Local asymptotic normality for qubit states

Mădălin Guţă[1] and Jonas Kahn[2]

[1]*University of Nijmegen, Toernooiveld 1, Postbus 9010, 6500 GL Nijmegen, The Netherlands*
[2]*Université Paris-Sud 11, Département de Mathématiques, Bât 425, 91405 Orsay Cedex, France*

We consider n identically prepared qubits and study the asymptotic properties of the joint state $\rho^{\otimes n}$. We show that for all individual states, ρ situated in a local neighborhood of size $1/\sqrt{n}$ of a fixed state $\rho^{\mathbf{0}}$, the joint state converges to a displaced thermal equilibrium state of a quantum harmonic oscillator. The precise meaning of the convergence is that there exists physical transformations T_n (trace preserving quantum channels) which map the qubits states asymptotically close to their corresponding oscillator state, uniformly over all states in the local neighborhood. A few consequences of the main result are derived. We show that the optimal joint measurement in the Bayesian setup is also optimal within the point-wise approach. Moreover, this measurement converges to the heterodyne measurement which is the optimal joint measurement of position and momentum for the quantum oscillator. A problem of local state discrimination is solved using local asymptotic normality.

I. INTRODUCTION

The quantum measurement theory brings together the quantum world of wave functions and incompatible observables with the classical world of random phenomena studied in probability and statistics. These fields have come ever closer due to the technological advances making it possible to perform measurements on individual quantum systems. Indeed, the engineering of a novel quantum state is typically accompanied by a verification procedure through which the state, or some aspect of it, is reconstructed from measurement data [1].

An important example of such a technique is that of quantum homodyne tomography in quantum optics [2]. This allows the estimation with arbitrary precision of the whole density matrix [3–6] of a monochromatic beam of light by repeatedly measuring a sufficiently large number of identically prepared beams [1,7,8].

In contrast to this "semiclassical" situation in which one fixed measurement is performed repeatedly on independent systems, the state estimation problem becomes more "quantum" if one is allowed to consider *joint measurements* on n identically prepared systems with the joint state $\rho^{\otimes n}$. It is known [9] that in the case of unknown *mixed* states ρ, joint measurements perform strictly better than separate measurements, in the sense that the asymptotical convergence rate of the optimal estimator $\hat{\rho}_n$ to ρ goes in both cases as $C/\sqrt{n}$ with a strictly smaller constant C in the case of joint measurements.

Let us look at this problem in more detail: We dispose of a number of n copies of an unknown state ρ and the task is to estimate ρ as well as possible. The first step is to specify a cost function $d(\hat{\rho}_n, \rho)$ which quantifies the deviation of the estimator $\hat{\rho}_n$ from the true state. Then, one tries to devise a measurement and an estimator which minimize the mean cost or risk in statistics jargon:

$$R(\rho, \hat{\rho}_n) := \langle d(\hat{\rho}_n(X), \rho) \rangle,$$

with the average taken over the measurement results X. Since this quantity still depends on the unknown state, one may choose a Bayesian approach and try to optimize the average risk with respect to some prior distribution π over the states

$$R_{n,\pi} = \int R(\rho, \hat{\rho}_n) \pi(d\rho).$$

Results of this type have been obtained in both the pure state case [10–17] and the mixed state case [18–24]. However, most of these papers use methods of group theory which depend upon the symmetry of the prior distribution and the form of the cost function, and thus cannot be extended to arbitrary priors.

In the point-wise approach [9,25–29], one tries to minimize $R(\rho, \hat{\rho}_n)$ for each fixed ρ. We can argue that even for a completely unknown state, as n becomes large, the problem ceases to be global and becomes a local one as the error in estimating the state parameters is of the order $1/\sqrt{n}$. For this reason, it makes sense to parametrize the state as $\rho := \rho(\theta)$ with θ belonging to some set in $\mathbb{R}^k$ and to replace the original cost with its quadratic approximation at θ:

$$d(\theta, \hat{\theta}_n) = (\theta - \hat{\theta}_n)^T G(\theta)(\theta - \hat{\theta}_n),$$

where G is a $k \times k$ positive real symmetric weight matrix.

Although seemingly different, the two approaches can be compared [30] and, in fact, for large n the prior distribution π of the Bayesian approach should become increasingly irrelevant and the optimal Bayesian estimator should be close to the maximum likelihood estimator. An instance of this asymptotic equivalence is proven in subsection VII A.

In this paper, we change the perspective and, instead of trying to devise optimal measurements and estimators for a particular statistical problem, we concentrate our attention on the *family* of joint states $\rho(\theta)^{\otimes n}$ which is the primary "carrier" of statistical information about θ. As suggested by the locality argument sketched above, we consider a neighborhood of size $1/\sqrt{n}$ around a fixed but arbitrary parameter θ_0, whose points can be written as $\theta = \theta_0 + \mathbf{u}/\sqrt{n}$ with $\mathbf{u} \in \mathbb{R}^k$—the "local parameter"—obtained by zooming into the smaller and smaller balls by a factor of $\sqrt{n}$. Briefly, the

principle of *local asymptotic normality* says that for large n, the local family

$$\rho_n^{\mathbf{u}} := \rho(\theta_0 + \mathbf{u}/\sqrt{n})^{\otimes n}, \quad \|\mathbf{u}\| < C,$$

converges to a family of displaced Gaussian states $\phi^{\mathbf{u}}$ of a quantum system consisting of a number of coupled quantum and classical harmonic oscillators.

The term local asymptotic normality comes from mathematical statistics [31], where the following result holds. We are given independent variables $X_1, \ldots, X_n \in \mathcal{X}$ drawn from the same probability distribution $P^{\theta_0 + \mathbf{u}/\sqrt{n}}$ over $\mathcal{X}$ depending upon the unknown parameter $\mathbf{u} \in \mathbb{R}^k$. Then, the statistical information contained in our data is asymptotically identical with the information contained in a *single* normally distributed $Y \in \mathbb{R}^k$ with mean $\mathbf{u}$ and variance $I(\theta_0)^{-1}$, the inverse Fisher information matrix. This means that for any statistical problem, we can replace the original data, $X_1, \ldots, X_n \in \mathcal{X}$, by the simpler Gaussian one, Y, with the same asymptotic results!

For the sake of clarity, let us consider the case of qubits with states parametrized by their Bloch vectors $\rho(\vec{r}) = 1/2(\mathbf{1} + \vec{r}\vec{\sigma})$ where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the Pauli matrices. Define now the two-dimensional family of identical spin states obtained by rotating the Bloch vector $\vec{r}_0 = (0, 0, 2\mu - 1)$ around an axis in the x-y plane

$$\rho_n^{\mathbf{u}} = \left[U\left(\frac{\mathbf{u}}{\sqrt{n}}\right) \begin{pmatrix} \mu & 0 \\ 0 & 1-\mu \end{pmatrix} U\left(\frac{\mathbf{u}}{\sqrt{n}}\right)^* \right]^{\otimes n}, \quad \mathbf{u} \in \mathbb{R}^2, \quad (1.1)$$

with unitary $U(\mathbf{v}) := \exp[i(v_x \sigma_x + v_y \sigma_y)]$ and $1/2 < \mu \leq 1$.

Consider now a quantum harmonic oscillator with position and momentum operators Q and P on $L^2(\mathbb{R})$ satisfying the commutation relations $[Q, P] = i\mathbf{1}$. We denote by $\{|n\rangle, n \geq 0\}$ the eigenbasis of the number operator, and define the thermal equilibrium state

$$\phi^0 = (1 - p) \sum_{k=0}^{\infty} p^k |k\rangle\langle k|,$$

where $p = (1 - \mu)/\mu$. We translate the state ϕ^0 by using the displacement operators $D(\mathbf{z}) = \exp(\mathbf{z}a^* - \bar{\mathbf{z}}a)$ with $\mathbf{z} \in \mathbb{C}$, which map the ground state $|0\rangle$ into the coherent state $|\mathbf{z}\rangle$:

$$\phi^{\mathbf{u}} := D(\sqrt{2\mu - 1}\alpha_{\mathbf{u}})\phi^0 D(\sqrt{2\mu - 1}\alpha_{\mathbf{u}})^*, \quad (1.2)$$

where $\alpha_{\mathbf{u}} := -u_y + iu_x$.

Theorem(1.1). *Let $\rho_n^{\mathbf{u}}$ be the family of states* (1.1) *on the Hilbert space* $(\mathbb{C}^2)^{\otimes n}$ *and $\phi^{\mathbf{u}}$ the family* (1.2) *of displaced thermal equilibrium states of a quantum oscillator. Then, for each n, there exist quantum channels (trace preserving completely positive maps)*

$$T_n : M((\mathbb{C}^2)^{\otimes n}) \to \mathcal{T}(L^2(\mathbb{R})),$$

$$S_n : \mathcal{T}(L^2(\mathbb{R})) \to M((\mathbb{C}^2)^{\otimes n}), \quad (1.3)$$

with $\mathcal{T}(L^2(\mathbb{R}))$ the trace-class operators, such that

$$\lim_{n \to \infty} \sup_{\mathbf{u} \in I^2} \|\phi^{\mathbf{u}} - T_n(\rho_n^{\mathbf{u}})\|_1 = 0,$$

$$\lim_{n \to \infty} \sup_{\mathbf{u} \in I^2} \|\rho_n^{\mathbf{u}} - S_n(\phi^{\mathbf{u}})\|_1 = 0. \quad (1.4)$$

for an arbitrary bounded interval $I \subset \mathbb{R}$.

Let us make a few comments on the significance of the above result.

(i) The "convergence" (1.4) of the qubit states holds in a strong way (uniformly in $\mathbf{u}$) with direct statistical and physical interpretation. Indeed, the channels T_n and S_n represent physical transformations which are analogues of randomizations of classical data [31]. The meaning of Eq. (1.4) is that the two quantum models are asymptotically equivalent from a statistical point of view.

(ii) Indeed, for any measurement M on $L^2(\mathbb{R})$, we can construct the measurement $M \circ T_n$ on the spin states by first mapping them to the oscillator space and then performing M. Then, the optimal solution of any statistical problem concerning the states $\rho_n^{\mathbf{u}}$ can be obtained by solving the same problem for $\phi^{\mathbf{u}}$ and pulling back the optimal measurement M as above. We illustrate this in Sec. VII for the estimation problem and for hypothesis testing.

(iii) The proposed technique may be useful for applications in the domain of coherent spin states [32] and squeezed spin states [33]. Indeed, it has been known since Dyson [34] that n spin-1/2 particles prepared in the spin-up state $|\uparrow\rangle^{\otimes n}$ behave asymptotically as the ground state of a quantum oscillator when considering the fluctuations of properly normalized total spin components in the directions orthogonal to z. Our Theorem extends this to spin directions making an "angle" $\mathbf{u}/\sqrt{n}$ with the z axis, as well as to mixed states, and gives a quantitative expression to heuristic pictures common in the physics literature (see Sec. III). We believe that a similar approach can be followed in the case of spin squeezed states and continuous time measurements with feedback control [35].

The next section gives an introduction to the statistical ideas motivating our work. In Sec. III, we give a heuristic picture of our main result based on the total spin vector representation of spin coherent states familiar in the physics literature.

The proof of Theorem 1.1 extends over the Secs. IV–VI, and uses methods of group theory and some ideas from [29,36–38].

Section VII describes a few applications of our main result. In subsection VII A, we compute the local asymptotic minimax risk for the statistical problem of qubit state estimation. An estimation scheme which achieves this risk asymptotically is optimal in the point-wise approach. We show that this figure of merit coincides with the risk of the heterodyne measurement, and that it is achieved by the optimal Bayesian measurement for the $SU(2)$-invariant prior [24,29]. This proves the asymptotic equivalence of the Bayesian and point-wise approaches.

In subsection VII B, we continue the investigation of the optimal Bayesian measurement, and show that it converges locally to the heterodyne measurement on the oscillator which is a optimal joint measurement of position and momentum [39].

Another application is the problem discriminating between two states $\rho_n^{\pm\mathbf{u}}$ which asymptotically converge to each other at rate of $1/\sqrt{n}$. In this case, the optimal measurement for the parameter $\mathbf{u}$ is not optimal for the testing problem,

showing in particular that the quantum Fisher information in general does not encode all statistical information.

II. LOCAL ASYMPTOTIC NORMALITY IN STATISTICS AND ITS EXTENSION TO QUANTUM MECHANICS

In this section, we introduce some statistical ideas which provide the motivation for deriving the main result.

Quantum statistical problems can be seen as a game between a statistician, or physicist in our case, and Nature. The latter tries to codify some information by preparing a quantum system in a state which depends on some parameter **u** unknown to the former. The physicist tries to guess the value of the parameter by devising measurements and estimators which work well for *all* choices of parameters that Nature may make. In a Bayesian setup, Nature may build her strategy by randomly choosing a state with some prior distribution. In order to solve the problem, the physicist is allowed to use the laws of quantum physics as well as those of classical stochastics and statistical inference. In particular, he may transform the quantum state by applying an arbitrary quantum channel T and obtain a new family $T(\rho^{\mathbf{u}})$. In general, such a transformation goes with a loss of information so one should have a good reason to do it, but there are nontrivial situations when no such loss occurs [40]; that is when there exists a channel S which reverses the effect of T *restricted* to the states of interest $S(T(\rho^{\mathbf{u}})) = \rho^{\mathbf{u}}$. If this is the case, we consider the two families of states $\rho^{\mathbf{u}}$ and $T(\rho^{\mathbf{u}})$ as statistically equivalent.

In statistics, such transformations are called *randomizations*, and a useful particular example is a *statistic*, which is just a function of the data we want to analyze. When this statistic contains all information about the unknown parameter, we say that it is sufficient; because knowing the value of this statistic alone suffices, and given this information the rest of the data is useless. For example, if $X_1, \dots X_n \in \{0, 1\}$ are results of independent coin tosses with a biased coin, then $\bar{X} = 1/n \Sigma_i X_i$ is a sufficient statistic and may be used for any statistical decision without loss of efficiency.

Quantum randomizations through quantum channels allow us to compare seemingly different families of states, and thus opens the possibility of solving a particular problem by casting it in a more familiar setting. The example of this paper is that of state estimation for n identical copies of a state which can be cast *asymptotically* into the problem of estimating the center of a quantum Gaussian which has a rather simple solution [39]. The term "asymptotically" means that for large n we can find quantum channels T_n, S_n, which almost map the families of states into each other as in Eq. (1.4).

The second main idea that we want to introduce is that of local asymptotic normality. Back in the coin toss example, we have that $\bar{X}$ is a good estimator of the probability μ of obtaining a 1, and by the Central Limit Theorem the error $\bar{X} - \mu$ asymptotically has a Gaussian distribution

$$\sqrt{n}(\bar{X} - \mu) \rightsquigarrow N(0, 1/\mu(1 - \mu)),$$

in particular the mean error is $\langle (\bar{X} - \mu)^2 \rangle = 1/(n\mu(1 - \mu))$. Now, if for each n the unknown parameter μ is restricted to

a local neighborhood of a fixed μ_0 of size $1/\sqrt{n}$, one might expect an improvement in the error because we know more about the parameter and we can use that information to built better estimators. However, this is not entirely true. Indeed, if we write $\mu = \mu_0 + u/\sqrt{n}$ then the estimator of the local parameter u is

$$\hat{u}_n = \sqrt{n}(\bar{X} - \mu_0) \rightsquigarrow N(u, 1/\mu_0(1 - \mu_0)),$$

which says that the problem of estimating μ in the local parameter model is as difficult as the original problem, i.e., the variance of the estimator is the same. The reason for this is that the additional information about the location of the parameter is nothing new as we could guess that directly from the data with very high probability. Thus, without changing the difficulty of the original problem, we can look at it locally and then we see that it transforms into that of estimating the center of a Gaussian with fixed variance $N(u, 1/\mu_0(1 - \mu_0))$, which is a classical statistical problem.

In general, we can formulate the following principle: Given $X_1, \dots, X_n \in \mathcal{X}$ independent with distribution $P^{\theta_0 + \mathbf{u}/\sqrt{n}}$ depending smoothly on the unknown parameter $\mathbf{u} \in \mathbb{R}^k$, then asymptotically this model is statistically equivalent (there exist explicit randomizations in both directions) with that of a single draw $Y \in \mathbb{R}^k$ from the Gaussian distribution $N(\mathbf{u}, I(\theta_0)^{-1})$ with fixed variance equal to the inverse of the Fisher information matrix [31].

In the quantum case, we replace the randomizations by quantum channels and the Gaussian limit model by its quantum equivalent which in the simplest case is a family of displaced thermal states of a quantum oscillator (see Theorem 1.1), but in general is a Gaussian state on a number of coupled quantum and classical oscillators, with canonical variables satisfying general commutation relations [41].

A simple extension of Theorem 1.1 is obtained by adding an additional local parameter $t \in \mathbb{R}$ for the density matrix eigenvalues such that $\mu = \mu_0 + t/\sqrt{n}$. This leads to a Gaussian limit model in which we are given a quantum oscillator is in state $\phi^{\mathbf{u}}$ and additionally, a classical Gaussian variable with distribution $N(t, 1/\mu_0(1 - \mu_0))$. The meaning of this quantum-classical coupling is the following: Asymptotically the problem of estimating the eigenvalues decouples from that of estimating the direction of the Bloch vector and becomes a *classical* statistical problem (identical with the coin toss discussed above), while that of estimating the direction remains quantum and converges to the estimation of a Gaussian state of a quantum oscillator. We note that this decoupling has been also observed in [24,29].

III. BIG BALL PICTURE OF COHERENT SPIN STATES

In this section, we give a heuristic argument for why Theorem (1.1) holds which will guide our intuition in later computations.

It is customary to represent the state of a two-dimensional quantum system by a vector $\vec{r}$ in the Bloch sphere such that the corresponding density matrix is
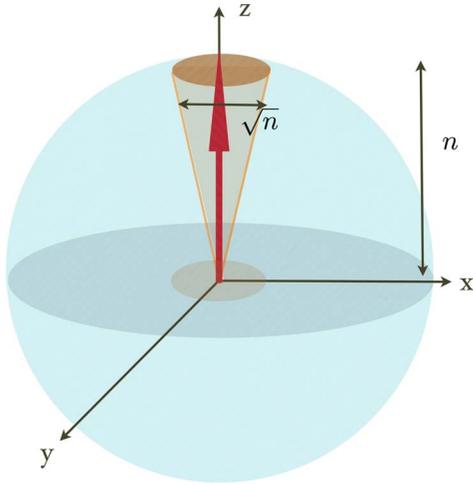
FIG. 1. (Color online) Quasiclassical representation of n spin-up qubits.

FIG. 2. (Color online) Rotated coherent state of n qubits.

$$\rho = \frac{1}{2}(1 + \vec{r}\vec{\sigma}) = \frac{1}{2}(1 + r_x\sigma_x + r_y\sigma_y + r_z\sigma_z),$$

where σ_i represent the Pauli matrices and satisfy the commutation relations $[\sigma_i,\sigma_j] = 2i\epsilon_{ijk}\sigma_k$. In particular if $\vec{r} = (0,0,\pm1)$ then the state is given by the spin-up $|\uparrow\rangle$ and, respectively, spin-down $|\downarrow\rangle$ basis vectors of $\mathbb{C}^2$, and the z component of the spin σ_z takes value ±1. As for the x and y spin components, each one may take the values ±1 with equal probabilities, such that on average $\langle\sigma_x\rangle = \langle\sigma_y\rangle = 0$ but the variances are $\langle\sigma_x^2\rangle = \langle\sigma_y^2\rangle = 1$. Moreover, σ_x and σ_y do not commute and thus cannot be measured simultaneously.

What happens with the Bloch sphere picture when we have more spins? Consider for the beginning n identical spins prepared in a coherent spin-up state $|\uparrow\rangle^{\otimes n}$, then we can think of the whole as a single-spin system and define the global observables $L_i^{(n)} = \Sigma_{k=1}^n \sigma_i^{(k)}$ for $i \in x,y,z$, where $\sigma_i^{(k)}$ is the spin component in the direction i of the k's spin. Intuitively, we can represent the joint state by a vector of length n pointing to the north pole of a large sphere as shown in Fig. 1. However, due to the quantum character of the spin observables, the x and y components cannot be equal to zero, and it is more instructive to think in terms of a vector whose tip lies on a small blob of the size of the uncertainties in x and y, sitting on the top of the sphere. Exactly how large is this blob? By using the Central Limit Theorem, we conclude that in the limit $n \to \infty$ the distribution of the "fluctuation operator"

$$S_x^{(n)} := \frac{1}{\sqrt{2n}}L_x^{(n)} = \frac{1}{\sqrt{2n}}\sum_{k=1}^n \sigma_x^{(k)},$$

converges to a $N(0,1/2)$ Gaussian, that is $\langle S_x\rangle = 0$ and $\langle S_x^2\rangle \approx 1/2$, and similarly for the component $S_y^{(n)}$. The width of the blob is thus of the order of $\sqrt{n}$ in both the x and y directions.

Now, the two fluctuations do not commute with each other

$$[S_x^{(n)}, S_y^{(n)}] = \frac{i}{n}L_z^{(n)} \approx i\mathbf{1}, \tag{3.1}$$

which is the well-known commutation relation for canonical variables of the quantum oscillator. In fact, the quantum extension of the Central Limit Theorem [36] makes this more precise

$$\lim_{n\to\infty} {}^{\otimes n}\langle\uparrow|\prod_{k=1}^p S_{i_k}^{(n)}|\uparrow\rangle^{\otimes n} = \langle\Omega,\prod_{k=1}^p X_{i_k}\Omega\rangle, \quad \forall\, i_k \in \{x,y\},$$

where $X_x := Q$ and $X_y := P$ satisfy $[Q,P] = i\mathbf{1}$ and Ω is the ground state of the oscillator.

The above description is not new in physics and goes back to Dyson's theory of spin-wave interaction [34]. More recently squeezed spin states [33] for which the variances $\langle S_x^2\rangle$ and $\langle S_y^2\rangle$ of spin variables are different, have been found to have important applications in various fields, such as magnetometry [35], and entanglement between many particles [42]. The connection with such applications will be discussed in more detail in Sec. VII.

We now rotate all spins by the same small angle for each particle, as shown in Fig. 2. As we will see, it makes sense to scale the angle by the factor of $1/\sqrt{n}$, i.e., to consider

$$\psi_n^{\mathbf{u}} = \left[\exp\left(\frac{i}{\sqrt{n}}(u_x\sigma_x + u_y\sigma_y)\right)|\uparrow\rangle\right]^{\otimes n}, \quad \mathbf{u} \in \mathbb{R}^2.$$

Indeed, for such angles, the z component of the vector will change by a small quantity of the order $\sqrt{n} \ll n$ so the commutation relations (3.1) remain the same, while the uncertainty blob will just shift its center such that the new averages of the renormalized spin components are $\langle S_x^{(n)}\rangle \approx -\sqrt{2}u_y$ and $\langle S_y^{(n)}\rangle \approx \sqrt{2}u_x$. All in all, the spins state converges to the coherent state $|\alpha_{\mathbf{u}}\rangle$ of the oscillator, where $\alpha_{\mathbf{u}} = (-u_y + iu_x) \in \mathbb{C}$ and, in general,
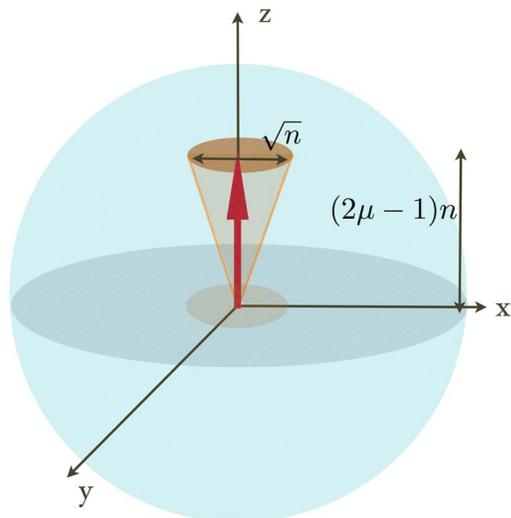
FIG. 3. (Color online) Quasiclassical representation of n qubit mixed states.

$$|\alpha\rangle := \exp(-|\alpha|^2/2)\sum_{j=0}^{\infty}\frac{\alpha^{\mathbf{j}}}{\sqrt{j!}}|j\rangle,$$

with $|j\rangle$ representing the j's energy level.

We consider now the case of qubits in an individual mixed state $\mu|\uparrow\rangle\langle\uparrow|+(1-\mu)|\downarrow\rangle\langle\downarrow|$ with $<1/2\mu<1$. Then, the "length" of L_z is $n(2\mu-1)$, but the size of the blob is the same (see Fig. 3). However, the commutation relations of S_x and S_y do not reproduce those of the harmonic oscillator, and we need to renormalize the spin as

$$S_x^{(n)} := \frac{1}{\sqrt{2(2\mu-1)n}}L_x, \quad S_y^{(n)} := \frac{1}{\sqrt{2(2\mu-1)n}}L_y.$$

The limit state will be a Gaussian state of the quantum oscillator with variance $\langle Q^2\rangle=\langle P^2\rangle=\frac{1}{2(2\mu-1)}<\frac{1}{2}$, that is a thermal equilibrium state

$$\phi^{\mathbf{0}} = (1-p)\sum_{k=0}^{\infty}p^k|k\rangle\langle k|, \quad p = \frac{1-\mu}{\mu}.$$

Finally, the rotation by $\exp\left(\frac{i}{\sqrt{n}}(u_x\sigma_x+u_y\sigma_y)\right)$ produces a displacement of the thermal state such that $\langle Q\rangle=-\sqrt{2}(2\mu-1)u_y$ and $\langle P\rangle=\sqrt{2}(2\mu-1)u_x$.

IV. LOCAL ASYMPTOTIC NORMALITY FOR MIXED QUBIT STATES

We now give a rigorous formulation of the heuristics presented in the previous section. Let

$$\rho^{\mathbf{0}} = \begin{pmatrix} \mu & 0 \\ 0 & 1-\mu \end{pmatrix} \qquad (4.1)$$

be a density matrix on $\mathbb{C}^2$ with $\mu>1/2$, representing a mixture of spin-up and spin-down states, and for every $\mathbf{u}=(u_x,u_y)\in\mathbb{R}^2$ consider the state $\rho^{\mathbf{u}}=U(\mathbf{u})\rho^{\mathbf{0}}U(\mathbf{u})^*$, where

$$U(\mathbf{u}) := \exp(i(u_x\sigma_x+u_y\sigma_y)) = \begin{pmatrix} \cos|\mathbf{u}| & -e^{-i\varphi}\sin|\mathbf{u}| \\ e^{i\varphi}\sin|\mathbf{u}| & \cos|\mathbf{u}| \end{pmatrix},$$

with $\varphi=\mathrm{Arg}(-u_y+iu_x)$. We are interested in the asymptotic behavior as $n\to\infty$ of the family

$$\mathcal{F}_n := \left\{\rho_n^{\mathbf{u}} = \left(\rho^{\mathbf{u}/\sqrt{n}}\right)^{\otimes n}, \mathbf{u}\in I^2\right\}, \qquad (4.2)$$

where $I=[-a,a]$ is a fixed finite interval.

The main result is that $\mathcal{F}_n$ is asymptotically normal, meaning that it converges as $n\to\infty$ to a limit family $\mathcal{G}_n := \{\phi^{\mathbf{u}}, \mathbf{u}\in I^2\}$ of Gaussian states of a quantum oscillator with creation and annihilation operators satisfying $[a,a^*]=\mathbf{1}$. Let

$$\phi^{\mathbf{0}} := (1-p)\sum_{k=0}^{\infty}p^k|k\rangle\langle k|, \qquad (4.3)$$

be a thermal equilibrium state with $|k\rangle$ denoting the ks' energy level of the oscillator and $p=\frac{1-\mu}{\mu}<1$. For every $\mathbf{u}\in I^2$, define

$$\phi^{\mathbf{u}} := D(\sqrt{2\mu-1}\,\alpha_{\mathbf{u}})[\phi^{\mathbf{0}}]D(-\sqrt{2\mu-1}\,\alpha_{\mathbf{u}}), \qquad (4.4)$$

where $D(\mathbf{z}):=\exp(\mathbf{z}a^*-\mathbf{z}^*a)$ is the displacement operator, mapping the vacuum vector $|\mathbf{0}\rangle$ to the coherent vector $|\mathbf{z}\rangle$ and $\alpha_{\mathbf{u}}=(-u_y+iu_x)$.

The exact formulation of the convergence is given in Theorem 1.1. Thus, the state $\rho_n^{\mathbf{u}}$ of the n qubits which depends on the unknown parameter $\mathbf{u}$ can be manipulated by applying a quantum channel T_n such that its image converges to the Gaussian state $\phi^{\mathbf{u}}$, uniformly in $\mathbf{u}\in I^2$. Conversely, by using the channel S_n, the state $\phi^{\mathbf{u}}$ can be mapped to a joint state of n qubits which is converges to $\rho_n^{\mathbf{u}}$ uniformly in $\mathbf{u}\in I^2$. By Stinespring's theorem, we know that the channels are of the form

$$T(\rho) = \mathrm{Tr}_{\mathcal{K}}(V\rho V^*),$$

$$S(\phi) = \mathrm{Tr}_{\mathcal{K}'}(W\phi W^*),$$

where the partial traces are taken over some ancillary Hilbert spaces $\mathcal{K}, \mathcal{K}'$ and

$$V:(\mathbb{C}^2)^{\otimes n} \to L^2(\mathbb{R})\otimes\mathcal{K},$$

$$W:L^2(\mathbb{R}) \to (\mathbb{C}^2)^{\otimes n}\otimes\mathcal{K}',$$

are isometries ($V^*V=\mathbf{1}$ and $W^*W=\mathbf{1}$).

Our task is now to identify the isometries V_n and W_n implementing the channels T_n and respectively S_n satisfying Eq. (1.4). The first step toward identifying these V_n is to use group representations methods so as to partially (block) diagonalize all the $\rho_n^{\mathbf{u}}$ simultaneously.

A. Block decomposition

In this subsection, we show that the states $\rho_n^{\mathbf{u}}$ have a block-diagonal form given by the decomposition of the space $(\mathbb{C}^2)^{\otimes n}$ into irreducible representations of the relevant symmetry groups. The main point is that for large n the weights of the different blocks concentrate around the representation with total spin $j_n=n(\mu-1/2)$.

The space $(\mathbb{C}^2)^{\otimes n}$ carries a unitary representation π_n of the one-spin symmetry group $SU(2)$ with $\pi_n(u) = u^{\otimes n}$ for any $u \in SU(2)$, and a unitary representation of the symmetric group $S(n)$ given by the permutation of factors

$$\pi_n(\tau):v_1 \otimes \cdots \otimes v_n \mapsto v_{\tau^{-1}(1)} \otimes \cdots \otimes v_{\tau^{-1}(n)}, \quad \tau \in S(n).$$

As $[\pi_n(u), \pi_n(\tau)] = 0$ for all $u \in SU(2), \tau \in S(n)$, we have the decomposition

$$(\mathbb{C}^2)^{\otimes n} = \bigoplus_{j=0,1/2}^{n/2} \mathcal{H}_j \otimes \mathcal{H}_n^j, \tag{4.5}$$

where the direct sum runs over all positive (half-)integers j up to $n/2$, and for each fixed j, $\mathcal{H}_j \cong \mathbb{C}^{2j+1}$ is a irreducible representation of $SU(2)$ with total angular momentum $J^2 = j(j+1)$, and $\mathcal{H}_n^j \cong \mathbb{C}^{n_j}$ is the irreducible representation of the symmetric group $S(n)$ with $n_j = \binom{n}{n/2-j} - \binom{n}{n/2-j-1}$. In particular, the density matrix $\rho_n^{\mathbf{u}}$ is invariant under permutations, and can be decomposed as a mixture of "block" density matrices

$$\rho_n^{\mathbf{u}} = \bigoplus_{j=0,1/2}^{n/2} p_n(j) \rho_{j,n}^{\mathbf{u}} \otimes \frac{\mathbf{1}}{n_j}, \tag{4.6}$$

with probability distribution $p_n(j)$ given by [24]:

$$p_n(j) := \frac{n_j}{2\mu - 1}(1-\mu)^{\frac{n}{2}-j}\mu^{\frac{n}{2}+j+1}(1-p^{2j+1}), \tag{4.7}$$

where $p := \frac{1-\mu}{\mu}$. A key observation is that for large n and in the relevant range of j's, $p_n(j)$ is essentially a binomial distribution

$$B_{n,\mu}(k) := \binom{n}{k}\mu^k(1-\mu)^{n-k}, \quad k = 0, \ldots, n.$$

Indeed, we can rewrite $p_n(j)$ as

$$p_n(j) := B_{n,\mu}(n/2 + j) \times K(j,n,\mu), \tag{4.8}$$

where the factor $K(j,n,\mu)$ is given by

$$K(j,n,\mu) := (1-p^{2j+1})\frac{n + (2(j-j_n)+1)/(2\mu-1)}{n + (j-j_n+1)/\mu},$$

and $j_n := n(\mu - 1/2)$. As $B_{n,\mu}$ is the distribution of the sum of n independent Bernoulli variables with individual distribution $(1-\mu,\mu)$ over $\{0,1\}$, we can use the Central Limit Theorem to conclude that its mass concentrates around the average μn with a width of order $\sqrt{n}$, in other words of any $0 < \epsilon < 1/2$ we have

$$\lim_{n\to\infty} \sum_{p=-n^{1/2+\epsilon}}^{n^{1/2+\epsilon}} B_{n,\mu}(\mu n + p) = 1. \tag{4.9}$$

Let us denote by $\mathcal{J}_{n,\epsilon}$ the set of values j of the total angular momentum of n qubits which lie in the interval $[j_n - n^{1/2+\epsilon}, j_n + n^{1/2+\epsilon}]$. Then, for large n, the factor $K(j,n,\mu)$ is close to 1 uniformly over $j \in \mathcal{J}_{n,\epsilon}$, and from formulas (4.8) and (4.9) we conclude that $p_n(j)$ concentrates asymptotically in an interval of order $n^{1/2+\epsilon}$ around j_n:

$$\lim_{n\to\infty} p_n(\mathcal{J}_{n,\epsilon}) = 1. \tag{4.10}$$

This justifies the big ball picture used in the previous section.

B. Irreducible representations of $SU(2)$

Here, we remind the reader of some details about the irreducible representation π_j of $SU(2)$ on $\mathcal{H}_j$. Let $\sigma_x, \sigma_y, \sigma_z$ be the Pauli matrices and denote $J_{j,l} := \pi_j(\sigma_l)$ for $l = x, y, z$ the generators of rotations in the irreducible representation π_j, such that the corresponding unitaries are $U_j(\mathbf{u}) := \exp(i(u_x J_{j,x} + u_y J_{j,y}))$. There exists an orthonormal basis $\{|j,m\rangle, m = -j, \ldots, j\}$ of $\mathcal{H}_j$ such that

$$J_{j,z}|j,m\rangle = m|j,m\rangle.$$

Moreover, with $J_{j,\pm} := J_{j,x} \pm iJ_{j,y}$, we have

$$J_{j,+}|j,m\rangle = \sqrt{j-m}\sqrt{j+m+1}|j,m+1\rangle,$$

$$J_{j,-}|j,m\rangle = \sqrt{j-m+1}\sqrt{j+m}|j,m-1\rangle.$$

With these notations and $p = \frac{1-\mu}{\mu}$ as before, the state $\rho_{j,n}^0$ can be written as [29]

$$\rho_{j,n}^{\mathbf{0}} = c_j(p) \sum_{m=-j}^{j} p^{j-m}|j,m\rangle\langle j,m|,$$

where the normalizing factor is $c_j(p) = (1-p)/(1-p^{2j+1})$. The rotated block states can be obtained by applying the unitary transformation

$$\rho_{j,n}^{\mathbf{u}} = U_j(\mathbf{u}/\sqrt{n})\rho_{j,n}^{\mathbf{0}}U_j(\mathbf{u}/\sqrt{n})^*,$$

with $U_j(\mathbf{u})$ as above. Finally, we define the vectors

$$|j,\mathbf{w}\rangle := U_j(\mathbf{w})|j,j\rangle \tag{4.11}$$

which will be used in later computations, and notice that their coordinates, with respect to the $|j,m\rangle$ basis, are given by [29]:

$$\langle j,m|j,\mathbf{w}\rangle = \sqrt{\binom{2j}{j+m}}\zeta^{j-m}(1-|\zeta|^2)^{\frac{j+m}{2}}. \tag{4.12}$$

where $\zeta = e^{i\varphi_w}\sin|\mathbf{w}|$ with $\varphi_w = \mathrm{Arg}(-w_y + iw_x)$.

V. CONSTRUCTION OF THE CHANNELS T_n

For each irreducible representation space $\mathcal{H}_j$, we define the isometry $V_j: \mathcal{H}_j \to L^2(\mathbb{R})$ by

$$V_j:|j,m\rangle \mapsto |j-m\rangle, \tag{5.1}$$

where $\{|n\rangle, n \geq 0\}$ represents the energy eigenbasis of the quantum oscillator with eigenfunctions $\psi_n(x) = H_n(x)e^{-x^2/2}/\sqrt{\sqrt{\pi}2^n n!} \in L^2(\mathbb{R})$. Using the decomposition (4.5), we put together the different blocks we construct for each $n \in \mathbb{N}$ the "global" isometry

$$V_n := \bigoplus_{j=0,1/2}^{n/2} V_j \otimes \mathbf{1}: \bigoplus_{j=0,1/2}^{n/2} \mathcal{H}_j \otimes \mathbb{C}^{n_j} \to L^2(\mathbb{R}) \otimes \mathcal{K}_n,$$

where $\mathcal{K}_n := \bigoplus_{j=0,1/2}^{n/2}\mathbb{C}^{n_j}$. By tracing over $\mathcal{K}_n$, we obtain the channel $T_n(\rho) := \mathrm{Tr}_{\mathcal{K}_n}(V_n\rho V_n^*)$ mapping a joint state of n spins

into a state of the quantum oscillator. This channel satisfies the convergence condition (1.4) as shown by the estimate

$$\|T_n(\rho_n^{\mathbf{u}}) - \phi^{\mathbf{u}}\|_1 = \left\| \sum_{j=0,1/2}^{n/2} p_n(j) V_j \rho_{n,j}^{\mathbf{u}} V_j^* - \phi^{\mathbf{u}} \right\|_1$$

$$\leq \sum_{j=0,1/2}^{n/2} p_n(j) \|V_j \rho_{n,j}^{\mathbf{u}} V_j^* - \phi^{\mathbf{u}}\|_1$$

$$\leq 2 \sum_{j \notin \mathcal{J}_{n,\epsilon}} p_n(j) + \sup_{\mathbf{u} \in I^2} \max_{j \in \mathcal{J}_{n,\epsilon}} \|V_j \rho_{j,n}^{\mathbf{u}} V_j^* - \phi^{\mathbf{u}}\|_1,$$

where the first term on the right side converges to 0 by (4.10), and for the second one we apply the following Proposition (5.1) which is the major technical contribution of this paper.

Proposition (5.1). *The following uniform convergence holds*

$$\lim_{n \to \infty} \sup_{\mathbf{u} \in I^2} \max_{j \in \mathcal{J}_{n,\epsilon}} \|V_j \rho_{j,n}^{\mathbf{u}} V_j^* - \phi^{\mathbf{u}}\|_1 = 0,$$

where $\mathcal{J}_{n,\epsilon}$ is the set defined above in Eq. (4.10).

The proof of the Proposition requires a few ingredients which, in our opinion, are important on their own and we formulate them apart and refer to relevant papers for the proofs.

Theorem(5.2). [36] *Let $a, b \in M(\mathbb{C}^d)$, satisfying $\mathrm{Tr}(a) = \mathrm{Tr}(b) = 0$ and define*

$$L(a,b) = \exp(ia)\exp(ib) - \exp(ia+ib)\exp\left(\frac{1}{2}[a,b]\right).$$

On $(\mathbb{C}^2)^{\otimes n}$, we define the fluctuation operator

$$F_n(a) = \frac{1}{\sqrt{n}} \sum a_i,$$

where $a_i = \mathbf{1} \otimes \cdots \otimes a \otimes \cdots \otimes \mathbf{1}$ with a acting on the i's position of the tensor product. Notice that $\exp(iF_n(a)) = \exp(ia/\sqrt{n})^{\otimes n}$ and $\sqrt{n}[F_n(a), F_n(b)] = F_n([a,b])$. Then,

$$\lim_{n \to \infty} \|L(F_n(a), F_n(b))\| = 0.$$

The convergence is uniform over $\|a\|$, $\|b\| < C$ for some constant C.

This Theorem is a key ingredient of the quantum Central Limit Theorem [36], and it is not surprising that it plays an important role in our quantum local asymptotic normality result which is an extension of the latter. We apply the Theorem to two unitaries of the form $U(\mathbf{u}/\sqrt{n})^{\otimes n} = \exp(i(u_x \sigma_x + u_y \sigma_y))/\sqrt{n})^{\otimes n}$. We thus get information on the effect of the $U_j(\mathbf{u}/\sqrt{n})$ on the highest weight vectors $|j,j\rangle$ of an irreducible representation.

Corollary (5.3). *For any unitary U and state τ, let*

$\mathrm{Ad}[U](\tau) := U\tau U^*$ *and consider the rotated states*

$$\tau(\mathbf{u}, \mathbf{v}, j, n) := \mathrm{Ad}\left[U_j\left(\frac{\mathbf{u}}{\sqrt{n}}\right) U_j\left(\frac{\mathbf{v}}{\sqrt{n}}\right) \right] (|jj\rangle\langle jj|),$$

$$\tau(\mathbf{u}+\mathbf{v}, j, n) := \mathrm{Ad}\left[U_j\left(\frac{\mathbf{u}+\mathbf{v}}{\sqrt{n}}\right) \right] (|jj\rangle\langle jj|).$$

Then, the following uniform convergence holds

$$\lim_{n \to \infty} \sup_{\mathbf{u}, \mathbf{v} \in I^2} \sup_{j \in \mathcal{J}_{n,\epsilon}} \|\tau(\mathbf{u}, \mathbf{v}, j, n) - \tau(\mathbf{u}+\mathbf{v}, j, n)\|_1 = 0.$$

Proof. By applying Theorem (5.2) to $U(\mathbf{u}/\sqrt{n})^{\otimes n}$ and $U(\mathbf{v}/\sqrt{n})^{\otimes n}$, the first term of $L(F_n(a))$ is

$$U_1 = U\left(\frac{\mathbf{u}}{\sqrt{n}}\right)^{\otimes n} U\left(\frac{\mathbf{v}}{\sqrt{n}}\right)^{\otimes n},$$

and the second is

$$U_2 = U\left(\frac{\mathbf{u}+\mathbf{v}}{\sqrt{n}}\right)^{\otimes n} \exp\left(\frac{F_n([\mathbf{u}, \mathbf{v}])}{2\sqrt{n}}\right)$$

with

$$[\mathbf{u}, \mathbf{v}] := [u_x\sigma_x + u_y\sigma_y, v_x\sigma_x + v_y\sigma_y] = 2(u_x v_y - u_y v_x)\sigma_z.$$

The norm one distance between these two operators is going to 0 as n is going to infinity, uniformly on $(\mathbf{u}, \mathbf{v}) \in I^2$. We may apply these operators on any pure state of $(\mathbb{C}^2)^{\otimes n}$, in particular on $|j,j\rangle\langle j,j|$ for any $j \in \mathcal{J}_{n,\epsilon}$ after block diagonalization and preserve the uniform limit

$$\|\mathrm{Ad}[U_1](|jj\rangle\langle jj|) - \mathrm{Ad}[U_2](|jj\rangle\langle jj|)\| \underset{n \to \infty}{\to} 0. \qquad (5.2)$$

Now, the action of $\exp\left(\frac{F_n([\mathbf{u}, \mathbf{v}])}{2\sqrt{n}}\right)$ on $|j,j\rangle\langle j,j|$ is simply identity because $|jj\rangle$ is an eigenvector of $J_{j,z}$. Thus,

$$\mathrm{Ad}[U_2](|jj\rangle\langle jj|) = \tau(\mathbf{u}+\mathbf{v}, j, n).$$

Together with Eq. (5.2), this ends the proof.

The following Lemma is a slight strengthening of a theorem by Hayashi and Matsumoto [29].

Lemma (5.4). *The uniform convergence holds*

$$\lim_{n \to \infty} \sup_{\mathbf{u} \in I^2} \sup_{j \in \mathcal{J}_{n,\epsilon}} \left\| V_j U_j\left(\frac{\mathbf{u}}{\sqrt{n}}\right) |jj\rangle - |\sqrt{2\mu - 1}\,\alpha_{\mathbf{u}}\rangle \right\| = 0,$$

where $|\mathbf{z}\rangle$ denotes a coherent state of the oscillator, and $\alpha_{\mathbf{u}} := (-u_y + iu_x)$. Moreover, for any sequence $j_n \to \infty$, we have

$$\lim_{n \to \infty} \|V_{j_n} \rho_{j_n}^{\mathbf{0}} V_{j_n}^* - \phi^{\mathbf{0}}\|_1 = 0. \qquad (5.3)$$

The convergence holds uniformly over all sequences j_n such that $j_n/n > c$ for some fixed constant $c > 0$, so, in particular for $j_n \in \mathcal{J}_{n,\epsilon}$.

Proof. We first prove the easier relation (5.3). As both density matrices are diagonal, we get

$$\|V_{j_n}\rho_{j_n}^{\mathbf{0}}V_{j_n}^* - \phi^{\mathbf{0}}\|_1 = \frac{(1-p)p^{2j_n+1}}{1-p^{2j_n+1}}\sum_{k=0}^{2j_n}p^k - (1-p)\sum_{k=2j_n+1}^{\infty}p^k$$

$$\leq \frac{p^{2j_n+1}}{1-p^{2j_n+1}} + p^{2j_n+1} \to 0,$$

as $n \to \infty$.

As for the first relation, let us denote $|\mathbf{u},j,n\rangle := V_j U_j\left(\frac{\mathbf{u}}{\sqrt{n}}\right)|j,j\rangle$, then by Eqs. (4.12) and (5.1), we have

$$\langle k|\mathbf{u},j,n\rangle = \sqrt{\binom{2j}{k}}(\sin(|\mathbf{u}|/\sqrt{n})e^{i\phi})^k(\cos(|\mathbf{u}|/\sqrt{n}))^{2j-k}.$$

Now, the following asymptotical relations hold uniformly over $j \in \mathcal{J}_{n,\epsilon}$:

$$\sin\left(\frac{|\mathbf{u}|}{\sqrt{n}}\right)^k = \left(\frac{|\mathbf{u}|}{\sqrt{n}}\right)^k(1 + O(|\mathbf{u}|^2 n^{-1})),$$

$$\cos\left(\frac{|\mathbf{u}|}{\sqrt{n}}\right)^{2j-k} = \exp\left(-\frac{(2\mu-1)|\mathbf{u}|^2}{2}\right)(1 + O(|\mathbf{u}|^2 n^{-\epsilon})),$$

$$\binom{2j}{k} = \frac{((2\mu-1)n)^k}{k!}(1 + O(n^{-\epsilon})),$$

and thus the coefficients converge uniformly to those of the corresponding coherent states as $n \to \infty$

$$\langle k|\mathbf{u},j,n\rangle \to \exp\left(-\frac{(2\mu-1)|\mathbf{u}|^2}{2}\right)\frac{(e^{i\phi}|\mathbf{u}|\sqrt{2\mu-1})^k}{\sqrt{k!}}.$$

Proof of Proposition (5.1). The main idea is to notice that ϕ^0 is a thermal equilibrium state of the oscillator and can be generated as a mixture of coherent states with centered Gaussian distribution over the displacements:

$$\phi^{\mathbf{0}} = \frac{1}{\sqrt{2\pi s^2}}\int e^{-|\mathbf{z}|^2/2s^2}|\mathbf{z}\rangle\langle\mathbf{z}|d^2\mathbf{z}. \quad (5.4)$$

The easiest way to see this is to think of the oscillator states in terms of their Wigner functions. Indeed, the Wigner function of a coherent state is

$$W_{\mathbf{z}}(q,p) = \exp(-(q-\sqrt{2}\mathrm{Re}\,\mathbf{z})^2 - (p-\sqrt{2}\mathrm{Im}\,\mathbf{z})^2),$$

and thus the state given by Eq. (5.4) has a Wigner function which is the convolution of two centered Gaussians, again a centered Gaussian with a variance equal to the sum of their

variances $2s^2 + 1/2$, which is equal to the variance of ϕ^0 for $s^2 := p/(2(1-p))$. Similarly,

$$\phi^{\mathbf{u}} = \frac{1}{2\pi s^2}\int e^{-|\mathbf{z}-\sqrt{2\mu-1}\alpha_{\mathbf{u}}|^2/2s^2}(|\mathbf{z}\rangle\langle\mathbf{z}|)d^2\mathbf{z}. \quad (5.5)$$

Let us first remark that

$$\|V_{j_n}\rho_{j_n}^{\mathbf{u}}V_{j_n}^* - \phi^{\mathbf{u}}\|_1 \leq \|\rho_{j_n}^{\mathbf{u}} - V_{j_n}^*\phi^{\mathbf{u}}V_{j_n}\|_1 + \|\phi^{\mathbf{u}} - P_{j_n}\phi^{\mathbf{u}}P_{j_n}\|_1,$$

where $P_{j_n} = V_{j_n}V_{j_n}^*$ is the projection onto the image of V_{j_n}, and

$$\lim_{n\to\infty}\sup_{j_n\in\mathcal{J}_{n,\epsilon}}\sup_{\mathbf{u}\in I^2}\|\phi^{\mathbf{u}} - P_{j_n}\phi^{\mathbf{u}}P_{j_n}\|_1 = 0,$$

because $j_n \to \infty$ uniformly and P_{j_n} converges to the identity in strong operator topology (a tightness property). Thus, it is enough to show that

$$\lim_{n\to\infty}\sup_{j_n\in\mathcal{J}_{n,\epsilon}}\sup_{\mathbf{u}\in I^2}\|\rho_{j_n}^{\mathbf{u}} - V_{j_n}^*\phi^{\mathbf{u}}V_{j_n}\|_1 = 0.$$

Now

$$\|\rho_{j_n}^{\mathbf{u}} - V_{j_n}^*\phi^{\mathbf{u}}V_{j_n}\|_1$$

$$= \left\|\mathrm{Ad}\left[U_{j_n}\left(\frac{\mathbf{u}}{\sqrt{n}}\right)\right](\rho_{j_n}^{\mathbf{0}}) - V_{j_n}^*\phi^{\mathbf{u}}V_{j_n}\right\|_1$$

$$\leq \|\rho_{j_n}^{\mathbf{0}} - V_{j_n}^*\phi^{\mathbf{0}}V_{j_n}\|_1$$

$$+ \left\|\mathrm{Ad}\left[U_{j_n}\left(\frac{\mathbf{u}}{\sqrt{n}}\right)\right](V_{j_n}^*\phi^{\mathbf{0}}V_{j_n}) - V_{j_n}^*\phi^{\mathbf{u}}V_{j_n}\right\|_1.$$

The first term on the right side of the inequality converges to zero by Lemma (5.4), uniformly for any sequence (j_n) such that $j_n \in \mathcal{J}_{n,\epsilon}$ and does not depend on $\mathbf{u}$. Using Eqs. (5.4) and (5.5), we bound the second term by

$$\frac{1}{s\sqrt{2\pi}}\int e^{-|\mathbf{z}|^2/2s^2}\|\Delta(\mathbf{u},\mathbf{z},j_n)\|_1 d^2\mathbf{z},$$

where the operator $\Delta(\mathbf{u},\mathbf{z},j_n)$ is given by

$$\Delta(\mathbf{u},\mathbf{z},j_n) := \mathrm{Ad}\left[U_{j_n}\left(\frac{\mathbf{u}}{\sqrt{n}}\right)\right](V_{j_n}^*|\mathbf{z}\rangle\langle\mathbf{z}|V_{j_n})$$

$$- V_{j_n}^*|\mathbf{z}+\sqrt{2\mu-1}\alpha_{\mathbf{u}}\rangle\langle\mathbf{z}+\sqrt{2\mu-1}\alpha_{\mathbf{u}}|V_{j_n}.$$

We analyze the expression under the integral. Let $\tilde{\mathbf{z}} \in \mathbb{R}^2$ be such that $\alpha_{\tilde{\mathbf{z}}} = \mathbf{z}/\sqrt{2\mu-1}$, then

$$\left\|\mathrm{Ad}\left[U_{j_n}\left(\frac{\mathbf{u}}{\sqrt{n}}\right)\right](V_{j_n}^*|\mathbf{z}\rangle\langle\mathbf{z}|V_{j_n}) - V_{j_n}^*|\mathbf{z}+\sqrt{2\mu-1}\alpha_{\mathbf{u}}\rangle\langle\mathbf{z}+\sqrt{2\mu-1}\alpha_{\mathbf{u}}|V_{j_n}\right\|_1$$

$$\leq \left\|\mathrm{Ad}\left[U_{j_n}\left(\frac{\mathbf{u}}{\sqrt{n}}\right)U_{j_n}\left(\frac{\tilde{\mathbf{z}}}{\sqrt{n}}\right)\right](|j_n j_n\rangle\langle j_n j_n|) - \mathrm{Ad}\left[U_{j_n}\left(\frac{\mathbf{u}+\tilde{\mathbf{z}}}{\sqrt{n}}\right)\right](|j_n j_n\rangle\langle j_n j_n|)\right\|_1$$

$$+ \left\| V_{j_n} \mathrm{Ad}\left[U_{j_n}\left(\frac{\tilde{\mathbf{z}}}{\sqrt{n}} \right) \right] (|j_n j_n\rangle\langle j_n j_n|) V_{j_n}^* - |\mathbf{z}\rangle\langle \mathbf{z}| \right\|_1$$

$$+ \left\| V_{j_n} \mathrm{Ad}\left[U_{j_n}\left(\frac{\mathbf{u}+\tilde{\mathbf{z}}}{\sqrt{n}} \right) \right] (|j_n j_n\rangle\langle j_n j_n|) V_{j_n}^* - |\mathbf{z}+\sqrt{2\mu-1}\,\alpha_{\mathbf{u}}\rangle\langle \mathbf{z}+\sqrt{2\mu-1}\,\alpha_{\mathbf{u}}| \right\|_1 .$$

By Corollary (5.3), the first term on the right side converges to zero uniformly in $(\mathbf{u}, j_n) \in I^2 \times \mathcal{J}_{n,\epsilon}$. By Lemma (5.4), we have that the last two terms converge to zero uniformly in $(\mathbf{u}, j_n) \in I^2 \times \mathcal{J}_{n,\epsilon}$. Thus, if we denote

$$F_n(\mathbf{z}) := \sup_{j_n \in \mathcal{J}_{n,\epsilon}} \sup_{\mathbf{u} \in I^2} \|\Delta(\mathbf{u}, \mathbf{z}, j_n)\|_1 ,$$

then $0 \leq F_n(\mathbf{z}) \leq 2$, $\lim_{n\to\infty} F_n(\mathbf{z}) = 0$ for all $\mathbf{z} \in \mathbb{R}^2$, and by the Lebesgue dominated convergence theorem, we get

$$\lim_{n\to\infty} \frac{1}{s\sqrt{2\pi}} \int e^{-|z|^2/2s^2} F_n(\mathbf{z}) d^2\mathbf{z} = 0 .$$

This implies the statement of the Proposition (5.1).

VI. CONSTRUCTION OF THE INVERSE CHANNEL S_n

To complete our proof of asymptotic equivalence as defined by Eq. (1.4), we must now exhibit the inverse channel S_n which maps the displaced thermal states $\phi^{\mathbf{u}}$ of the oscillator into approximations of the rotated spin states. As the latter are block diagonal with weights $p_n(j)$ as defined in Eq. (4.7), it is natural to look for S_n of the form

$$S_n(\phi) = \bigoplus_{j=0,1/2}^{n/2} p_n(j) S_n^j(\phi) \otimes \frac{\mathbf{1}}{n_j} ,$$

where S_n^j are channels with outputs in $\mathcal{H}_j$. Moreover, because $V_j : \mathcal{H}_j \to L^2(\mathbb{R})$ is an isometry, we can choose S_n^j such that

$$S_n^j(V_j \rho V_j^*) = \rho , \qquad (6.1)$$

for all density matrices ρ on $\mathcal{H}_j$. This property does not fix the channel completely, but it is sufficient for our purposes. Basically, what we want is an inverse of the embedding $V_j \cdot V_j^*$ used for the direct channel, and one way to get this is as follows: First, block diagonalize ϕ to get $P_j \phi P_j + P_j^\perp \phi P_j^\perp$, where P_j is the projection onto the image of V_j, i.e., $P_j = V_j V_j^*$, and note that this is a trace preserving completely positive map. This block-diagonal state can be now seen as a state on the direct sum algebra $\mathcal{B}(P_j L^2(\mathbb{R})) \oplus \mathcal{B}(P_j^\perp L^2(\mathbb{R}))$ and can be mapped to a state on $\mathcal{B}(\mathcal{H}_j)$ by the channel $V_j^* \cdot V_j \oplus S_j^\perp$ with $S_j^\perp$ arbitrary on the "upper block". The resulting composition is the channel S_n^j satisfying property (6.1).

Theorem (6.1). *The following holds*

$$\limsup_{n\to\infty} \sup_{u\in I^2} \|S_n(\phi^{\mathbf{u}}) - \rho_n^{\mathbf{u}}\|_1 = 0 .$$

Proof. As both $\rho_n^{\mathbf{u}}$ and $S_n(\phi^{\mathbf{u}})$ are block diagonal, we may decompose their distance as

$$\|S_n(\phi^{\mathbf{u}}) - \rho_n^{\mathbf{u}}\|_1 = \sum_{j=0,1/2}^{n/2} p_n(j) \|S_n^j(\phi^{\mathbf{u}}) - \rho_{j,n}^{\mathbf{u}}\|_1 \leq \sum_{j \notin \mathcal{J}_{n,\epsilon}} 2 p_n(j) + \sum_{j \in \mathcal{J}_{n,\epsilon}} p_n(j) \|S_n^j(\phi^{\mathbf{u}}) - S_n^j(V_j \rho_{j,n}^{\mathbf{u}} V_j^*)\|_1$$

$$+ \sum_{j \in \mathcal{J}_{n,\epsilon}} p_n(j) \|S_n^j(V_j \rho_{j,n}^{\mathbf{u}} V_j^*) - \rho_{j,n}^{\mathbf{u}}\|_1 \leq 2 \sum_{j \notin \mathcal{J}_{n,\epsilon}} p_n(j) + \sum_{j \in \mathcal{J}_{n,\epsilon}} p_n(j) \|\phi^{\mathbf{u}} - V_j \rho_{j,n}^{\mathbf{u}} V_j^*\|_1 ,$$

where we have used at the last line that S_n^j is a contraction and property (6.1) of S_n^j. Now, the first sum is going to 0 by Eq. (4.10) and the second sum is also uniformly going to 0 by use of Proposition (5.1).

VII. APPLICATIONS

A. Optimal Bayes measurement is also asymptotically local minimax

In this subsection, we will introduce some ideas from the point-wise approach to state estimation. We show that the measurement, which is known to be optimal for a uniform prior in the Bayesian setup, is also asymptotically optimal in the point-wise sense.

Using the jargon of mathematical statistics, we will call the *quantum statistical experiment (model)* [40] a family $\{\rho^\theta \in M(\mathbb{C}^d) : \theta \in \Theta\}$ of density matrices indexed by a parameter belonging to a set Θ. The main examples of quantum statistical experiments considered so far are that of n identical qubits

$$\mathcal{F} := \{\rho^{\otimes n} : \rho \in M(\mathbb{C}^2)\},$$

the local model

$$\mathcal{F}_n^I := \{\rho_n^{\mathbf{u}} = (\rho^{\mathbf{u}/\sqrt{n}})^{\otimes n}, \mathbf{u} \in I^2\},$$

and its "limit" model

$$\mathcal{G}^I := \{\phi^{\mathbf{u}}, \mathbf{u} \in I^2\},$$

where $I = [-a, a]$, and $\rho_n^{\mathbf{u}}$ and $\phi^{\mathbf{u}}$ are defined by Eqs. (1.1) and (1.2). More generally, we can replace the square I^2 by an arbitrary region K in the parameter space and obtain:

$$\mathcal{G}^K := \{\phi^{\mathbf{u}}, \mathbf{u} \in K \subset \mathbb{R}^2\}.$$

We shall also make use of

$$\mathcal{G} := \{\phi^{\mathbf{u}}, \mathbf{u} \in \mathbb{R}^2\}.$$

A natural choice of distance between density matrices is related to the fidelity square

$$F(\rho, \sigma)^2 = [\mathrm{Tr}((\sqrt{\rho}\sigma\sqrt{\rho})^{1/2})]^2,$$

which is locally quadratic in first-order approximation, i.e.,

$$1 - F(\rho_n^{\mathbf{u}}, \rho_n^{\mathbf{v}})^2 \approx \frac{1}{n}\|\mathbf{u} - \mathbf{v}\|^2.$$

As we expect that reasonable estimators are in a local neighborhood of the true state, we will replace the fidelity square by the local distance

$$d(\mathbf{u}, \hat{\mathbf{u}}) = \|\hat{\mathbf{u}} - \mathbf{u}\|^2,$$

and define the risk of a measurement-estimator pair as $R_M(\mathbf{u}, \hat{\mathbf{u}}) = \langle d(\mathbf{u}, \hat{\mathbf{u}}) \rangle$, keeping in mind the factor $1/n$ relating the risks expressed in local and global parameters.

Similar to the Bayesian approach, we are interested in estimators which have small risk *everywhere* in the parameter space, and we define a worst case figure of merit called the minimax risk.

Definition (7.1). *The minimax risk of a quantum statisti-*

cal *experiment $\mathcal{E}$ over the parameter space Θ for loss function $d(\theta, \hat{\theta})$, is defined as*

$$C(\mathcal{E}) = \inf_{M, \hat{\theta}} \sup_{\theta \in \Theta} R_M(\theta, \hat{\theta}), \qquad (7.1)$$

where the infimum is taken over all measurement-estimator pairs $(M, \hat{\theta})$, and $R_M(\theta, \hat{\theta}) = \langle d(\theta, \hat{\theta}) \rangle$.

The minimax risk tells us how difficult the model is and thus we expect that if two models are "close" to each other then their minimax risks are almost equal. The "statistical distance" between quantum experiments is defined in a natural way with direct physical interpretation, and such a problem has been already addressed in [43] for the case of a quantum statistical experiment consisting of a finite family of pure states.

Definition (7.2). *Let $\mathcal{E} = \{\rho^\theta \in M(\mathbb{C}^d) : \theta \in \Theta\}$ and $\mathcal{F} = \{\tau^\theta \in M(\mathbb{C}^p) : \theta \in \Theta\}$ be two quantum statistical experiments (models) with the same parameter space Θ. We define the deficiencies*

$$\delta(\mathcal{E}, \mathcal{F}) = \inf_T \sup_{\theta \in \Theta} \|T(\rho^\theta) - \tau^\theta\|_1,$$

$$\delta(\mathcal{F}, \mathcal{E}) = \inf_S \sup_{\theta \in \Theta} \|\rho^\theta - S(\tau^\theta)\|_1,$$

where the infimum is taken over all trace preserving channels $T : M(\mathbb{C}^d) \to M(\mathbb{C}^p)$ and $S : M(\mathbb{C}^p) \to M(\mathbb{C}^d)$.

With this terminology, our main result states that for any bounded interval I:

$$\lim_{n \to \infty} \max(\delta(\mathcal{F}_n^I, \mathcal{G}^I), \delta(\mathcal{G}^I, \mathcal{F}_n^I)) = 0. \qquad (7.2)$$

As suggested above, the deficiency has a direct statistical interpretation: If we want to estimate θ in both statistical experiments $\mathcal{E}$ and $\mathcal{F}$ and we choose a bounded loss function $d(\theta, \hat{\theta}) \leqslant K$, then for any measurement and estimator $\hat{\theta}$ for $\mathcal{F}$ with risk $R_M(\theta, \hat{\theta}) = \langle d(\theta, \hat{\theta}) \rangle$ we can find a measurement N on $\mathcal{E}$ whose risk is at most $R_M(\theta, \hat{\theta}) + K\delta(\mathcal{E}, \mathcal{F})$. Indeed, if we choose T such that the infimum in the definition of $\delta(\mathcal{E}, \mathcal{F})$ is achieved, we can map the state ρ^θ through the channel T and then perform M to obtain an estimator $\tilde{\theta}$ for which

$$R_N(\theta, \tilde{\theta}) = \langle d(\theta, \tilde{\theta}) \rangle = \int_\Theta d(\theta, \tilde{\theta}) \mathrm{Tr}(T(\rho^\theta)M(d\tilde{\theta})) \leqslant \int_\Theta d(\theta, \tilde{\theta}) \mathrm{Tr}(\tau^\theta M(d\tilde{\theta})) + \|d\|_\infty \|T(\rho^\theta) - \tau^\theta\|_1 \leqslant R_M(\theta, \hat{\theta}) + K\delta(\mathcal{E}, \mathcal{F}).$$

This means that the difficulty of estimating the parameter θ in the two models is comparable within a factor $\delta(\mathcal{E}, \mathcal{F})$. With the above definition of the minimax risk and using the convergence (7.2), we obtain the following lemma.

Lemma (7.3). *Let $I = [-a, a]$ with $0 < a < \infty$, then*

$$\lim_{n \to \infty} C(\mathcal{F}_n^I) = C(\mathcal{G}^I).$$

The minimax risk for the local family $\mathcal{F}_n^I$ is a figure of merit for the "local difficulty" of the global model $\mathcal{F}_n$. It converges asymptotically to the minimax risk of a family of

thermal states. However, this quantity depends on the arbitrary parameter $I=[-a,a]$ which we would like to remove as our last step in defining the *local asymptotic minimax risk*:

$$C_{\text{l.a.m.}}(\mathcal{F}_n : n \in \mathbb{N}) := \lim_{a \to \infty} \lim_{n \to \infty} C(\mathcal{F}_n^I) = \lim_{a \to \infty} C(\mathcal{G}^I).$$

This quantity depends in principle on the state which is at the center of the local neighborhood. However, by invariance under rotations, the risk is constant for a given pair of eigenvalues of the density matrix. Now, as one might expect the minimax risks for the restricted families of thermal states will converge to that of the experiment with no restrictions on the parameters. The proof of this fact is however nontrivial.

Lemma (7.4). *Let $I=[-a,a]$, then we have*

$$\lim_{a \to \infty} C(\mathcal{G}^I) = C(\mathcal{G}).$$

Moreover the heterodyne measurement saturates $C(\mathcal{G})$, and thus $C(\mathcal{G})$ is equal to the Holevo bound.

Proof. The inequality in one direction is easy. For any estimator, $\sup_{\mathbf{u} \in I^2} R_M(\mathbf{u}, \hat{\mathbf{u}}) \leq \sup_{\mathbf{u} \in \mathbb{R}^2} R_M(\mathbf{u}, \hat{\mathbf{u}})$, so that $C(\mathcal{G}^I) \leq C(\mathcal{G})$ and the same holds for the limit. By the same reasoning, for any $K_1 \subset K_2 \subset \mathbb{R}^2$, we have $C(\mathcal{G}^{K_1}) \leq C(\mathcal{G}^{K_2})$.

When calculating minimax bounds, we are interested in the worst risk of estimators within some parameter region K, and this worst risk is obviously higher than the Bayes risk with respect to the probability distribution with constant density on K. We shall work on $B(0, c+b)$ the ball of center 0 and radius $(c+b)$, with $b > c$, and denote our measurement M with density $m(\hat{\mathbf{u}})d\hat{\mathbf{u}}$. In general, M need not have a density, but this will ease notations. Then,

$$\sup_{\mathbf{u} \in B(0,c+b)} R_M(\mathbf{u}, \hat{\mathbf{u}}) \geq$$

$$\int_{B(0,c+b) \times \mathbb{R}^2} \frac{d\mathbf{u}d\hat{\mathbf{u}}}{\pi(c+b)^2} \|\mathbf{u} - \hat{\mathbf{u}}\|^2 \text{Tr}(\phi^{\mathbf{u}} m(\hat{\mathbf{u}})). \quad (7.3)$$

We fix the following notations

$$f(\mathcal{D}) = \int_{\mathcal{D}} d\mathbf{u}d\mathbf{v} \|\mathbf{x} - \mathbf{y}\|^2 \text{Tr}(\phi^{\mathbf{u}} m(\mathbf{v})),$$

$$g(\mathcal{D}) = \int_{\mathcal{D}} d\mathbf{u}d\mathbf{v} \text{Tr}(\phi^{\mathbf{u}} m(\mathbf{v})),$$

and define the domains

$$\mathcal{D}_1 = \{(\mathbf{u}, \hat{\mathbf{u}}) | \mathbf{u} \in B(0, c+b), \hat{\mathbf{u}} \in \mathbb{R}^2\},$$

$$\mathcal{D}_2 = \{(\mathbf{u} + \mathbf{k}, \mathbf{k}) | \mathbf{u} \in B(0,c), \mathbf{k} \in B(0,b)\},$$

$$\mathcal{D}_3 = \{(\mathbf{u}, \mathbf{u} + \mathbf{h}) | \mathbf{u} \in B(0, b-c), \mathbf{h} \in B(0,c)\},$$

$$\mathcal{D}_4 = \{(\mathbf{u}, \mathbf{u} + \mathbf{h}) | \mathbf{u} \in B(0, b-c), \mathbf{h} \in \mathbb{R}^2 \setminus B(0,c)\}.$$

Notice the following relations:

$$\mathcal{D}_3 \subset \mathcal{D}_2 \subset \mathcal{D}_1, \quad \mathcal{D}_4 \subset \mathcal{D}_1 \setminus \mathcal{D}_2. \quad (7.4)$$

Then Eq. (7.3) can be rewritten as

$$\sup_{\mathbf{u} \in B(0,c+b)} R_M(\mathbf{u}, \hat{\mathbf{u}}) \geq \frac{1}{\pi(b+c)^2} f(\mathcal{D}_1).$$

The following inequalities follow directly from the definitions:

$$f(\mathcal{D}_2) \leq c^2 g(\mathcal{D}_2), \qquad f(\mathcal{D}_3) \leq c^2 g(\mathcal{D}_3),$$
$$f(\mathcal{D}_4) \geq c^2 g(\mathcal{D}_4), \qquad g(\mathcal{D}_4) + g(\mathcal{D}_3) = \pi(b-c)^2.$$

Using these and Eq. (7.4), we may write:

$$\frac{1}{\pi(c+b)^2} f(\mathcal{D}_1) \geq \frac{1}{\pi(c+b)^2}(f(\mathcal{D}_2) + f(\mathcal{D}_4)) \geq \frac{1}{\pi(c+b)^2}(f(\mathcal{D}_2) + c^2 g(\mathcal{D}_4)) = \frac{(b-c)^2}{(b+c)^2}\left(\frac{f(\mathcal{D}_2)}{g(\mathcal{D}_2)}\frac{g(\mathcal{D}_2)}{\pi(b-c)^2} + c^2 - c^2 \frac{g(\mathcal{D}_3)}{\pi(b-c)^2}\right)$$

$$\geq \frac{(b-c)^2}{(b+c)^2}\left(c^2 + \frac{g(\mathcal{D}_3)}{\pi(b-c)^2}\left(\frac{f(\mathcal{D}_2)}{g(\mathcal{D}_2)} - c^2\right)\right) \geq \frac{(b-c)^2}{(b+c)^2}\frac{f(\mathcal{D}_2)}{g(\mathcal{D}_2)}. \quad (7.5)$$

We analyze now the expression $f(\mathcal{D}_2)/g(\mathcal{D}_2)$. By using the definition (1.2) of the displaced thermal states $\phi^{\mathbf{u}}$ we get that $\text{Tr}[\phi^{\mathbf{u}+\mathbf{k}} m(\mathbf{l})] = \text{Tr}[\phi^{\mathbf{k}} m_{\mathbf{u}}(\mathbf{l})]$, where

$$m_{\mathbf{u}}(\mathbf{l}) := D(-\sqrt{2\mu - 1}\alpha_{\mathbf{u}}) m(\mathbf{l}) D(\sqrt{2\mu - 1}\alpha_{\mathbf{u}}).$$

Then,

$$g(\mathcal{D}_2) = \int_{B(0,c) \times B(0,b)} d\mathbf{u}d\mathbf{k} \text{Tr}[\phi^{\mathbf{u}+\mathbf{k}} m(\mathbf{k})] = \text{Tr}[\tilde{\phi}_c \tilde{m}_b],$$

where we have written

$$\tilde{\phi}_c = \int_{B(0,c)} \phi^{\mathbf{u}} d\mathbf{u}, \quad \tilde{m}_b = \int_{B(0,b)} m_{\mathbf{k}}(\mathbf{k}) d\mathbf{k}.$$

Upon writing $v_c := \int_{B(0,c)} \|\mathbf{u}\|^2 \phi^{\mathbf{u}} d\mathbf{u}$, we get similarly $f(\mathcal{D}_2) = \text{Tr}[v_c \tilde{m}_b]$. Note that by rotational symmetry v_c and $\tilde{\phi}_c$ are diagonal in the number operator eigenbasis, so without restricting the generality, we may assume that $\tilde{m}_b$ is also diagonal in that basis: $\tilde{m}_b = \sum_k p_k |k\rangle\langle k|$. We have then

$$\frac{f(\mathcal{D}_2)}{g(\mathcal{D}_2)} = \frac{\sum_{k \in \mathbb{N}} p_k \langle k | v_c | k \rangle}{\sum_{k \in \mathbb{N}} p_k \langle k | \tilde{\phi}_c | k \rangle} \geq \inf_{k \in \mathbb{N}} \frac{\langle k | v_c | k \rangle}{\langle k | \tilde{\phi}_c | k \rangle}.$$

The infimum on the right side is achieved by the vacuum vector. By Lemma (7.5), this fact follows from the inequality

$$\frac{\langle k | \phi^{\mathbf{u}_1} | k \rangle}{\langle k | \phi^{\mathbf{u}_2} | k \rangle} \geq \frac{\langle 0 | \phi^{\mathbf{u}_1} | 0 \rangle}{\langle 0 | \phi^{\mathbf{u}_2} | 0 \rangle}, \quad \|\mathbf{u}_1\| \geq \|\mathbf{u}_2\|,$$

which can be checked by explicit calculations.

Letting now c and b go to infinity with $c = o(b)$ and using Eq. (7.5), we obtain that

$$\lim_{a \to \infty} C(\mathcal{G}_a) \geq \frac{\int_{\mathbb{R}^2} \langle 0 | \phi^{\mathbf{u}} | 0 \rangle \|\mathbf{u}\|^2 d\mathbf{u}}{\int_{\mathbb{R}^2} \langle 0 | \phi^{\mathbf{u}} | 0 \rangle d\mathbf{u}},$$

which is exactly the point-wise risk of the heterodyne measurement $H(d\mathbf{u}) = h(\mathbf{u}) d\mathbf{u}$ whose density is

$$h(\mathbf{u}) = (2\mu - 1) D(-\sqrt{2\mu - 1} \, \alpha_{\mathbf{u}}) |0\rangle \langle 0| D(-\sqrt{2\mu - 1} \, \alpha_{\mathbf{u}}).$$

By symmetry, this point-wise risk does not depend on the point, so that $C(\mathcal{G}) \leq R_H(\mathbf{u}, \hat{\mathbf{u}})$. And we have our second inequality: $\lim_{a \to \infty} C(\mathcal{G}_a) \geq C(\mathcal{G})$.

Moreover, the heterodyne measurement is known to saturate the Holevo bound for $G = Id$ and the Cramér–Rao bound for locally unbiased estimators [29,39]. We conclude that the local minimax risk for qubits is equal to the minimax risk for the limit Gaussian quantum experiment which is achieved by the heterodyne measurement.

Lemma (7.5). *Let p and q be two probability densities on $[0, 1]$ and assume that*

$$\frac{p(x_1)}{p(x_2)} \geq \frac{q(x_1)}{q(x_2)}, \quad x_1 \geq x_2.$$

Then, $\int x^2 p(x) dx \geq \int x^2 q(x) dx$.

Proof. It is enough to show that there exists a point $x_0 \in [0, 1]$ such that $p(x) \leq q(x)$ for $x \leq x_0$ and $p(x) \geq q(x)$ for $x \geq x_0$. Now, if $p(x) \leq q(x)$ then by using the assumption, we get that $p(y) \leq q(y)$ for all $y \leq x$. Similarly, if $p(x) \geq q(x)$ then $p(y) \leq q(y)$ for all $y \geq x$. This implies the existence of the crossing point x_0.

By putting the last two lemmas together, we obtain the following.

Proposition (7.6). *The local asymptotic minimax risk $C_{l.a.m}(\mathcal{F}_n : n \in \mathbb{N})$ for the qubit state estimation problem is equal to the minimax risk $C(\mathcal{G})$ for the corresponding quantum Gaussian shift experiment which is achieved by the heterodyne measurement.*

The natural question is now the following: Is there a sequence of measurement-estimator pairs for the qubits which achieves this the risk $C_{l.a.m}(\mathcal{F}_n : n \in \mathbb{N})$ asymptotically for *all* local neighborhoods simultaneously, i.e., without prior knowledge of the center ρ^0 of the $1/\sqrt{n}$ ball within which the true state lies. Intuitively, the following procedure seems natural: Use the local asymptotic normality to transfer the

heterodyne measurement from the space of the oscillator to that of the qubits and in this way achieve the desired asymptotic risk. However, this requires the knowledge of the local neighborhood on which the convergence holds. In order to obtain this information about the state we need to "localize" the state by performing a first stage of (rough) measurements on a small proportion of the systems of order $o(n)$ and then perform the (optimal) heterodyne-type measurements corresponding to the local neighborhood of the first stage estimator. In order to make this argument rigurous, we need some finer estimates on the region in which local asymptotic normality holds and we leave this problem for a separate work.

However, there exists another measurement which achieves the risk $C_{l.a.m}(\mathcal{F}_n : n \in \mathbb{N})$, namely the optimal measurement from the Bayesian point of view discussed in [24,29]. The connection between the local and Bayesian approaches is discussed in more detail in the next subsection to which we refer for the appropriate definitions. In particular, the next proposition can be better understood after reading the next subsection, but we state it here because it is a direct consequence of the results derived in this subsection.

Let us denote by $(M_n, \hat{\rho}_n)$ the measurement-estimator pair from [24,29] which are optimal from the Bayesian point of view.

Proposition (7.7). *The optimal measurement-estimator pair $(M_n, \hat{\mathbf{u}}_n)$ in the Bayesian setup is an asymptotically local minimax estimation scheme. That is for any ρ*

$$\lim_{n \to \infty} n R_{M_n}(\rho, \hat{\rho}_n) = C_{l.a.m}(\mathcal{F}_n : n \in \mathbb{N}),$$

where $R_{M_n}(\rho, \hat{\rho}_n)$ is the risk with respect to the fidelity distance.

Proof. The pointwise risk of $(M_n, \hat{\rho}_n)$ is known to converge to that of the heterodyne measurement [24]. The rest follows from Lemma (7.3) and Lemma (7.4).

B. Local asymptotic equivalence of the optimal Bayesian measurement and the heterodyne measurement

In this subsection, we will continue our comparison of the pointwise (local) point of view with the global one used in the Bayesian approach. The result is that the optimal $SU(2)$ covariant measurement [24,29] converges locally to the optimal measurement for the family of displaced Gaussian states which is a heterodyne measurement [39]. Together with the results on the asymptotic local minimax optimality of this measurement, this closes a circle of ideas relating the different optimality notions and the relations between the optimal measurements.

Let us recall what are the ingredients of the state estimation problem in the Bayesian framework [24]. We choose as cost function the fidelity squared $F(\rho, \sigma)^2 = \text{Tr}(\sqrt{\sqrt{\rho} \sigma \sqrt{\rho}})^2$ and fix a prior distribution π over all states in $\mathbb{C}^2$ which is invariant under the $SU(2)$ symmetry group. Given n identical systems $\rho^{\otimes n}$ we would like to find a measurement M_n—whose outcome is the estimator $\hat{\rho}_n$—which *maximizes*

$$R_{\pi,n} := \int \langle F(\hat{\rho}_n, \rho)^2 \rangle \pi(d\rho).$$

By the $SU(2)$ invariance of π, the optimal measurement can be chosen to be $SU(2)$ covariant, i.e.,

$$UM_n(d\sigma)U^* = M_n(U^* d\sigma U),$$

and can be described as follows. First, we use the decomposition (4.5) to make a "which block" measurement and obtain a result j and the conditional state $\rho_{j,n}$ as in Eq. (4.6). This part will provide us the eigenvalues of the estimator. Next, we perform block wise the covariant measurement $M_{j,n}(d\vec{s}) = m_{j,n}(\vec{s})d\vec{s}$ with

$$m_{j,n}(\vec{s}) := (2j+1)U_j(\vec{s})^* |j\rangle\langle j| U_j(\vec{s}) \otimes \mathbf{1}_j,$$

whose result is a unit vector $\vec{s}$ where $U(\vec{s})$ is a unitary rotating the vector state $|\vec{s}\rangle$ to $|\uparrow\rangle$. The complete estimator is then $\hat{\rho}_n = \frac{1}{2}(\mathbf{1} + \frac{2j}{n}\vec{s}\vec{\sigma})$.

We pass now to the description of the heterodyne measurement for the quantum harmonic oscillator. This measurement has outcomes $\mathbf{u} \in \mathbb{R}^2$ and is covariant with respect to the translations induced by the displacement operators $D(\mathbf{z})$ such that $H(d\mathbf{u}) = h(\mathbf{u})d\mathbf{u}$ with

$$h(\mathbf{u}) := (2\mu - 1)D(-\sqrt{2\mu - 1}\,\alpha_{\mathbf{u}})|0\rangle\langle 0|D(\sqrt{2\mu - 1}\,\alpha_{\mathbf{u}}).$$

Using Theorem (1.1), we can map H into a measurement on the n-spin system as follows: First we perform the which block step as in the case of the $SU(2)$-covariant measurements. Then we map $\rho_{j,n}$ into an oscillator state using the isometry V_j [see Eq. (5.1)], and subsequently we perform H. The result $\mathbf{u}$ will define our estimator for the local state, i.e.,

$$\hat{\rho}_n = U\left(\frac{\mathbf{u}}{\sqrt{n}}\right)\begin{pmatrix} \frac{1}{2} + \frac{j}{n} & 0 \\ 0 & \frac{1}{2} - \frac{j}{n} \end{pmatrix} U\left(\frac{\mathbf{u}}{\sqrt{n}}\right)^*. \quad (7.6)$$

We denote by H_n the resulting measurement with values in the states on $\mathbb{C}^2$.

The next Theorem shows that in a *local neighborhood* of a fixed state ρ^0, the $SU(2)$-covariant measurement M_n and the heterodyne-type measurement H_n are asymptotically equivalent in the sense that the probability distributions $P(M_n, \rho)$ and $P(H_n, \rho)$ are close to each other uniformly over all local states ρ such that $\|\rho - \rho^0\|_1 \leq \frac{C}{\sqrt{n}}$ for a fixed but arbitrary constant $C < \infty$.

Theorem (7.8). *Let ρ^0 be as in Eq. (4.1), and let*

$$B_n(I) = \{\rho^{\mathbf{v}/\sqrt{n}} : \mathbf{v} \in I^2\}, \quad ,|I| < \infty,$$

be a local family of states around ρ^0. Then,

$$\lim_{n\to\infty} \sup_{\rho \in B_n(I)} \|P(M_n, \rho) - P(H_n, \rho)\|_1 = 0.$$

Proof. Note first that both $P(M_n, \rho)$ and $P(H_n, \rho)$ are distributions over the Bloch sphere and the marginals over the length of the Bloch vectors are identical because by construction the first step of both measurements is the same. Then,

$$\|P(M_n, \rho) - P(H_n, \rho)\|_1 =$$

$$\sum_j p_n(j) \int |\text{Tr}(\rho_{j,n}(m_{j,n}(\vec{s}) - h_{j,n}(\vec{s})))| d\vec{s}.$$

According to Eq. (4.10), we can restrict the summation to the interval $\mathcal{J}_{n,\epsilon}$ around $j = n(\mu - \frac{1}{2})$. By Theorem (1.1), we can replace (whenever needed) the local states $\rho_{j,n}^{\mathbf{v}/\sqrt{n}}$ by their limits in the oscillator space $\phi^{\mathbf{v}}$ with an asymptotically vanishing error, uniformly over $\mathbf{v} \in I^2$.

We make now the change of variable $\vec{s} \to \mathbf{u}$, where $\mathbf{u} \in \mathbb{R}^2$ belongs to the ball $|\mathbf{u}| < 2\sqrt{n}\pi$, and is the smallest vector such that $U(\frac{\mathbf{u}}{\sqrt{n}}) = U(\vec{s})$.

The density of the $SU(2)$ estimator with respect to the measure $d\mathbf{u}$ is

$$m_{j,n}(\mathbf{u}) := \frac{2j+1}{n}U_j\left(\frac{\mathbf{u}}{\sqrt{n}}\right)^* |j\rangle\langle j| U_j\left(\frac{\mathbf{u}}{\sqrt{n}}\right) J\left(\frac{\mathbf{u}}{\sqrt{n}}\right),$$

where J is the determinant of a Jacobian related with the change of variables such that $J(0) = 1$.

Similarly, the density of the homodyne-type estimator becomes

$$h_{j,n}(\mathbf{u}) := \sum_{k \in \mathbb{N}} V_j^* h\left(\mathbf{u} + 2k\sqrt{n}\pi\frac{\mathbf{u}}{|\mathbf{u}|}\right) V_j |J_{k,n}(\mathbf{u})|,$$

because displacements in the same direction which differ by multiples of $2\sqrt{n}\pi$ lead to the same unitary on the qubits. Here, $J_{k,n}(\mathbf{u})$ is again the determinant of the Jacobian of the map from the kth ring to the disk, in particular $J_{0,n}(\mathbf{u}) = 1$.

The integral becomes then

$$\int_{|\mathbf{u}| \leq 2\pi\sqrt{n}} |\text{Tr}(\rho_{j,n}^{\mathbf{v}/\sqrt{n}}(m_{j,n}(\mathbf{u}) - h_{j,n}(\mathbf{u})))| d\mathbf{u}.$$

We bound this integral by the sum of two terms, the first one being

$$\int_{|\mathbf{u}| \leq 2\pi\sqrt{n}} |\text{Tr}(\rho_{j,n}^{\mathbf{v}/\sqrt{n}}(m_{j,n}(\mathbf{u}) - \widetilde{h}_j(\mathbf{u})))| d\mathbf{u},$$

where $\widetilde{h}_j(\mathbf{u})$ is just the term with $k=0$ in $h_{j,n}(\mathbf{u})$. By Lemma (5.4), for any fixed $\mathbf{u}$, we have $m_{j,n}(\mathbf{u}) \to h(\mathbf{u})$ uniformly over $j \in \mathcal{J}_{n,\epsilon}$. Using similar estimates as in Lemma (5.4), it can be shown that the function under the integral is bounded by a fixed integrable function $g(\mathbf{u})$ uniformly over $\mathbf{v} \in I^2$, and then we can use dominated convergence to conclude that the integral converges to 0 uniformly over $\mathbf{v} \in I^2$ and $j \in \mathcal{J}_{n,\epsilon}$.

The second integral is

$$\int_{|\mathbf{u}| \leq 2\pi\sqrt{n}} |\text{Tr}(\rho_{j,n}^{\mathbf{v}/\sqrt{n}}(\widetilde{h}_j(\mathbf{u}) - h_{j,n}(\mathbf{u})))| d\mathbf{u},$$

which is smaller than

$$\int_{|\mathbf{u}| > 2\pi\sqrt{n}} |\text{Tr}(\rho_{j,n}^{\mathbf{v}/\sqrt{n}} h(\mathbf{u}))| d\mathbf{u},$$

which converges uniformly to 0. This can be seen by replacing the states with $\phi^{\mathbf{v}}$ which are "confined" to a fixed region of the size I^2 in the phase space, while the terms $h(\mathbf{u})$ are Gaussians located at distance at least $2\pi\sqrt{n}$ from the origin.

Putting these two estimates together, we obtain the desired result. $\square$

Remark. The result in the above theorem holds more generally for all states in a local neighborhood of $\rho^{\mathbf{0}}$ but for the proof we need a slightly more general version of Theorem (1.1), where the eigenvalues of the density matrices are not fixed but allowed to vary in a local neighborhood of $(\mu, 1-\mu)$. This result will be presented in a future work concerning the general case of d-dimensional states.

C. Discrimination of states

Another illustration of the local asymptotic normality Theorem is the problem of discriminating between two states ρ^+ and ρ^-. When the two states are fixed, this problem has been solved by Helstrom [44], and if we are given n systems in state $\rho_{\pm}^{\otimes n}$ then the probability of error converge to 0 exponentially. Here, we consider the problem of distinguishing between two states $\rho_n^{\pm}$ which approach each other as $n \to \infty$ with rate $\|\rho_n^+ - \rho_n^-\|_1 \approx \frac{1}{\sqrt{n}}$. In this case, the probability of error does not go to 0 because the problem becomes more difficult as we have more systems, and converges to the limit problem of distinguishing between two fixed Gaussian states of a quantum oscillator.

This problem is interesting for several reasons. First, it shows that the convergence in Theorem (1.1) can be used for finding asymptotically optimal procedures for various statistical problems such as that of parameter estimation and hypothesis testing. Second, for any fixed n, the optimal discrimination is performed by a rather complicated *joint* measurement and the hope is that the asymptotic problem of discriminating between two Gaussian states may provide a more realistic measurement which can be implemented in the lab. Third, this example shows that a noncommuting one-parameter families of states is not "classical" as it is sometimes argued, but should be considered as a quantum "resource" which cannot be transformed into a classical one without loss of information. More explicitly, the optimal measurement for estimating the parameter is not optimal for other statistical problems, such as the one considered here, and thus different statistical decision problems are accompanied by mutually incompatible optimal measurements.

Let us recall the framework of quantum hypothesis testing for two states $\rho^{\pm}$: We consider two-outcomes POVM's $M = (M_-, M_+)$ with $0 \leq M_+ \leq \mathbf{1}$ and $M_- = \mathbf{1} - M_+$ such that the probability of error when the state is ρ^- is given by $\mathrm{Tr}(M_+ \rho^-)$, and similarly for ρ^+. As we do not know the state, we want to minimize our worst-case probability error. Our figure of merit (the lower, the better) is therefore:

$$R(\rho^+, \rho^-) = \inf_M \max\{\mathrm{Tr}(\rho_+ M_-), \mathrm{Tr}(\rho_+ M_-)\}.$$

Now we are interested in the case when $\rho^{\pm} = \rho^{\pm \mathbf{u}_n}$ as defined in Eq. (1.1), and in the limit $\rho_{\pm} = \phi^{\pm \mathbf{u}}$ (recall that both $\rho_n^{\mathbf{u}}$ and $\phi^{\mathbf{u}}$ depend on μ). We then have:

Theorem (7.9). *The following limit holds*

$$\lim_{n \to \infty} R(\rho_n^{\mathbf{u}}, \rho_n^{-\mathbf{u}}) = R(\phi^{\mathbf{u}}, \phi^{-\mathbf{u}}).$$

Moreover for pure states this limit is equal to $(1 - (1 - e^{-4|\mathbf{u}|^2})^{1/2})/2$ *which is strictly smaller than* $1/2 - \mathrm{erf}(|\mathbf{u}|)$

which is the limit if we do not use collective measurements on the qubits. Here we have used this convention for the error function: $\mathrm{erf}(x) = \int_0^x e^{-t^2}/\sqrt{\pi} dt$.

Proof. Let M be the optimal discrimination procedure $\phi^{\pm \mathbf{u}}$. Then, we use the channel T_n to send $\rho_n^{\pm \mathbf{u}}$ to states of the oscillator and then perform the measurement M. By Theorem (1.1), $\|\phi^{\pm \mathbf{u}} - T_n(\rho_n^{\pm \mathbf{u}})\|_1 \to 0$ so that $\mathrm{Tr}(T_n(\rho_n^{\pm \mathbf{u}}) M_{\mp}) \to \mathrm{Tr}(\phi^{\pm \mathbf{u}} M_{\mp})$. Thus, $M \circ T_n$ is asymptotically optimal for $\rho_n^{\pm \mathbf{u}}$.

Now for pure states $|\psi_+\rangle$ and $|\psi_-\rangle$, the optimal measurement is well known [45,46]. It is unique on the span of these pure states and arbitrary on the orthogonal. If we choose the phase such that $\langle \psi_- | \psi_+ \rangle > 0$, then M_+ is the projector on the vector

$$\frac{|\psi_+\rangle + |\psi_-\rangle}{2\sqrt{1 + \langle \psi_- | \psi_+ \rangle}} + \frac{|\psi_+\rangle - |\psi_-\rangle}{2\sqrt{1 - \langle \psi_- | \psi_+ \rangle}},$$

and the associated risk is

$$\frac{1}{2}(1 - \sqrt{1 - |\langle \psi_+ | \psi_- \rangle|^2}).$$

Now in our case, in the limit experiment, $\phi^{\mathbf{u}}$ is the coherent state $|\psi_{\mathbf{u}}\rangle = e^{-|\mathbf{u}|^2/2} \Sigma_n |\mathbf{u}|^n / \sqrt{n!} |n\rangle$. So that

$$\langle \psi_{\mathbf{u}} | \psi_{-\mathbf{u}} \rangle = e^{-|\mathbf{u}|^2} \sum_n \frac{(-|\mathbf{u}|^2)^n}{n!} = e^{-2|\mathbf{u}|^2},$$

and $R(\phi^{\mathbf{u}}, \phi^{-\mathbf{u}}) = \frac{1}{2}(1 - \sqrt{1 - e^{-4|\mathbf{u}|^2}})$.

We would like to insist here that the best measurement for discrimination is not measuring the positive part of the position observable Q (we assume by symmetry that $\pm \mathbf{u}$ is on the first coordinate), as one might expect from the analogy with the classical problem. Indeed, if we meausure Q then we obtain a classical Gaussian variable with density $p(x) = e^{-(x-|\mathbf{u}|)^2}/\sqrt{\pi}$ and the best guess at the sign $\pm$ has in this case the risk $1/2 - \mathrm{erf}(|\mathbf{u}|)$.

This may be a bit surprising considering that measuring Q preserves the quantum Fisher information. The conclusion is simply that the quantum Fisher information is not an exhaustive indicator of the statistical information in a family of states, as it may remain unchanged even when there is a clear degradation in the inference power. This example fits in a more general framework of a theory of quantum statistical experiments and quantum decisions [47].

D. Spin squeezed states and continuous time measurements

In an emblematic experiment for the field of quantum filtering and control [35], it is shown how spin squeezed states can be prepared deterministically by using continuous time measurements performed in the environment and real time feedback on the spins. Without going in the details, the basic idea is to describe the evolution of identically prepared spins by passing first to the coherent state picture. There one can easily solve the stochastic Schrödinger equation describing the evolution (quantum trajectory) of the quantum oscillator conditioned on the continuous signal of the measurement device. The solution is a Gaussian state whose center

evolves stochastically while one of the quadratures gets more and more squeezed as one obtains more information through the measurement. Using feedback, one can then stabilize the center of the state around a fixed point.

This description is of course approximative and holds as long as the errors in identifying the spins with Gaussian states are not significant. The framework developed in the proof of Theorem (1.1) can then be used to make more precise statements about the validity of the results, including the squeezing process.

Perhaps more interesting for quantum estimation, such measurements may be used to perform optimal estimation of spin states. The idea would be to first localize the state in a small region by performing a weak measurement and then in a second stage, one performs a heterodyne type measurement after rotating the spins so that they point approximately in the z direction. We believe that this type of procedure has better chances of being implemented in practice compared with the abstract covariant measurement of [24,29].

ACKNOWLEDGMENTS

[1] G. Breitenbach, S. Schiller, and J. Mlynek, Nature (London) **387**, 471 (1997).

[2] K. Vogel and H. Risken, Phys. Rev. A **40**, R2847 (1989).

[3] G. M. D'Ariano, U. Leonhardt, and H. Paul, Phys. Rev. A **52**, R1801 (1995).

[4] U. Leonhardt, H. Paul, and G. M. D'Ariano, Phys. Rev. A **52**, 4899 (1995).

[5] U. Leonhardt, M. Munroe, T. Kiss, T. Richter, and M. G. Raymer, Opt. Commun. **127**, 144 (1996).

[6] L. Artiles, M. I. Guţă, and R. D. Gill, J. R. Stat. Soc. Ser. B. Methodol. **67**, 109 (2005).

[7] D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani, Phys. Rev. Lett. **70**, 1244 (1993).

[8] A. Zavatta, S. Viciani, and M. Bellini, Science **306**, 660 (2004).

[9] R. D. Gill and S. Massar,Phys. Rev. A **61**, 042312 (2000).

[10] K. R. W. Jones, Phys. Rev. A **50**, 3682 (1994).

[11] S. Field, N. Venturi, and F. Nori, Phys. Rev. Lett. **74**, 1259 (1995).

[12] J. I. Latorre, P. Pascual, and R. Tarrach, Phys. Rev. Lett. **81**, 1351 (1998).

[13] D. G. Fischer, S. H. Kienle, and M. Freyberger, Phys. Rev. A **61**, 032306 (2000).

[14] T. Hannemann, D. Reiss, C. Balzer, W. Neuhauser, P. E. Toschek, and C. Wunderlich, Phys. Rev. A **65**, 050303(R) (2002).

[15] E. Bagan, M. Baig, and R. Munoz-Tapia, Phys. Rev. Lett. **89**, 277904 (2002).

[16] F. Embacher and H. Narnhofer, Ann. Phys. (N.Y.) **311**, 220 (2004).

[17] E. Bagan, A. Monras, and R. Munoz-Tapia, Phys. Rev. A **71**, 062318 (2005).

[18] J. I. Cirac, A. K. Ekert, and C. Macchiavello, Phys. Rev. Lett. **82**, 4344 (1999).

[19] G. Vidal, J. I. Latorre, P. Pascual, and R. Tarrach, Phys. Rev. A **60**, 126 (1999).

[20] H. Mack, D. G. Fischer, and M. Freyberger, Phys. Rev. A **62**, 042301 (2000).

[21] M. Keyl and R. F. Werner, Phys. Rev. A **64**, 052311 (2001).

[22] E. Bagan, M. Baig, R. Munoz-Tapia, and A. Rodriguez, Phys. Rev. A **69**, 010304(R) (2004).

[23] K. Zyczkowski and H. J. Sommers, Phys. Rev. A **71**, 032313 (2005).

[24] E. Bagan, M. A. Ballester, R. D. Gill, A. Monras, and R. Munõz Tapia, quanth-ph/0510158.

[25] M. Hayashi, J. Phys. A **31**, 4633 (1998).

[26] O. E. Barndorff-Nielsen and R. Gill, J. Phys. A **33**, 1 (2000).

[27] K. Matsumoto, J. Phys. A **35**, 3111 (2002).

[28] O. E. Barndorff-Nielsen, R. Gill, and P. E. Jupp, J. R. Stat. Soc. Ser. B. Methodol. **65**, 775 (2003).

[29] M. Hayashi and K. Mastumoto, quant-ph/0411073.

[30] R. Gill, math.ST/0512443 (unpublished).

[31] A. van der Vaart, *Asymptotic Statistics* (Cambridge University Press, Cambridge, U.K., 1998).

[32] R. Holtz and J. Hanus, J. Phys. A **7**, 37 (1974).

[33] M. Kitagawa and M. Ueda, Phys. Rev. A **47**, 5138 (1992).

[34] F. J. Dyson, Phys. Rev. **102**, 1217 (1956).

[35] J. Geremia, J. K. Stockton, and H. Mabuchi, Science **304**, 270 (2004).

[36] M. Ohya and D. Petz, *Quantum Entropy and its Use* (Springer Verlag, Berlin-Heidelberg, 2004).

[37] L. Accardi and A. Bach, in *Quantum Probability and Applications IV*, edited by Luigi Accardi and Wilhelm von Wandelfels (Springer, New York, 1987), Lecture Notes in Mathematics, Vol. 1396, pp. 7–19.

[38] L. Accardi and A. Bach, Probab. Theory Relat. Fields **68**, 393 (1985).

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

[40] D. Petz and A. Jencova, Commun. Math. Phys. **263**, 259 (2006).

[41] D. Petz, *An Invitation to the Algebra of Canonical Commutation Relations* (Leuven University Press, Leuven, Belgium, 1990).

[42] J. K. Stockton, J. M. Geremia, A. C. Doherty, and H. Mabuchi, Phys. Rev. A **67**, 022112 (2003).

[43] A. Chefles, R. Josza, and A. Winter, quant-ph/0307227.

[44] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).

[45] G. M. D'Ariano, M. F. Sacchi, and J. Kahn, Phys. Rev. A **72**, 032310 (2005).

[46] A. Chefles, J. Contemp. Phys. **41**, 401 (2000).

[47] M. Guţă, (unpublished).