec{r}_0^{\,\rm opt}\}$ of this plane, each unit vector $\vec{a}_k$ is taken with coordinates $\vec{a}_k = [\cos(\phi_k), \sin(\phi_k)]^\top$, with the azimuth angle ϕ_k, for $k = 1$ to K. This leads for the classical Fisher information of Eq. (39) to

$$F_c(\xi) = \frac{\alpha^2}{K} \sum_{k=1}^{K} \frac{\sin^2(\phi_k - \xi)}{1 + \alpha\cos(\phi_k - \xi)}. \qquad (42)$$

When the phase angle ξ of the rotated Bloch vector $\vec{r}_1(\xi)$ is fully unknown in the plane orthogonal to $\vec{n}$, there is no motivation for a choice other than uniformly distributing the K vectors $\vec{a}_k$ in this plane, with $\phi_k = 2\pi(k-1)/K$ for $k = 1$ to K. With this choice, Fig. 1 shows the classical Fisher information $F_c(\xi)$ from Eq. (42), as a function of the number K of POVM elements, and for different values of the phase angle ξ to be estimated.

In Fig. 1, it is observed that at small number $K \geqslant 2$ of POVM elements, there is a clear dependence of the performance $F_c(\xi)$ on the unknown phase ξ. This is expected when there is a small number K of measurement vectors $\vec{a}_k$ to tract the rotated Bloch vector $\vec{r}_1(\xi)$ at an angle ξ in the plane orthogonal to $\vec{n}$. However, at sufficiently large K, around $K = 8$ to 10 in Fig. 1, the uniform distribution in the plane of a larger number K of $\vec{a}_k$ reduces the dependence on ξ, with $F_c(\xi)$ tending to stabilize at an asymptotic limiting value. At large K, when the sum in Eq. (42) tends to an integral over $[0, 2\pi)$, the Fisher information $F_c(\xi)$ tends to the limit expressible as

$$F_c^\infty = \frac{\alpha^2}{2\pi} \int_0^{2\pi} \frac{\sin^2(\phi)}{1 + \alpha\cos(\phi)} d\phi = 1 - \sqrt{1 - \alpha^2}, \qquad (43)$$

independent of the unknown phase ξ. In practice, $K = 8$ to 10 allows $F_c(\xi)$ of Eq. (42) to come close to F_c^∞ of Eq. (43) for any ξ and α, since numerically, starting with $K = 8$, it is found that $F_c(\xi)$ of Eq. (42) never drops below $0.9F_c^\infty$.

It is a significant finding that the limit F_c^∞ of Eq. (43), as visible in Fig. 1, does not reach the overall maximum dictated

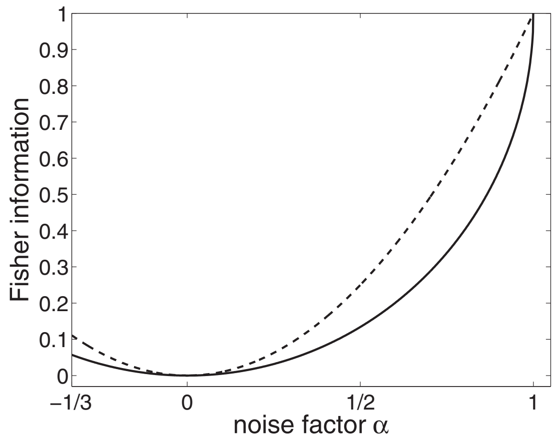

FIG. 2. Classical Fisher information $F_c^\infty = 1 - \sqrt{1-\alpha^2}$ from Eq. (43) at large number K of POVM elements (solid line), and overall maximum dictated by the quantum Fisher information $F_q^{\max} = \alpha^2$ (dashed line), as a function of the depolarizing noise factor $\alpha \in [-1/3,1]$.

by the quantum Fisher information $F_q^{\max} = \alpha^2$, but remains generally below it. This is also illustrated by Fig. 2, confronting the classical Fisher information F_c^∞ of Eq. (43) at large K, with the overall maximum dictated by the quantum Fisher information $F_q^{\max} = \alpha^2$, over the whole range of feasible values for the depolarizing noise factor $\alpha = 1 - 4p/3 \in [-1/3,1]$ related to Eq. (36) with a probability $p \in [0,1]$.

For any noise factor α in Fig. 2, the performance F_c^∞ is below the maximum $F_q^{\max}$, demonstrating that, in the presence of depolarizing noise, even at the price of a large number K of POVM elements, it is not possible to obtain a ξ-independent measurement associated with a classical Fisher information $F_c(\xi)$ reaching the maximum $F_q^{\max}$ uniformly for all ξ. As we saw above in this section, the optimal measurement achieving $F_c(\xi) = F_q^{\max}$ is a ξ-dependent measurement, and our analysis now shows that there is no ξ-independent generalized measurement that can be as efficient as the optimal ξ-dependent measurement precisely matched to the specific value of ξ to be estimated. A ξ-independent optimal measurement was however possible with no noise, as established in Sec. III B, with this same conclusion also established by Fig. 2 at the extreme point $\alpha = 1$ of vanishing noise where locally $F_c^\infty = F_q^{\max}$.

A concrete mechanism for this loss of estimation efficacy in the presence of noise can be pictured in the following way. With no noise, an optimal input probe with $\vec{r}_0 \perp \vec{n}$ is a pure state $|\psi_0\rangle$. The rotated state is $|\psi_1\rangle = U_\xi |\psi_0\rangle = \cos(\xi/2)|\psi_0\rangle + \sin(\xi/2)|\psi_0^\perp\rangle$, with $|\psi_0^\perp\rangle$ a vector of $\mathcal{H}_2$ orthogonal to $|\psi_0\rangle$. An optimal measurement is realized by the spin observable $\Omega = \vec{r}_0 \cdot \vec{\sigma}$ which projects $|\psi_1\rangle$ on the orthonormal basis $\{|\psi_0\rangle, |\psi_0^\perp\rangle\}$, allowing the ultimate estimation efficacy as we now know. By contrast with depolarizing noise, the noisy state to be measured is the pure state $|\psi_1\rangle$ degraded by the noise into the mixed state $\rho_\xi = (I_2 + \vec{r}_\xi \cdot \vec{\sigma})/2$ with Bloch vector $\vec{r}_\xi = \alpha \vec{r}_1(\xi)$. By spectral decomposition ρ_ξ has $|\psi_1\rangle$ and $|\psi_1^\perp\rangle$ as eigenvectors, with respectively $(1+\alpha)/2$ and $(1-\alpha)/2$ as eigenvalues, with $|\psi_1^\perp\rangle$ orthogonal to $|\psi_1\rangle$. The mixed state ρ_ξ is thus equivalent to the statistical ensemble $\{|\psi_1\rangle, (1+\alpha)/2; |\psi_1^\perp\rangle, (1-\alpha)/2\}$ of two probabilistically weighted pure states. At low noise when $\alpha \lesssim 1$, most of the time the noisy state ρ_ξ is seen by the measurement as $|\psi_1\rangle$ just as if there were no noise; but on some occasions, with the small probability $(1-\alpha)/2$, the noisy state ρ_ξ is seen by the measurement as the orthogonal vector $|\psi_1^\perp\rangle$. It is this possibility which inevitably degrades the performance in estimation and renders ξ dependent the optimal measurement. With no noise, only $|\psi_1\rangle$ is presented to the measurement, with for estimation the statistical performance ensuing from the inherent probabilistic nature of quantum measurement. With noise, on some random occasions $|\psi_1^\perp\rangle$ is presented to the measurement instead of $|\psi_1\rangle$, and this constitutes an added source of equivocation which inevitably degrades the statistical performance in estimation. Moreover, with no noise the optimal measurement has to be optimal for measuring the pure state $|\psi_1\rangle$ alone, and this does not come with a stringent ξ-dependent position of the measurement relative to $|\psi_1\rangle$ for optimality. With noise the optimal measurement has to be optimal for measuring two orthogonal pure states $|\psi_1\rangle$ and $|\psi_1^\perp\rangle$ occurring randomly, which is a more demanding configuration, whence the reduced performance, and this imposes a specific ξ-dependent position of the measurement relative to $|\psi_1\rangle$ and $|\psi_1^\perp\rangle$ for optimality.

One can also envisage, at any number $K \geqslant 2$ of POVM elements, the averaging of $F_c(\xi)$ of Eq. (42) over the unknown phase ξ uniformly distributed over $[0,2\pi)$, defining $\overline{F}_c = \int_0^{2\pi} F_c(\xi) d\xi/2\pi$. It is the same integral as in Eq. (43) which is involved, yielding also $\overline{F}_c = 1 - \sqrt{1-\alpha^2} = F_c^\infty$, and this is true for any $K \geqslant 2$, indicating that when averaged over the unknown phase ξ, each number K of POVM elements leads to the same average performance $\overline{F}_c = F_c^\infty$, although with less variability for $F_c(\xi)$ around $\overline{F}_c$ at larger K.

It is also interesting to study Eq. (42) at $K = 2$, when $\vec{a}_1 = -\vec{a}_2$ and $\phi_1 = 0$ and $\phi_2 = \pi$, which corresponds to the case of a two-element POVM as previously considered at the occasion of Eq. (40). The classical Fisher information of Eq. (42) at $K = 2$ becomes

$$F_c(\xi) = \alpha^2 \frac{\sin^2(\xi)}{1 - \alpha^2 \cos^2(\xi)}, \tag{44}$$

and Fig. 3 represents this $F_c(\xi)$ at various values of the noise factor α, which is useful to better understand the limit $\alpha \to 1$ at vanishing depolarizing noise.

With no noise at $\alpha = 1$, the classical Fisher information $F_c(\xi)$ of Eq. (44) is the constant $F_c(\xi) = F_q^{\max} = 1$ independent of the phase ξ. This is the noise-free case studied in Sec. III B, where any unit vector $\vec{a}$ in the plane orthogonal to $\vec{n}$ defines an optimal two-element POVM or optimal spin observable $\Omega = \vec{a} \cdot \vec{\sigma}$. As soon as a small amount of noise comes into play, with $\alpha \lesssim 1$, a sharp and narrow dip of $F_c(\xi)$ down to zero appears in Fig. 3 around $\xi = 0$ and $\xi = \pi$. At the same time, $F_c(\xi)$ takes a maximum $F_c(\xi) = F_q^{\max} = \alpha^2$ localized at $\xi = \pi/2$ and $\xi = 3\pi/2$, but this maximum remains flat and broad for α close to 1, as visible in Fig. 3. This expresses, as explained above while analyzing Eq. (40), that strictly speaking, at $\alpha < 1$, reaching the maximum $F_c(\xi) = F_q^{\max} = \alpha^2$ requires a specific ξ-dependent $\vec{a}$ orthogonal to $\vec{r}_1(\xi)$, as marked by the maxima at $\xi = \pi/2$ and $\xi = 3\pi/2$

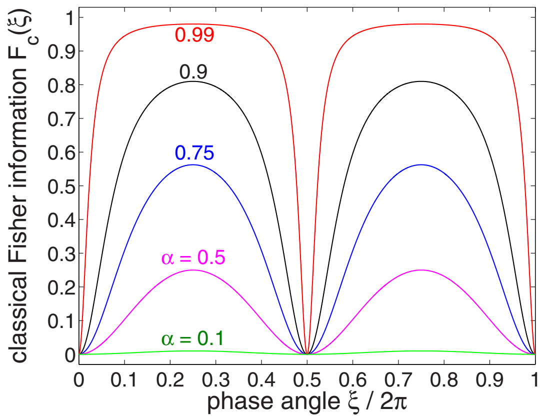

FIG. 3. Classical Fisher information $F_c(\xi)$ from Eq. (44) for a POVM with $K = 2$ elements, as a function of the phase angle $\xi \in [0,2\pi)$ to be estimated, and for various values of the depolarizing noise factor $\alpha = 0.1$, 0.5, 0.75, 0.9, and 0.99.

in Fig. 3; also, $\vec{a}$ colinear to $\vec{r}_1(\xi)$ realizes the minimum $F_c(\xi) = 0$, as it occurs at $\xi = 0$ and $\xi = \pi$ in Fig. 3. However, at α close to 1, the maximum at $F_c(\xi) = F_q^{\max} = \alpha^2$ is flat and broad, along $\xi \in [0,2\pi)$, while the minimum at $F_c(\xi) = 0$ is sharp and narrow, thus explaining how one gradually departs from a ξ-independent optimal two-element POVM at $\alpha = 1$ to jump to a ξ-dependent optimal two-element POVM as soon as $\alpha < 1$.

Relying on a two-element POVM, characterized by a fixed vector $\vec{a}$ lying in the plane orthogonal to $\vec{n}$, exposes one to experience a very poor performance close to $F_c(\xi) = 0$ for some values of the unknown phase ξ to be estimated [associated with a rotated vector $\vec{r}_1(\xi)$ quasicolinear to $\vec{a}$]. But with the same $\vec{a}$, the performance may be close to the overall maximum $F_c(\xi) = F_q^{\max} = \alpha^2$ for other values of ξ [associated with a rotated vector $\vec{r}_1(\xi)$ quasiorthogonal to $\vec{a}$]. On average, when averaged over ξ uniform in $[0,2\pi)$, the performance is $\overline{F}_c = 1 - \sqrt{1-\alpha^2} \leqslant F_q^{\max} = \alpha^2$, for any $\vec{a}$ in the plane orthogonal to $\vec{n}$, as visible in Fig. 2 since $\overline{F}_c = F_c^{\infty}$. A POVM with a larger number K of elements, typically $K = 8$ to 10 from Fig. 1, allows one to reach the same average performance $\overline{F}_c = 1 - \sqrt{1-\alpha^2}$, yet with much smaller deviation of $F_c(\xi)$ around $\overline{F}_c$ for various ξ, compared to the two-element POVM which shares the same $\overline{F}_c$ but with the most extreme possible deviation, between $F_c(\xi) = F_q^{\max} = \alpha^2$ and $F_c(\xi) = 0$ according to ξ. With no noise at $\alpha = 1$, the K-element POVM illustrated in Fig. 1 always achieves in Eq. (42) a classical Fisher information $F_c(\xi) = F_q^{\max} = 1$ for any $K \geqslant 2$, meaning that any such POVM realizes a ξ-independent optimal POVM for any $K \geqslant 2$. So if one is not sure of the noise level α, a conservative choice can be a POVM with large K around $K = 8$ to 10, achieving the ξ-independent performance $F_c(\xi) = F_c^{\infty} = \overline{F}_c$ at any noise level α, while being also an optimal POVM reaching $F_c(\xi) = F_q^{\max} = 1$ with no noise at $\alpha = 1$.

For constructing a ξ-independent POVM maximizing the classical Fisher information $F_c(\xi)$, as we have just seen, the K measurement operators $M_k = (I_2 + \vec{a}_k \cdot \vec{\sigma})/K$ should have their $\vec{a}_k$ in the plane orthogonal to $\vec{n}$, and there is no possibility of a uniform performance better than that of the set of azimuth angles $\phi_k = 2\pi(k-1)/K$ for $k = 1$ to K. It however still remains the latitude of considering nonunit vectors $\vec{a}_k$. For positivity of each M_k, one has only the option of $\|\vec{a}_k\| < 1$. Also, with an unknown phase ξ, there is no reason that could motivate in the plane $\perp \vec{n}$ a nonisotropic distribution of the norms $\|\vec{a}_k\|$ for $k = 1$ to K. So, as for the azimuths ϕ_k, there is no motivation for a choice other than the uniform constant $\|\vec{a}_k\| = a < 1$ for $k = 1$ to K. This choice leads for the classical Fisher information $F_c(\xi)$ of Eq. (39) to a form similar to that of Eq. (42) but with the noise factor α replaced by αa. With $0 < a < 1$ the effect is just as if the estimation process was accomplished at a higher noise level characterized by a smaller noise factor, consistently reducing $F_c(\xi)$, so that there is no benefit to be gained by selecting $\|\vec{a}_k\| < 1$ since $\|\vec{a}_k\| = 1$ is always more efficient. Taking measurement vectors $\vec{a}_k$ nonorthogonal to $\vec{n}$ is also similar to operating at a higher noise level characterized by a smaller noise factor reduced by the amount by which the components orthogonal to $\vec{n}$ of the $\vec{a}_k$ are reduced; so measurement vectors $\vec{a}_k$ orthogonal to $\vec{n}$ are generally better to maximize $F_c(\xi)$ of Eq. (39).

We can therefore conclude that, with quantum depolarizing noise affecting the qubit, to construct a ξ-independent POVM maximizing the classical Fisher information $F_c(\xi)$, it is not possible to do better than the set of K measurement operators $M_k = (I_2 + \vec{a}_k \cdot \vec{\sigma})/K$ proportional to projectors in $\mathcal{H}_2$, with the $\vec{a}_k$ in the plane orthogonal to $\vec{n}$ at azimuths $\phi_k = 2\pi(k-1)/K$ for $k = 1$ to K. It is preferable to take a large number K around $K = 8$ to 10 or larger, this allowing to reach the ξ-independent performance $F_c(\xi) = F_c^{\infty} = \overline{F}_c = 1 - \sqrt{1-\alpha^2}$, and this similarly at any noise factor α. This performance of $F_c(\xi)$ will usually remain below the overall maximum $F_q^{\max} = \alpha^2$ as shown by Fig. 2 (except at vanishing noise $\alpha = 1$), but there exists no possibility of another POVM doing better uniformly for any unknown phase ξ. For this reason, we can qualify this approach as defining the ξ-independent optimal POVM or measurement protocol. As in the noise-free case of Sec. III, we are now interested in an estimator with a performance matching the classical Fisher information $F_c(\xi) = F_c^{\infty}$ set by this ξ-independent optimal measurement protocol.

D. Optimizing the estimator with depolarizing noise

We know by principle that the targeted optimal performance is realized by the maximum likelihood estimator for the phase ξ, at least in the limit of a large number N of independent measurements. With the ξ-independent optimal measurement protocol determined in Sec. IV C, the probability for obtaining the measurement outcome k corresponding to operator $M_k = (I_2 + \vec{a}_k \cdot \vec{\sigma})/K$ is, following Eq. (23),

$$P(k;\xi) = \frac{1}{K}[1 + \alpha \vec{a}_k \vec{r}_1(\xi)]. \tag{45}$$

From N independent copies of the qubit state ρ_ξ, a sequence of N independent measurements are performed, yielding a number N_k of outcomes k, for $k = 1$ to K.

From such a sequence of N measurement outcomes, the likelihood $\mathcal{L}(\xi)$ for the parameter ξ follows as

$$\mathcal{L}(\xi) = \prod_{k=1}^{K} P(k;\xi)^{N_k} = \frac{1}{K^N} \prod_{k=1}^{K} [1 + \alpha \vec{a}_k \vec{r}_1(\xi)]^{N_k}. \quad (46)$$

We now seek the maximum likelihood estimator $\widehat{\xi}_{\rm ML}$ as that value of ξ maximizing $\mathcal{L}(\xi)$ in Eq. (46). When seen as a function of the K probabilities $P(k;\xi)$ from Eq. (45) considered as free variables, it is known that the likelihood $\mathcal{L}(\xi)$ of Eq. (46) is maximized by $P(k;\xi) = N_k/N$ for $k = 1$ to K. This is a standard result known for estimation of the probabilities of a multinomial distribution. However, this is not quite what we seek for our specific task of estimating the qubit phase ξ. The K variables $P(k;\xi)$ from Eq. (45), constrained to sum to 1, form a set with dimension $K - 1$, supported on a $(K - 1)$-dimensional hyperplane in $\mathbb{R}^K$. In this $(K - 1)$-dimensional set, the variable ξ, through the relations of Eq. (45) determined for the $P(k;\xi)$ by the quantum measurement, defines a one-dimensional manifold. It is over this one-dimensional manifold [and not over the whole embedding $(K - 1)$-dimensional set] that the maximization of $\mathcal{L}(\xi)$ of Eq. (46) has to be accomplished for estimating the phase ξ.

Due to the nonlinearities involved in Eq. (46), it is usually not possible to obtain an analytical resolution of the maximization of $\mathcal{L}(\xi)$ to yield a closed-form expression for the phase estimator $\widehat{\xi}_{\rm ML}$. Nevertheless, the circumstance is quite favorable for a numerical resolution. The one-dimensional maximization of $\mathcal{L}(\xi)$ has to be performed over a known bounded domain $\xi \in [0,2\pi)$, with a suitable precision on ξ determined by the root-mean-squared error $(N F_c^\infty)^{-1/2} = \sigma^\infty$ as set by the classical Fisher information F_c^∞ of Eq. (43). For each sequence of N measurements, it is thus easy to compute the resulting likelihood $\mathcal{L}(\xi)$ of Eq. (46) over a fine grid for $\xi \in [0,2\pi)$, so as to locate the maximum and maximizer yielding the value of the estimate by $\widehat{\xi}_{\rm ML}$.

Figure 4 presents the results of a numerical simulation which illustrates the operation of the maximum likelihood estimator $\widehat{\xi}_{\rm ML}$ with $K = 10$ POMV elements. The simulation for the quantum measurement basically implements N independent random draws according to the probability distribution of Eq. (45), and it counts the resulting K measurement outcomes N_k that next serve to the numerical maximization of the likelihood of Eq. (46) delivering the value of the estimate $\widehat{\xi}_{\rm ML}$. Two noise levels α are tested in Fig. 4 with $N = 30$ repeated measurements.

The results of Fig. 4 show a reasonable behavior for the maximum likelihood estimator $\widehat{\xi}_{\rm ML}$ derived here, and for its performance. We especially observe an experimental rms estimation error $\sigma_{\rm exper}$ in good agreement with the theoretical prediction $\sigma^\infty = (N F_c^\infty)^{-1/2}$ from Eq. (43) based on the analysis of the Fisher information.

In the noise-free case of Sec. III, to overcome the limitation of the two-element POVM or spin observable associated with the optimal estimator of Eq. (31) returning a value of ξ in $[0,\pi]$, as explained at the end of Sec. III C, for a determination of ξ in $[0,2\pi)$ one has to resort to a POVM with at least $K = 3$ elements. Any such POVM at $K \geqslant 3$ when constructed as described in Sec. IV C, is equally optimal as it achieves

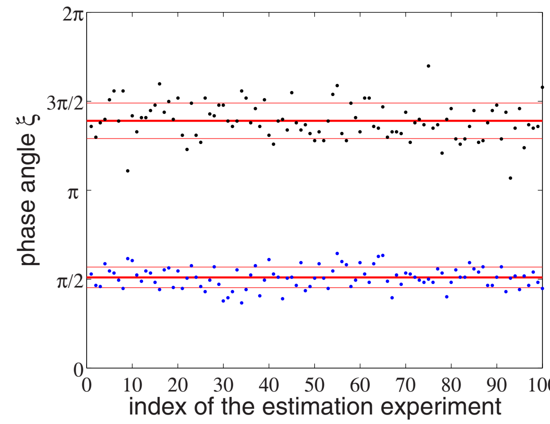

FIG. 4. Two values of the phase angle are successively used for estimation, $\xi = 0.51\pi$ with no noise at $\alpha = 1$, and $\xi = 1.39\pi$ with noise at $\alpha = 0.75$. Each of these two values of ξ is localized by a thick horizontal line, each surrounded by two thin horizontal lines at ordinates $\xi \pm \sigma^\infty$, with $\sigma^\infty = (N F_c^\infty)^{-1/2}$ the rms error theoretically predicted from Eq. (43) for the best estimator, with for $\xi = 0.51\pi$ and $\alpha = 1$ the value $\sigma^\infty \approx 0.1826$, and for $\xi = 1.39\pi$ and $\alpha = 0.75$ the value $\sigma^\infty \approx 0.3138$. Each discrete data point represents an estimate $\widehat{\xi}_{\rm ML}$ computed from $N = 30$ independent copies of the qubit state ρ_ξ measured and followed by numerical maximization of the likelihood of Eq. (46), this constituting one estimation experiment. For each of the two values of ξ a total of 100 estimation experiments have been repeated, as displayed in abscissa. Over these 100 repetitions, the experimental rms error is $\sigma_{\rm exper} = 0.1890$ when $\xi = 0.51\pi$, and $\sigma_{\rm exper} = 0.3256$ when $\xi = 1.39\pi$.

$F_c(\xi) = F_q^{\rm max} = 1$ uniformly for any ξ, as we have seen in Sec. IV C. At the same time, such a POVM will usually require the above numerical maximization of the likelihood, in place of the analytical expression of Eq. (31), to implement the maximum likelihood estimator $\widehat{\xi}_{\rm ML}$.

V. SUMMARY AND CONCLUSION

The theory of parameter estimation from a quantum state has been reviewed in Sec. II to obtain with Eqs. (3)–(6) general expressions for the quantum score L_ξ and quantum Fisher information $F_q(\xi)$ determining the ultimate best performance for this task. In Sec. III, these expressions for L_ξ and $F_q(\xi)$ have been specified in Bloch representation for parameter estimation on a qubit, with Eqs. (13)–(16) as previously derived in [20]. Next, for estimation of the phase ξ acquired by a qubit in a rotation around an arbitrary axis $\vec{n}$, the expressions of Eqs. (13)–(16) for L_ξ and $F_q(\xi)$ have been transformed to concise geometric forms, with Eqs. (18), (19), (21), and (37)–(38), not contained in [20]. Then this geometric formulation of Eqs. (18), (19), (21), and (37)–(38) served here to characterize the best strategies for qubit phase estimation, together with their performance, consistently in terms of optimal probe, optimal measurement, and optimal estimator, first with no noise and then with noise. Comparatively, Ref. [20] essentially concentrated on optimization of the input probe, for different noise models.

With no noise, in Sec. III, it was established that the optimal estimation strategy is to operate as follows. Choose an optimal pure probe ρ_0 with unit Bloch vector $\vec{r}_0$ orthogonal to the rotation axis $\vec{n}$; implement an optimal quantum measurement by measuring the spin observable $\Omega = \vec{r}_0 \cdot \vec{\sigma}$; use the optimal estimator of Eq. (31). These are optimal steps allowing together in succession to reach the ultimate best performance in qubit phase estimation, with a minimal mean-squared estimation error matching the least possible value of $1/(NF_q^{\max}) = 1/N$ when estimating from N independent measurements, at large N. It is an important property that this strategy is a ξ-independent solution, uniformly optimal for any unknown phase ξ.

With quantum noise on the qubit, we have observed that such parameter-independent optimal solutions are usually no longer possible. For arbitrary quantum noise on the qubit, it was argued in Sec. IV A that in general to achieve a performance reaching the overall maximum of the quantum Fisher information $F_q(\xi)$ requires a ξ-dependent strategy and adaptive implementation. However, for an isotropic depolarizing noise with compression factor α, we have shown that the performance $F_q(\xi)$ can be raised to its maximum $F_q^{\max} = \alpha^2$ by a ξ-independent pure input probe ρ_0 with a unit Bloch vector $\vec{r}_0$ orthogonal to the rotation axis $\vec{n}$. There is however no ξ-independent measurement protocol able to raise the classical Fisher information $F_c(\xi)$ at the level $F_q^{\max} = \alpha^2$ of the maximal quantum Fisher information uniformly for any unknown phase ξ. We have characterized in Sec. IV C a ξ-independent optimal measurement protocol and the uniform maximum $F_c^{\infty} = 1 - \sqrt{1-\alpha^2} \leqslant F_q^{\max} = \alpha^2$ it can reach for the classical Fisher information. This measurement protocol is a generalized measurement with a large number K (starting around 8 to 10) of POVM elements; and there is no spin observable or von Neumann projective measurement that would suffice to achieve this same performance. Finally, a maximum likelihood estimator associated with this ξ-independent optimal measurement is able to achieve the minimal mean-squared estimation error saturating the classical Cramér-Rao inequality, as analyzed in Sec. IV D.

In principle, a comparable approach examining in conjunction optimization of the input probe for maximum quantum Fisher information, of the quantum measurement and of the estimator, could be developed for estimation of multidimensional parameters and/or estimation on quantum systems with dimension higher than the dimension $D = 2$ of the qubit. However, for multiple parameters generally optimality cannot be achieved simultaneously for all parameters because the optimal quantum measurements for them typically do not commute and are therefore incompatible [7]. Also, in such higher-dimensional conditions, the analytical characterization of the optimal performances and strategies are generally much more difficult to handle. By contrast, for the qubit which is a fundamental system of quantum information, we have seen here that the situation of single-parameter estimation and its optimization is analytically tractable to a large extent, especially with quantum noise on the qubit, and represents in this respect a useful reference.

We have considered here the very common approach to estimation which consists of gathering the results of N independent measurements so as to collect more information to serve the estimation task. This mode of operation is especially convenient for practical implementation and for analytical treatment. For classical estimation, when estimating from N successive measurements, independent measurements are generally preferable to correlated measurements, because correlation tends to replace useful original information with unhelpful repetition or redundancy. For quantum estimation, this is no longer necessarily true, thanks to an unparalleled type of correlation under the form of quantum entanglement. Instead of reproducing the probe state ρ_0 as N independent copies forming the N-fold separable state $\rho_0^{\otimes N}$, a composite probe could be prepared in an N-fold entangled state. These represent N-qubit composite states which are more difficult to prepare and control experimentally, and to handle analytically. Yet such entangled composite states can provide in definite conditions enhanced performance, typically with a mean-squared estimation error decreasing as $1/N^2$ (designated as the Heisenberg limit) instead of $1/N$ with N independent probes (the standard quantum limit) [25,33,34]. Such quantum enhancement does not occur for all parametric dependence [35], yet it can occur for estimation of parameters from unitary transformations [25,33–35]. However, such quantum enhancement is very fragile, and a small amount of depolarizing noise as considered here is sufficient to ruin the $1/N^2$ performance and return it to $1/N$ in the asymptotic limit of large N [35–37]. The analysis of quantum estimation with depolarizing noise we have carried out here with N independent measurements is thus an essential reference for estimation in realistic noisy conditions. Intermediate configurations, with many repetitions of measurements on composite quantum states of finite size, may also provide interesting alternatives for enhanced performance in estimation. Such proposals have been investigated recently for instance in [38,39] for parameter estimation from multiple-photon states. When quantum noise is taken into account, such intermediate configurations can be specially relevant. Since composite entangled states of asymptotic size N are no longer capable of enhancing the performance above the standard quantum limit of separable states, maximal enhancement of the performance may possibly be obtained with intermediate configurations of finite size. Optimization of such intermediate configurations in noisy conditions can be tackled with an approach similar to that followed here, especially with the quantum and classical Fisher information for setting the ultimate best performance, and it remains open for further research.

APPENDIX: NOISE LEVEL ESTIMATION

We consider the situation where the unitary transformation U_ξ is absent and only the depolarizing noise channel is present and acts on the input probe ρ_0. We apply the approach of the paper to derive the optimal strategy and its performance for estimation of the depolarizing noise factor α which now plays the role of the unknown parameter in place of ξ. The input probe ρ_0 is transformed by the depolarizing noise channel into the output noisy state denoted ρ_α. From Eq. (14), in Bloch representation, we now have for estimating α the quantum Fisher information

$$F_q(\alpha) = \frac{(\vec{r}_\alpha\, \partial_\alpha \vec{r}_\alpha)^2}{1 - \vec{r}_\alpha^{\,2}} + (\partial_\alpha \vec{r}_\alpha)^2. \qquad \text{(A1)}$$

The noisy Bloch vector is $\vec{r}_\alpha = \alpha\vec{r}_0$, so that $\partial_\alpha \vec{r}_\alpha = \vec{r}_0$, leading in Eq. (A1) to

$$F_q(\alpha) = \frac{\|\vec{r}_0\|^2}{1-\alpha^2\|\vec{r}_0\|^2}. \tag{A2}$$

Maximization of the quantum Fisher information $F_q(\alpha)$ of Eq. (A2) is obtained simply by $\|\vec{r}_0\| = 1$, i.e., by a pure input probe ρ_0 with an arbitrary orientation for $\vec{r}_0$ in $\mathbb{R}^3$, to achieve the maximum

$$F_q^{\max} = \frac{1}{1-\alpha^2} \tag{A3}$$

in Eq. (A2). To obtain an optimal measurement it is enough to measure a spin observable $\Omega = \vec{a}\cdot\vec{\sigma}$. As in Sec. III B, the resulting two measurement outcomes ± 1 have probabilities

$$P(\pm 1) = \tfrac{1}{2}(1 \pm \vec{r}_\alpha \vec{a}) = \tfrac{1}{2}(1 \pm \alpha\vec{r}_0\vec{a}), \tag{A4}$$

associated, from Eq. (24), with the classical Fisher information

$$F_c(\alpha) = \frac{[\partial_\alpha P(+1)]^2}{P(+1)} + \frac{[\partial_\alpha P(-1)]^2}{P(-1)} = \frac{(\vec{r}_0\vec{a})^2}{1-\alpha^2(\vec{r}_0\vec{a})^2}. \tag{A5}$$

Maximization of the classical Fisher information $F_c(\alpha)$ of Eq. (A5) is obtained by $(\vec{r}_0\vec{a})^2 = 1$, i.e., by $\vec{a} = \pm\vec{r}_0$, achieving for $F_c(\alpha)$ the maximum $F_q^{\max}$ of Eq. (A3). This establishes the observable $\Omega = \vec{r}_0\cdot\vec{\sigma}$ (or $\Omega = -\vec{r}_0\cdot\vec{\sigma}$) as an optimal measurement, uniformly for any α.

For an optimal estimator for α, we turn to the maximum likelihood estimator as in Sec. III C, which follows as

$$\widehat{\alpha}_{\rm ML} = 2\frac{N_+}{N} - 1, \tag{A6}$$

unbiased and with mean-squared error $\langle(\widehat{\alpha}_{\rm ML} - \alpha)^2\rangle = (1-\alpha^2)/N = 1/[NF_c(\alpha)] = 1/[NF_q^{\max}]$, proving that $\widehat{\alpha}_{\rm ML}$ of Eq. (A6) is an optimal estimator for any number N of measurements. This defines the optimal probe, optimal measurement, and optimal estimator, with their maximal performance, for the factor α of a qubit depolarizing noise.

[1] C. W. Helstrom, The minimum variance of estimates in quantum signal detection, IEEE Trans. Inf. Theory **14**, 234 (1968).
[2] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).
[3] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).
[4] *Quantum State Estimation*, Lecture Notes in Physics Vol. 649, edited by M. G. A. Paris and J. Řeháček (Springer, Berlin, 2004).
[5] M. Hayashi, *Quantum Information Theory* (Springer, Berlin, 2006).
[6] M. G. A. Paris, Quantum estimation for quantum technology, Int. J. Quantum. Inf. **07**, 125 (2009).
[7] O. E. Barndorff-Nielsen and R. D. Gill, Fisher information in quantum statistics, J. Phys. A **33**, 4481 (2000).
[8] A. Luati, Maximum Fisher information in mixed state quantum systems, Ann. Stat. **32**, 1770 (2004).
[9] P. Gibilisco, F. Hiai, and D. Petz, Quantum covariance, quantum Fisher information, and the uncertainty relations, IEEE Trans. Inf. Theory **55**, 439 (2009).
[10] P. Facchi, R. Kulkarni, V. I. Man'ko, G. Marmo, E. C. G. Sudarshan, and F. Ventriglia, Classical and quantum Fisher information in the geometrical formulation of quantum mechanics, Phys. Lett. A **374**, 4801 (2010).
[11] M. Walter and J. M. Renes, Lower bounds for quantum parameter estimation, IEEE Trans. Inf. Theory **60**, 8007 (2014).
[12] Z. Hradil, Quantum-state estimation, Phys. Rev. A **55**, R1561 (1997).
[13] G. M. D'Ariano, Universal quantum estimation, Phys. Lett. A **268**, 151 (2000).
[14] Z. Hradil, J. Summhammer, G. Badurek, and H. Rauch, Reconstruction of the spin state, Phys. Rev. A **62**, 014101 (2000).
[15] R. Blume-Kohout, Optimal, reliable estimation of quantum states, New J. Phys. **12**, 043034 (2010).
[16] J. Řeháček, Z. Hradil, and M. Ježek, Iterative algorithm for reconstruction of entangled states, Phys. Rev. A **63**, 040303(R) (2001).
[17] D. Brivio, S. Cialdi, S. Vezzoli, B. T. Gebrehiwot, M. G. Genoni, S. Olivares, and M. G. A. Paris, Experimental estimation of one-parameter qubit gates in the presence of phase diffusion, Phys. Rev. A **81**, 012305 (2010).
[18] E. Tesio, S. Olivares, and M. G. A. Paris, Optimized qubit phase estimation in noisy quantum channels, Int. J. Quantum Inf. **09**, 379 (2011).
[19] M. G. Genoni, S. Olivares, and M. G. A. Paris, Optical Phase Estimation in the Presence of Phase Diffusion, Phys. Rev. Lett. **106**, 153603 (2011).
[20] F. Chapeau-Blondeau, Optimized probing states for qubit phase estimation with general quantum noise, Phys. Rev. A **91**, 052310 (2015).
[21] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).
[22] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory* (Prentice-Hall, Englewood Cliffs, NJ, 1993).
[23] S. L. Braunstein and C. M. Caves, Statistical Distance and the Geometry of Quantum States, Phys. Rev. Lett. **72**, 3439 (1994).
[24] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).
[25] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum-enhanced measurements: Beating the standard quantum limit, Science **306**, 1330 (2004).
[26] F. Chapeau-Blondeau, Optimization of quantum states for signaling across an arbitrary qubit noise channel with minimum-error detection, IEEE Trans. Inf. Theory **61**, 4500 (2015).
[27] A. Fujiwara, Quantum channel identification problem, Phys. Rev. A **63**, 042304 (2001).
[28] S. Alipour and A. T. Rezakhani, Extended convexity of quantum Fisher information in quantum metrology, Phys. Rev. A **91**, 042104 (2015).
[29] M. A. Armen, J. K. Au, J. K. Stockton, A. C. Doherty, and H. Mabuchi, Adaptive Homodyne Measurement of Optical Phase, Phys. Rev. Lett. **89**, 133602 (2002).

[30] A. Fujiwara, Strong consistency and asymptotic efficiency for adaptive quantum estimation problems, J. Phys. A **39**, 12489 (2006).
[31] R. Okamoto, M. Iefuji, S. Oyama, K. Yamagata, H. Imai, A. Fujiwara, and S. Takeuchi, Experimental Demonstration of Adaptive Quantum State Estimation, Phys. Rev. Lett. **109**, 130404 (2012).
[32] M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, Cambridge, England, 2013).
[33] G. M. D'Ariano, P. Lo Presti, and M. G. A. Paris, Using Entanglement Improves the Precision of Quantum Measurements, Phys. Rev. Lett. **87**, 270404 (2001).
[34] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum Metrology, Phys. Rev. Lett. **96**, 010401 (2006).
[35] Z. Ji, G. Wang, R. Duan, Y. Feng, and M. Ying, Parameter estimation of quantum channels, IEEE Trans. Inf. Theory **54**, 5172 (2008).
[36] V. Giovannetti, S. Lloyd, and L. Maccone, Advances in quantum metrology, Nat. Photon. **5**, 222 (2011).
[37] R. Demkowicz-Dobrzański, J. Kolodyński, and M. Gutá, The elusive Heisenberg limit in quantum-enhanced metrology, Nat. Commun. **3**, 1063 (2012).
[38] D. W. Berry, B. L. Higgins, S. D. Bartlett, M. W. Mitchell, G. J. Pryde, and H. M. Wiseman, How to perform the most accurate possible phase measurements, Phys. Rev. A **80**, 052114 (2009).
[39] T. A. Wheatley, M. Tsang, I. R. Petersen, and E. H. Huntington, Improved mirror position estimation using resonant quantum smoothing, Eur. Phys. J. Quantum Technol. **2**, 13 (2015).