Adaptive quantum tomography in high dimensions

L. Pereira,[1,2] L. Zambrano,[1,2] J. Cortés-Vega,[1,2] S. Niklitschek,[3] and A. Delgado[1,2,*]

[1]*Centro de Óptica y Fotónica e Instituto Milenio de Investigación en Óptica, Universidad de Concepción, Concepción, Chile*

[2]*Facultad de Ciencias Físicas y Matemáticas, Departamento de Física, Universidad de Concepción, Concepción, Chile*

[3]*Facultad de Ciencias Físicas y Matemáticas, Departamento de Estadística, Universidad de Concepción, Concepción, Chile*

(Received 23 March 2018; published 31 July 2018)

Standard quantum tomography of a single qudit achieves an infidelity that scales in the worst case as $O(1/\sqrt{N})$ for a sample of size N. Here, we propose a suitable generalization of the two-stage adaptive quantum tomography for a qubit to the case of a single qudit. This achieves an infidelity of the order of $O[1/\sqrt{N_0(N-N_0)}]$ for all quantum states, where N_0 and $N-N_0$ are the ensemble sizes employed in the two stages of the method. This result is based on a second-order Taylor series expansion of the infidelity that is obtained by means of the Fréchet derivative and measurement outcomes modeled by a multinomial distribution. Numerical simulations indicate that the choice $N_0 = N/2$ leads to an infidelity that scales approximately as $O(1/N)$ for all quantum states in a wide range of dimensions, that is, a quadratic improvement of the infidelity when compared to standard quantum tomography in the case of low-rank states.

DOI: 10.1103/PhysRevA.98.012339

I. INTRODUCTION

Study and applications of high-dimensional quantum systems are intensive research topics in quantum information theory. This class of systems provides several advantages over two-dimensional quantum systems. For instance, it has been shown that the use of d-dimensional quantum systems (qudits) in quantum communications leads to more efficient, secure, and noise-resistant protocols [1–3]. Similarly, quantum computing algorithms can be implemented through qudits [4–6] more efficiently. Qudits also allow us to design noise-tolerant Bell-like inequalities that lead to larger violations [7,8].

In spite of the advantages provided by qudits, the assessment of experimentally generated quantum states of higher dimensional systems and quantum devices has proven to be a remarkably difficult task [9]. To accomplish this goal, several quantum tomographic methods to estimate unknown states have been designed and experimentally demonstrated such as, for instance, standard quantum tomography [10,11], mutually unbiased bases (MUBs)–based tomography [12–18], and symmetric informationally complete (SIC) positive-operator valued measure (POVM)–based quantum tomography [19–25], among others [26,27]. In general, these methods are based on the measurement of several POVMs on an ensemble of N identically prepared copies of the state ρ to be estimated. The experimentally acquired data is postprocessed by means of statistical inference methods [28–31] to provide a physically acceptable estimate $\tilde{\rho}$ of the unknown state.

In order to reduce the experimental complexity in the determination of quantum states, quantum tomographic methods were initially designed to employ the smallest possible number of measurement outcomes. Nevertheless, this number scales exponentially in the case of multipartite systems, which makes quantum tomography unfeasible but in systems with few components. This led to the search for new methods [32–34] with a better scaling. Typically, these exploit *a priori* information about the state to be estimated such as, for instance, rank or purity.

Recently, quantum tomographic methods have been studied from the point of view of the achievable accuracy as a function of the ensemble size N. This is another costly resource in experimental realizations. It has been noted that standard quantum tomography of a single qubit has a rank-dependent accuracy. The infidelity between a state and its estimate can be used to quantify the accuracy of a tomographic method. In this case, standard quantum tomography achieves an infidelity that scales as $O(1/\sqrt{N})$ and $O(1/N)$ for pure and mixed states, respectively. The former resembles a classical behavior, while the latter approaches the Gill-Massar [35] bound. This establishes the highest possible accuracy achievable by means of separable measurements on a given ensemble.

This unwanted feature of standard quantum tomography can be overcome with the help of two-stage adaptive quantum tomography for a qubit [36–38]. In the first stage, standard quantum tomography is carried out on an ensemble of size N_0 and an estimate $\tilde{\rho}_0$ is obtained. The eigenstates of this estimate are then used to construct three new mutually unbiased bases, which are subsequently employed to perform a second stage of standard quantum tomography on an ensemble of size $N-N_0$ followed by the generation of a maximum likelihood estimate $\tilde{\rho}$. This tomographic method attains an infidelity that scales as $O(1/N)$ for all states. Furthermore, two-stage adaptive quantum tomography becomes very close to the Gill-Massar bound for several figures of merit for the accuracy [39] such as, for instance, mean square error or Bures distance. Besides, it has been proven that an adaptive method improves quadratically the accuracy in the estimation of the spectrum of a density matrix [40].

*aldelgado@udec.cl

Here, we present a generalization of the two-stage adaptive quantum tomographic method to the case of a single qudit and analyze its properties by means of analytical calculations and numerical simulations. This is mainly motivated by the need to improve the accuracy in the assessment of states and quantum devices of higher dimensional systems that find application in quantum information theory. We employ the infidelity [41] between the unknown state ρ and its estimate $\tilde{\rho}$ as figure of merit for the accuracy. This choice is dictated by the agreement between the infidelity and the Bures distance in the infinitesimal case [42]. Furthermore, the inverse of the infidelity can be identified with the ensamble size required to achieve a prescribed accuracy [37]. First, we calculate a Taylor series expansion for the infidelity up to second order in the error $\rho - \tilde{\rho}$ by means of the Fréchet derivative [43], which does not requires a parametrization of quantum states. This expansion shows clearly that the behavior of the infidelity depends on the rank of the state to be estimated. For instance, in the case of rank-1 states dominates the first-order term while for full rank states dominates the second-order term. The expansion also allows us to show that the first-order term vanishes for any finite, arbitrary dimension d when standard tomography for a qudit is carried out representing the generators of the group SU(d) in the eigenbasis of ρ. Then, we show that the second-order term scales as $O(1/N)$ for all states. However, such a tomography on the eigenbasis of ρ is impossible since ρ is unknown. This leads to the two-stage adaptive tomography. In the first stage, standard quantum tomography is performed on the unknown state employing an ensemble of size N_0. After a linear inversion process, an estimate $\tilde{\rho}_0$ is obtained and diagonalized. The difference between the eigenvalues of $\tilde{\rho}_0$ and ρ scales as $O(1/N_0)$. Thereafter, the eigenstates of $\tilde{\rho}_0$ are employed to represent the generators of the group SU(d). These are then measured in a second stage of standard quantum tomography. This tomographic method does not cancel the first-order term in the Taylor series expansion of the infidelity. This, however, scales now as $O[1/\sqrt{(N-N_0)N_0}]$. With the choice $N_0 = N/2$, the first-order term scales akin to the the second-order term, that is, as $O(1/N)$. Thereby, the two-stage adaptive quantum tomographic method attains an infidelity which scales as $O(1/N)$ for all states of a single qudit. We also perform numerical simulations in a wide range of dimensions to study the accuracy in the determination of the infidelity, the impact of the choice of N_0, and the role played by the rank of the states to be estimated.

This article is organized as follows: In Sec. I, we briefly recall the expansion of a single-qudit state in terms of the generators of SU(d). In Sec. II, we deduce the Taylor series expansion of the infidelity employing the notion of Fréchet derivative. In Sec. III, we introduce the two-stage adaptive quantum tomographic method for a single qudit and study its main properties. In Sec. IV, we illustrate the properties of the method by means of numerical simulations addressing the relevant cases of pure states, pure states affected by white noise, and full-rank states. In Sec. V, we summarize, comment possible experimental implementations and extensions of the two-stage adaptive tomography, and conclude.

II. STATES OF A SINGLE QUDIT

States of a d-dimensional quantum system are mathematically described by means of positive semidefinite operators with unitary trace. A representation of quantum states can be obtained by means of the generalized Gell-Mann representation of the $d^2 - 1$ Hermitian, traceless generators of the special unitary group SU(d) together with the identity operator.

The generators of SU(d) are given by a set of $d - 1$ diagonal operators

$$\sigma_k^z = \frac{\sqrt{2}}{\sqrt{k(k+1)}}\left(\sum_{j=1}^{k} |j\rangle\langle j| - k|k+1\rangle\langle k+1|\right), \tag{1}$$

with $k = 1, \ldots, d-1$ and $d(d-1)$ antidiagonal operators given by

$$\sigma_{ij}^x = |i\rangle\langle j| + |j\rangle\langle i| \tag{2}$$

and

$$\sigma_{ij}^y = -i(|i\rangle\langle j| - |j\rangle\langle i|), \tag{3}$$

with $1 \leqslant i < j \leqslant d$. The set $\{|i\rangle\}$ of states with $i = 1, \ldots, d-1$ is an arbitrary orthonormal basis. For sake of simplicity, these operators can be labeled with a single index as

$$\sigma_{(j-1)^2+2(i-1)} = \sigma_{ij}^x, \tag{4}$$

$$\sigma_{(j-1)^2+2i-1} = \sigma_{ij}^y, \tag{5}$$

$$\sigma_{j^2-1} = \sigma_{j-1}^z, \tag{6}$$

where operations involving subindexes are carried out mod(d).

Any quantum state ρ can be uniquely represented as

$$\rho = \frac{1}{d}\mathbb{I} + \frac{1}{2}\left[\sum_{ij}\left(S_{ij}^x\sigma_{ij}^x + S_{ij}^y\sigma_{ij}^y\right) + \sum_k S_k^z\sigma_k^z\right], \tag{7}$$

where the $d^2 - 1$ real coefficients $\{S_{ij}^x, S_{ij}^y, S_k^z\}$ entering in Eq. (7) are the average of the observables $\{\sigma_{ij}^x, \sigma_{ij}^y, \sigma_k^z\}$ on the state ρ, respectively. In an experimental implementation of standard quantum tomography, each average is obtained by projecting onto the eigenstates of the observable to be measured. Considering the eigenvalue problem $\sigma_k|i,k\rangle = \lambda_i^k|i,k\rangle$, the projection probabilities are given by $p_i^k = \mathrm{Tr}(\rho|i,k\rangle\langle i,k|)$ and the averages by $S_k = \sum_i \lambda_i^k p_i^k$. These quantities are estimated as $\tilde{p}_i^k = n_i^k/N_k$ and $\tilde{S}_k = \sum_i \lambda_i^k \tilde{p}_i^k$, where n_i^k is the number of projections onto the eigenstate $|i,k\rangle$ when employing an ensemble of size N_k.

Assuming an ideal experiment, that is, an experiment where the only source of error is the size N of the ensemble of identically prepared copies, the projection probabilities exhibit multinomial noise with covariance matrix given by

$$\mathrm{Cov}\left(\tilde{p}_i^k, \tilde{p}_j^k\right) = \frac{1}{N_k}p_i^k\left[\delta_{ij} - p_j^k\right], \tag{8}$$

for $k = 1, \ldots, d^2 - 1$. Thereby, the variance associated to the

estimation of the average of each observable becomes

$$\mathrm{Var}(\tilde{S}_{ij}^x) = \frac{1}{N_{ij}^x}\left[p_i^z + p_j^z - \left(S_{ij}^x\right)^2\right], \tag{9}$$

$$\mathrm{Var}(\tilde{S}_{ij}^y) = \frac{1}{N_{ij}^y}\left[p_i^z + p_j^z - \left(S_{ij}^y\right)^2\right], \tag{10}$$

$$\mathrm{Var}(\tilde{S}_k^z) = \frac{1}{N_k^z}\left[\frac{2\sum_{n=1}^k p_n^z + 2k^2 p_{k+1}^z}{k(k+1)} - \left(S_k^z\right)^2\right], \tag{11}$$

where $p_i^z = \mathrm{Tr}(\rho|i\rangle\langle i|)$ are the probabilities of measuring the eigenstates of the operators σ_k^z and N_{ij}^x, N_{ij}^y, and N_k^z are the ensemble sizes used to measure the operators σ_{ij}^x, σ_{ij}^y, and σ_k^z, respectively. In what follows, we choose to estimate the statistics associated to the generators employing the same ensemble size, that is, $N_{ij}^x = N_{ij}^y = N_k^z = N/(d^2-1)$. From Eqs. (9), (10), and (11), we see that the uncertainty in the estimation of the averages $\{S_{ij}^x, S_{ij}^y, S_k^z\}$, described by the elements of the covariance matrix, also depends on these parameters, that is, the uncertainty in the estimation of the parameters defining ρ depends on ρ itself.

From the Bienaymé-Chebyshev inequality,

$$p(|\tilde{S}_i - S_i| \leqslant \alpha\sqrt{\mathrm{Var}(\tilde{S}_i)}) \geqslant 1 - \frac{1}{\alpha^2}, \tag{12}$$

where $\alpha \geqslant 1$, it can be seen that the error of the estimated parameters scale as the square root of the variance,

$$\tilde{S}_i - S_i \sim \sqrt{\mathrm{Var}(\tilde{S}_i)}. \tag{13}$$

For instance, the error is bounded by $5\sqrt{\mathrm{Var}(\tilde{S}_i)}$ with a probability at least of $24/25$.

The set of estimates $\{\tilde{S}_{ij}^x, \tilde{S}_{ij}^y, \tilde{S}_k^z\}$ leads to an estimate $\tilde{\rho}$ that differs from ρ by an error operator $\Delta = \tilde{\rho} - \rho$. This traceless, Hermitian operator is given by

$$\Delta = \frac{1}{2}\sum_{i=1}^{d^2-1}(\tilde{S}_i - S_i)\sigma_i. \tag{14}$$

The norm of this operator $||\Delta|| = \sqrt{\mathrm{Tr}(\Delta^2)}$ is

$$||\Delta|| = \left[\frac{1}{2}\sum_{i=1}^{d^2-1}(\tilde{S}_i - S_i)^2\right]^{1/2}. \tag{15}$$

Thereby, considering Eq. (13), the norm $||\Delta||$ of the error operator scales approximately as $O(1/\sqrt{N})$.

III. INFIDELITY

In order to study the tomographic precision, we employ the infidelity as figure of merit. This is a simple and well-motivated metric [41,42,44,45]. The infidelity between two arbitrary states ρ and $\tilde{\rho}$ is defined by

$$I(\rho, \tilde{\rho}) = 1 - \mathrm{Tr}(\sqrt{\sqrt{\rho}\tilde{\rho}\sqrt{\rho}})^2. \tag{16}$$

The infidelity has several properties that motivate its use to quantify the accuracy of a tomographic process. The Bures metric $B(\rho, \tilde{\rho})$, a distance defined by the expression

$$B(\rho, \tilde{\rho})^2 = 2[1 - \mathrm{Tr}(\sqrt{\sqrt{\rho}\tilde{\rho}\sqrt{\rho}})], \tag{17}$$

agrees with the infidelity for infinitesimally close states [42]. The trace distance $T(\rho, \tilde{\rho})$ [37,46,47], given by the expression

$$T(\rho, \tilde{\rho}) = (1/2)\mathrm{Tr}|\rho - \tilde{\rho}|, \tag{18}$$

can be bounded by the infidelity as

$$1 - \sqrt{1 - I(\rho, \tilde{\rho})} \leqslant \tfrac{1}{2}\mathrm{Tr}|\rho - \tilde{\rho}| \leqslant \sqrt{I(\rho, \tilde{\rho})}, \tag{19}$$

where $|A| = \sqrt{A^\dagger A}$. Furthermore, the inverse $I(\rho, \tilde{\rho})^{-1}$ of the infidelity can be related to the ensemble size required to reliably distinguish two quantum states [37,48,49], allowing us to have a notion of the ensemble size necessary to reach a prescribed accuracy. Besides, the infidelity provides a simple analytical picture. Nevertheless, the infidelity is not exempt of criticism [46,47,50]. It has been shown that high-dimensional states with very low infidelity might exhibit very different physical properties, for instance, quantum discord.

Considering the infidelity between ρ and $\tilde{\rho}$ and assuming that ρ is a rank-r density matrix, the infidelity becomes

$$I(\rho, \tilde{\rho}) = 1 - \mathrm{Tr}\left(\sqrt{\rho_r^2 + \sqrt{\rho_r}\Delta_r\sqrt{\rho_r}}\right)^2, \tag{20}$$

where ρ_r and Δ_r are restrictions of ρ and Δ to the r-dimensional eigenspace of ρ, respectively. Since the norm of Δ decreases as $O(1/\sqrt{N})$, we can develop the infidelity in a Taylor series expansion. The Taylor series of a function f of bounded operators is given by

$$f(A + B) = f(A) + \sum_{m=1}^{\infty}\frac{1}{m!}D^m f(A)([B]^m), \tag{21}$$

where $D^m f(A)([B]^m)$ is the mth directional Fréchet derivative of $f(A)$ along B [43]. Let us note that this expansion does not depend on any parametrization of density matrices. Considering that $f(A) = \sqrt{A}$, the infidelity becomes

$$I(\rho, \tilde{\rho}) = 1 - \mathrm{Tr}\left[f\left(\rho_r^2 + \sqrt{\rho_r}\Delta_r\sqrt{\rho_r}\right)\right]^2. \tag{22}$$

Let $\{\lambda_i\}$ be the nonvanishing eigenvalues of ρ_r with eigenvectors $\{|\lambda_j^r\rangle\}$. Thus, the first and second derivatives of f are (see Appendix A)

$$\left[Df\left(\rho_r^2\right)\left(\sqrt{\rho_r}\Delta_r\sqrt{\rho_r}\right)\right]_{ij} = \frac{\sqrt{\lambda_i\lambda_j}\langle\lambda_i^r|\Delta_r|\lambda_j^r\rangle}{\lambda_i + \lambda_j} \tag{23}$$

and

$$\left[D^2 f\left(\rho_r^2\right)([\sqrt{\rho_r}\Delta_r\sqrt{\rho_r}]^2)\right]_{ij} = -\frac{\sqrt{\lambda_i\lambda_j}}{\lambda_i + \lambda_j}\sum_{k=1}^{r}\frac{\lambda_k\langle\lambda_i^r|\Delta_r|\lambda_k^r\rangle\langle\lambda_k^r|\Delta_r|\lambda_j^r\rangle}{(\lambda_i + \lambda_k)(\lambda_k + \lambda_j)}. \tag{24}$$

Let $\{|\lambda_i\rangle\}$ the eigenvectors of ρ, where the states with index $i = 1, \ldots, r$ are the non-null eigenvectors of ρ and the states with $i = r+1, \ldots, d$ are orthonormal states acting onto the kernel of ρ. Thereby, the Taylor series expansion of the infidelity up to second order is

$$I(\rho, \tilde{\rho}) = \sum_{i=r+1}^{d}\langle\lambda_i|\Delta|\lambda_i\rangle + \frac{1}{2}\sum_{i,k=1}^{r}\frac{|\langle\lambda_i|\Delta|\lambda_k\rangle|^2}{\lambda_i + \lambda_k} - \frac{1}{4}\left[\sum_{i=r+1}^{d}\langle\lambda_i|\Delta|\lambda_i\rangle\right]^2 + O(||\Delta||^3). \tag{25}$$

Let us note that this series holds for quantum states of arbitrary rank. In the next section, we use this approximated expression for the infidelity to extend the two-stage adaptive quantum tomographic method to the case of a single qudit.

IV. TWO-STAGE ADAPTIVE QUANTUM TOMOGRAPHY OF A QUDIT

The previous expression for the Taylor series of the infidelity, Eq. (25), exhibits a behavior that depends on the rank of the quantum state. For instance, for a rank-1 state, the first-order term dominates and thus the infidelity scales as $O(1/\sqrt{N})$. For a full-rank quantum state, the first-order term vanishes and the infidelity scales as $O(1/N)$.

This behavior of the infidelity can be corrected by eliminating the influence of the first-order term. This depends on the diagonal coefficients of Δ in the kernel of ρ. This suggests performing standard tomography in a basis that agrees with the eigenstates of ρ, that is, replacing the set $\{|i\rangle\}$ (with $i = 1, \ldots, d$) by $\{|\lambda_i\rangle\}$ (with $i = 1, \ldots, d$) in Eqs. (1), (2), and (3). The first-order term can be cast as

$$\sum_{i=r+1}^{d} \langle\lambda_i|\Delta|\lambda_i\rangle = \frac{1}{2}\sum_{j=1}^{d^2-1}\sum_{i=r+1}^{d} (\tilde{S}_i - S_i)\langle\lambda_i|\sigma_j|\lambda_i\rangle. \quad (26)$$

The nondiagonal generators vanish from this expression since it involves only diagonal parts. Thus,

$$\sum_{i=r+1}^{d} \langle\lambda_i|\Delta|\lambda_i\rangle = \frac{1}{2}\sum_{j=1}^{d-1}\sum_{i=r+1}^{d} \left(\tilde{S}_i^z - S_i^z\right)\langle\lambda_i|\sigma_j^z|\lambda_i\rangle. \quad (27)$$

This expression can be separated into two contributions:

$$\sum_{i=r+1}^{d} \langle\lambda_i|\Delta|\lambda_i\rangle = \frac{1}{2}\sum_{j=1}^{r-1}\sum_{i=r+1}^{d} \left(\tilde{S}_i^z - S_i^z\right)\langle\lambda_i|\sigma_j^z|\lambda_i\rangle + \frac{1}{2}\sum_{j=r}^{d-1}\sum_{i=r+1}^{d} \left(\tilde{S}_i^z - S_i^z\right)\langle\lambda_i|\sigma_j^z|\lambda_i\rangle. \quad (28)$$

The first term at the right-hand side vanishes. The diagonal generators entering in this term are lineal combinations of projectors onto the eigenstates of ρ, that is, the set $\{|\lambda_i\rangle\}$ with $i = 1, \ldots, r$. Thereby, the expectation values $\langle\lambda_i|\sigma_j^z|\lambda_i\rangle$ with $j = 1, \ldots, r-1$ and $i = r+1, \ldots, d$ vanish. The second term at the right-hand side contains the difference $\tilde{S}_i^z - S_i^z$, which is of the order of magnitude of the square root of the variance. Replacing p_n^z by λ_n in Eq. (11), this becomes

$$\mathrm{Var}(\tilde{S}_i^z) = \frac{1}{N_i}\left[\frac{2}{i(i+1)} - \left(S_i^z\right)^2\right] \quad (29)$$

for $i \geqslant r$. In this particular case, we have that

$$(S_i^z)^2 = \frac{2}{i(i+1)}\left(\sum_{j=1}^{i}\lambda_j - i\lambda_{i+1}\right)^2, \quad (30)$$

or equivalently

$$\left(S_i^z\right)^2 = \frac{2}{i(i+1)}, \quad (31)$$

and thus the variance cancels. Consequently, the first-order term entering in the Taylor series of the infidelity vanishes.

Standard tomography performed on the eigenbasis of ρ eliminates the first-order term of the Taylor series expansion of the infidelity. Consequently, the infidelity scales as $O(1/N)$ for all quantum states. However, since the quantum state ρ is unknown, this tomographic method cannot be carried out. Nevertheless, this result suggests a workaround. In a first stage, standard quantum tomography is applied on an ensemble of size N_0. This provides an estimate $\tilde{\rho}_0$ that can be diagonalized to obtain its eigenbasis. In a second stage, the generators of SU(d) are adapted to the eigenbasis of $\tilde{\rho}_0$, that is, represented in the eigenbasis of $\tilde{\rho}_0$, and employed to perform standard quantum tomography on an ensemble of size $N - N_0$. Thereafter, a new estimate $\tilde{\rho}$ is obtained by postprocessing all experimentally acquired data via Maximum likelihood estimation [28]. This is the two-stage adaptive quantum tomographic method for a single qudit.

Let us now analyze the scaling of the infidelity as function of N_0 and N provided by this method. In the two-stage adaptive tomography, the first-order term of the Taylor series for the infidelity, Eq. (25), does not vanish. This is because the eigenvectors of $\tilde{\rho}_0$ are estimates of the eigenvectors of ρ. The eigenvectors $\{|\tilde{\lambda}_i^0\rangle\}$ of $\tilde{\rho}_0$ can be represented as a rotation of those of ρ, that is,

$$\left|\tilde{\lambda}_i^0\right\rangle = e^{-i\Omega_0}|\lambda_i\rangle, \quad (32)$$

where Ω_0 is a Hermitian matrix. The error δ_i in the estimation of the eigenvalue λ_i is given by $\tilde{\lambda}_i^0 = \lambda_i + \delta_i$. Then, the spectral decomposition of the estimate $\tilde{\rho}_0$,

$$\tilde{\rho}_0 = \sum_{i=1}^{d} \tilde{\lambda}_i^0 \left|\tilde{\lambda}_i^0\right\rangle\left\langle\tilde{\lambda}_i^0\right|, \quad (33)$$

becomes

$$\tilde{\rho}_0 = \rho + \sum_{i=1}^{d}(\delta_i|\lambda_i\rangle\langle\lambda_i| + i\lambda_i[\Omega_0, |\lambda_i\rangle\langle\lambda_i|]) + \sum_{i=1}^{d} i\delta_i[\Omega_0, |\lambda_i\rangle\langle\lambda_i|] + O(||\Omega_0||^2), \quad (34)$$

where the Baker-Campbell-Hausdorff formula has been applied to the unitary transformation entering in Eq. (32). Since $||\Delta_0|| = ||\tilde{\rho}_0 - \rho||$ scales in the worst case as $O(1/\sqrt{N_0})$, we have that $\delta_i \sim O(1/\sqrt{N_0})$ or $||\Omega_0|| \sim O(1/\sqrt{N_0})$. Since only the eigenvectors will intervene in the protocol, we have that

$$||\Omega_0|| \sim O(1/\sqrt{N_0}). \quad (35)$$

Thereby, the inner product between the eigenvectors of ρ and their estimates obtained from $\tilde{\rho}_0$ is

$$\left|\left\langle\tilde{\lambda}_i^0|\lambda_j\right\rangle\right|^2 = \delta_{ij}\left(1 - \langle\lambda_i|\Omega_0^2|\lambda_i\rangle\right) + |\langle\lambda_j|\Omega_0|\lambda_i\rangle|^2 + O(||\Omega_0||^3). \quad (36)$$

Then, the probability $p_i^0 = \langle\tilde{\lambda}_i^0|\rho|\tilde{\lambda}_i^0\rangle$ of projecting the unknown state ρ onto the eigenstate $|\tilde{\lambda}_i^0\rangle$ of $\tilde{\rho}_0$ is

given by

$$p_i^0 = \lambda_i - \tfrac{1}{2}\langle\lambda_i|[\Omega_0,[\Omega_0,\rho]]|\lambda_j\rangle + O(||\Omega_0||^3). \quad (37)$$

Therefore, in the second stage of the tomographic method, both terms entering at the right-hand side of Eq. (28) do not vanish. In this step, the variance is of the order of $O[1/(N-N_0)]$ since the measurements are performed on a ensemble of size $N-N_0$. The expected values of the generators $\{\sigma_j^z\}$, with $j=1,\ldots,r-1$ onto states $\{|\lambda_i\rangle\}$, with $i=r+1,\ldots,d$, do not vanish; instead

$$\langle\tilde{\lambda}_i^0|\sigma_j^z|\tilde{\lambda}_i^0\rangle = \frac{\sqrt{2}}{\sqrt{j(j+1)}}\sum_{k=1}^{j}\left|\langle\tilde{\lambda}_i^0|\lambda_k\rangle\right|^2 - \frac{j\sqrt{2}}{\sqrt{j(j+1)}}\left|\langle\tilde{\lambda}_i^0|\lambda_{j+1}\rangle\right|^2. \quad (38)$$

This expression is a linear combination of the inner products in Eq. (36) and therefore we have that

$$\langle\tilde{\lambda}_i^0|\sigma_j^z|\tilde{\lambda}_i^0\rangle \sim O(1/N_0). \quad (39)$$

Thereby, the first term at the right-hand side of Eq. (28) scales as $O(1/N_0\sqrt{N-N_0})$, since this term is the multiplication of Eq. (38) and $\tilde{S}_i^z - S_i^z \sim O(1/\sqrt{N-N_0})$. The parameters S_i^z of ρ, with $i=r+1,\ldots,d$, on the basis of eigenstates of $\tilde{\rho}_0$ are

$$\left(S_i^z\right)^2 = \frac{2}{i(i+1)}\left(\sum_{j=1}^{i} p_j^0 - i p_{i+1}^0\right)^2, \quad (40)$$

which employing Eq. (37) become

$$\left(S_i^z\right)^2 = \frac{2}{i(i+1)} + O(1/N_0). \quad (41)$$

Replacing this result in Eq. (29), it can be seen that the variance $\mathrm{Var}(\tilde{S}_i^z)$ scale as $O[1/(N-N_0)N_0]$, for $i=r+1,\ldots,d$. Thereby, the second term of Eq. (28) is of the order of $O[1/\sqrt{N_0(N-N_0)}]$. Therefore, the first-order term of the Taylor series expansion of the infidelity scales as

$$\sum_{i=r+1}^{d}\langle\lambda_i|\Delta|\lambda_i\rangle \sim O[1/\sqrt{N_0(N-N_0)}], \quad (42)$$

which is similar to the scaling of the second-order term in the Taylor series. Thereby, the two-stage adaptive quantum tomographic method attains an infidelity that scales as $O[1/\sqrt{(N-N_0)N_0}]$ for any quantum state. In spite of the increase in accuracy, the estimates provided by two-stage adaptive quantum tomography might not be physically acceptable. Thus, the method must be complemented with a postprocessing stage by means of an statistical inference method, such as, for instance, maximum likelihood estimation [28].

The choice of N_0 plays a key role in the performance of the two-stage adaptive quantum tomography. In the case of a qubit, it has been proposed to choose the preliminary ensemble as a fraction [37] or a power [36] of the total sample size N, that is, $N_0 = aN$ or $N_0 = N^b$ with $b \geqslant 2/3$. The first choice leads to an infidelity that scales as $O[1/\sqrt{a(1-a)N}]$, while with the second choice the infidelity scales as $O(1/N^{(1+b)/2})$.

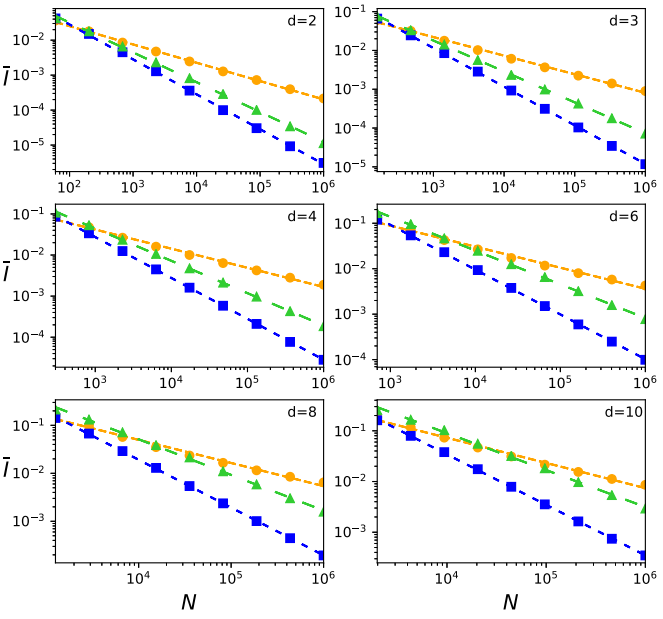

FIG. 1. Mean infidelity $\bar{I}$ as a function of ensemble size N for two-stage adaptive tomography with $N_0 = N/2$ (blue squares), $N_0 = N^{2/3}$ (green triangles), and standard quantum tomography (orange circles), for several dimensions d. Average considers uniformly distributed pure states. For $d=2,3$ and $d=4$ average infidelity was obtained from a sample of 5×10^3 states. For $d=6,8$ and $d=10$ average infidelity was obtained from a sample of 10^3 states.

In the next section, we study the impact of the choices $N_0 = aN$ and $N_0 = N^b$ in the performance of the method by means of numerical simulations.

V. NUMERICAL SIMULATIONS

In order to study the properties of the two-stage adaptive quantum tomography, we have carried out numerical simulations. In a given dimension d, we have uniformly generated (Haar measure) quantum states. Estimates of these states are obtained by means of the two-stage adaptive quantum tomography considering a multinomial distribution for the measurement results and an ensemble size N. Thereafter, we calculate the average $\bar{I}$ of the infidelity Eq. (16) as a function of d and N.

Figure 1 shows log-log graphs for the mean infidelity $\bar{I}$ obtained in the determination of pure states for dimensions $d=2,3,4,6,8$ and $d=10$ as a function of N. Squares and triangles indicate the mean infidelity obtained by means of two-stage adaptive quantum tomography with $N_0 = N/2$ and $N_0 = N^{2/3}$, respectively. Circles indicate the mean infidelity obtained by standard quantum tomography. Dashed lines correspond to the best linear fit. Figure 1 shows that the slope of the mean infidelity $\bar{I}$ provided by two-stage adaptive tomography with $N_0 = N/2$ is approximately twice the slope of the mean infidelity achieved by standard quantum tomography. This indicates that two-stage adaptive quantum tomography provides a quadratic reduction in the mean infidelity when compared to standard quantum tomography. This result is in agreement with our previous estimate for the scaling of the infidelity for two-stage adaptive quantum tomography. Thus, two-stage adaptive

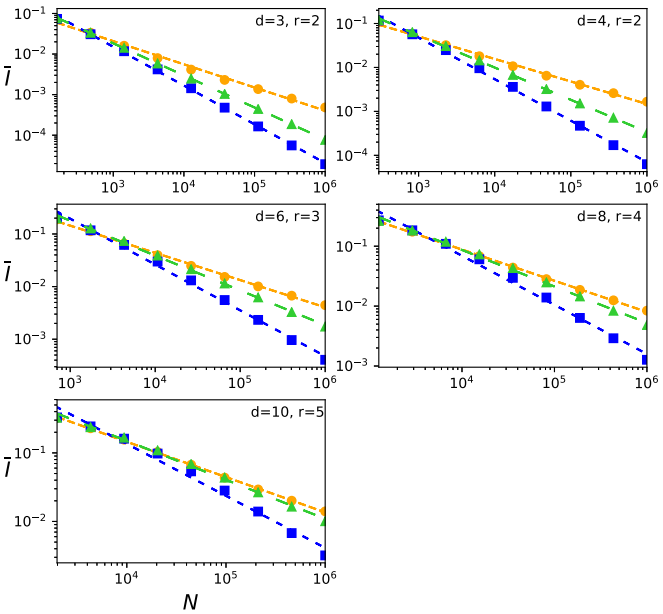

FIG. 2. Mean infidelity $\bar{I}$ as a function of ensemble size N for two-stage adaptive tomography with $N_0 = N/2$ (blue squares), $N_0 = N^{2/3}$ (green triangles), and standard quantum tomography (orange circles), for several dimensions d. Average considers uniformly distributed states of rank $r = [d/2]$. For $d = 2, 3$ and $d = 4$ average infidelity was obtained from a sample of 5×10^3 states. For $d = 6, 8$ and $d = 10$ average infidelity was obtained from a sample of 10^3 states.

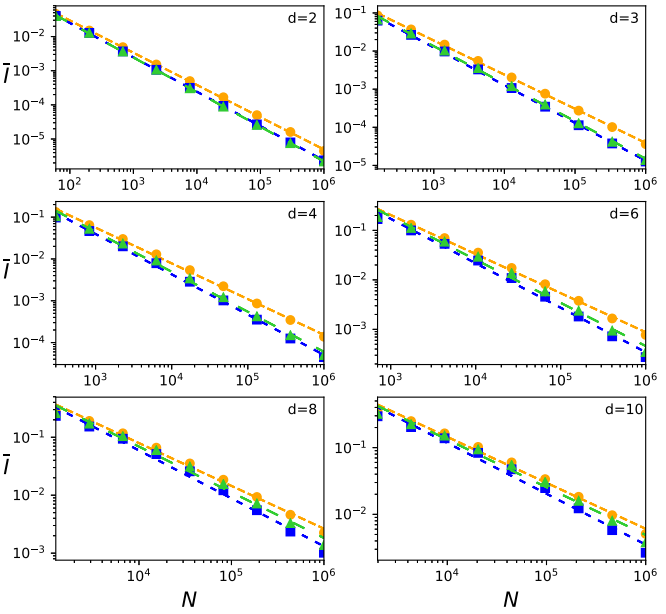

FIG. 3. Mean infidelity $\bar{I}$ as a function of ensemble size N for two-stage adaptive tomography with $N_0 = N/2$ (blue squares), $N_0 = N^{2/3}$ (green triangles), and standard quantum tomography (orange circles), for several dimensions d. Average considers uniformly distributed full-rank states. For $d = 2, 3$ and $d = 4$ average infidelity was obtained from a sample of 5×10^3 states. For $d = 6, 8$ and $d = 10$ average infidelity was obtained from a sample of 10^3 states.

tomography with $N_0 = N/2$ achieves a higher accuracy in the determination of rank-1 quantum states when compared to standard tomography. The case $N_0 = N^{2/3}$ also leads to an improvement with respect to standard tomography. However, for this particular choice the improvement becomes marginal as the dimension increases in the inspected interval of ensamble size N.

Figure 2 illustrates the behavior of the average infidelity $\bar{I}$ as a function of N in the case of quantum states with rank $r = [d/2]$. Two-stage adaptive tomography with $N_0 = N/2$ provides the best accuracy, as in the case of pure states, albeit the gain in accuracy is lower than the obtained in the case of pure states. Furthermore, as the dimension d increases the gain in accuracy diminishes and the three tomographic methods tend to provide a similar accuracy.

Figure 3 displays the behavior of the mean infidelity for the case of full-rank states. Here, the dominant feature is that standard tomography and two-stage adaptive tomography provide a similar accuracy for all dimensions and ensemble sizes. In particular, the mean infidelity scales approximately as $O(1/N)$ for both tomographic methods. The choices $N_0 = N^{2/3}$ and $N_0 = N/2$ seems to perform marginally better than standard tomography in the inspected interval of ensemble sizes. Nevertheless, the linear fits indicate that for larger ensemble sizes the choice $N_0 = N/2$ might provide a non-negligible improvement over standard tomography.

Finally, in Fig. 4 we have studied the case of full-rank, highly pure states of the form

$$\rho_d = \lambda|\psi\rangle\langle\psi| + \frac{1-\lambda}{d}\mathbb{I} \tag{43}$$

for $\lambda = 0.99$. This class of states is important due to the fact that it is usually employed to model errors in certain experiments. As Fig. 4 illustrates, for lower dimensions the choices $N_0 = N/2$ and $N_0 = N^{2/3}$ achieve a similar accuracy, which is higher than the accuracy achieved by standard quantum tomography. As the dimension increases and several eigenvalues decrease in magnitude, two-stage adaptive tomography with $N_0 = N/2$ leads to better results than standard tomography. However, as the dimension increases, the improvement becomes marginal within the interval of inspected sample sizes.

We have also estimated the variance of the mean infidelity. This is characterized by a very small value, which makes difficult its depiction. Instead, we have also studied median infidelity and interquartile range (see Appendix B). Mean and median have a similar behavior while interquartile range is of the same order of magnitude as mean or median infidelity. With the exception of dimension $d = 2$ and full-rank states in all dimensions, the choice $N_0 = N/2$ achieves a median infidelity whose interquartile range does not overlap with the interquartile range of the median infidelity provided by standard quantum tomography. Thus, 75% of all reconstructed states are much better estimated by two-stage standard quantum tomography than standard quantum tomography.

Thus, numerical simulations indicate that the choice $N_0 = N/2$ provides the best results in two-stage adaptive tomography. Furthermore, this tomographic methods provides a quadratic improvement in the accuracy when compared to standard tomography. This result seems to hold for states of rank smaller than or equal to $[d/2]$. Thereby, in the absence of *a priori* information about the rank, the choice $N_0 = N/2$

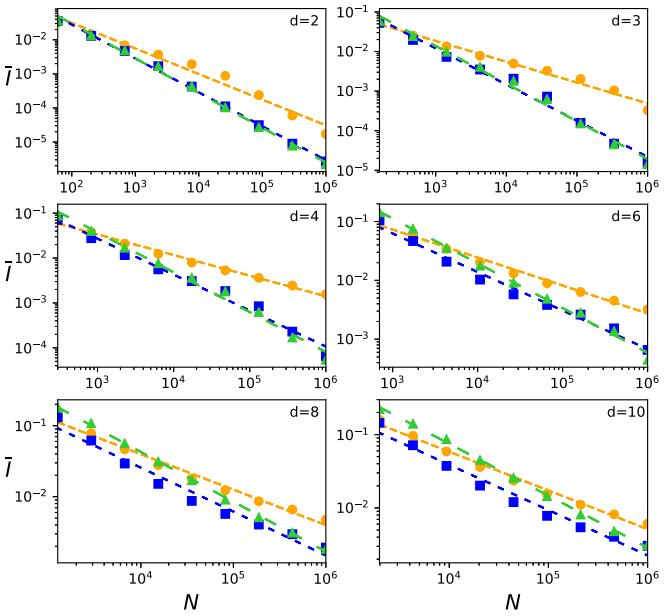

FIG. 4. Mean infidelity $\bar{I}$ as a function of ensemble size N for two-stage adaptive tomography with $N_0 = N/2$ (blue squares), $N_0 = N^{2/3}$ (green triangles), and standard quantum tomography (orange circles), for several dimensions d. Average considers pure states uniformly distributed affected by white noise Eq. (43) with $\lambda = 0.99$. For $d = 2, 3$ and $d = 4$ average infidelity was obtained from a sample of 5×10^3 states. For $d = 6, 8$ and $d = 10$ average infidelity was obtained from a sample of 10^3 states.

should be employed in the determination of the states of a single qudit by means of two-stage adaptive tomography.

VI. CONCLUSIONS

In this article, we have extended the validity of two-stage quantum tomography to the case of a single qudit. In a first stage, standard quantum tomography is carried out on a ensemble of size N_0. After linear inversion, an estimate $\tilde{\rho}_0$ is obtained and diagonalized. The eigenbasis of $\tilde{\rho}_0$ is then employed to represent the generators of SU(d), which are then subsequently used in a second stage of standard quantum tomography on an ensemble of size $N - N_0$. We have shown that the estimate $\tilde{\rho}$ provided by this latter stage leads to an infidelity that scales as $O[1/\sqrt{N_0(N - N_0)}]$ for all quantum states. Numerical simulations indicate that the choice $N_0 = N/2$ attains an accuracy of the order of $O(1/N)$ that is higher than the accuracy achieved by standard quantum tomography. Largest performance gain is obtained for low-rank states, where a quadratic improvement of the infidelity is observed. According to the lineal fits for the mean infidelity, for a large enough ensemble size N two-stage adaptive tomography leads to a higher accuracy than standard tomography. Nevertheless, our numerical experiments indicate that for a small ensemble size N two-stage adaptive tomography and standard quantum tomography exhibit a similar performance. Here a word of caution is necessary. Our calculations are based on the Taylor series expansion of the infidelity, which requires a large ensemble size. For a small ensemble size, a different approach is required.

The choice $N_0 = N/2$ seems to perform well in the numerical simulations. It is, however, not necessarily optimal. It has been shown that in the case of a single qubit, the choice $N_0 = N/3$ allows two-stage quantum tomography to saturate the Gill-Massar bound for several figures of merit for the accuracy, such as, for instance, mean square error, mean square Bures distance, and weighted mean square error. Therefore, it is plausible that the performance gain attained by two-stage adaptive quantum tomography can be further increased by optimizing N_0 for every dimension.

The dimension of quantum systems can be conceived as a valuable resource for information processing since higher dimensional systems exhibit several advantages in protocols of quantum computation and quantum communications. This drives the need to design accurate tomographic methods for higher dimensional quantum systems. In this scenario, our results find direct application. In recent years, higher dimensional systems have been experimentally generated, for instance, 1024-dimensional photonic states [51], 16-dimensional photonic states [52], entangled photons with high angular momenta [53], eight-photon entanglement [54], and 14-qubit entanglement [55]. In the case of photonic systems, higher dimensional quantum systems are generated by means of single photons manipulated with the help of optical diffractive elements such as physical slits [56] or spatial light modulators (SLM). A sequence of SLMs allows us to implement projective measurements on a d-dimensional photonic qudit [18,57], which can be electronically controlled in real time. This makes feasible an experimental implementation of two-stage adaptive quantum tomography in higher dimensions.

We have employed the infidelity as a measure of the accuracy of the tomographic scheme. It is possible, however, to extend our analysis to other figures of merits, for instance, the family of functionals defined by the Weighted mean square error [39].

Finally, we would like to comment on the extension of two-stage adaptive quantum tomography to the case of multiqudit systems. Standard tomography can be extended to the case of multipartite systems considering a decomposition on tensor products of local SU(d) generators [11], that is,

$$\rho = \frac{1}{2^n} \sum_{i_0,\ldots,i_n=0}^{d^2-1} S_{i_1\cdots i_n} \sigma_{i_1} \otimes \cdots \otimes \sigma_{i_n}, \tag{44}$$

where $\sigma_0 = 2\mathbb{I}/d$ and $\{S_{i_1\ldots i_n}\}$ are the averages of the local observables $\{\sigma_{i_1} \otimes \cdots \otimes \sigma_{i_n}\}$. Therefore, by performing measurements on these d^{2n} observables, the state ρ can be reconstructed. The main characteristic and appeal of this tomographic scheme is that it only requires local measurements on each subsystem. Analogously, the two-stage adaptive tomography could be extended to multipartite systems. First, an estimate is obtained from a fraction of the total ensemble by local standard multipartite tomography following the previous equation. Thereafter, a second standard multipartite tomography in the remaining ensemble is carried out. Here, the measurement operators are adapted to the base of eigenstates of the preliminary estimate. However, this simple generalization leads to a significant drawback: The tomographic scheme is no longer local. Since the state to be reconstructed might be entangled, the method would require measuring on bases

containing entangled states. This is precisely what standard quantum tomography avoids. Clearly, a generalization of standard multipartite quantum tomography to two-stage adaptive multipartite standard quantum tomography requires further consideration [58].

ACKNOWLEDGMENTS

This work was funded by the Millennium Institute for Research in Optics (MIRO) and CONICYT FONDECYT 1140635 and 1180558. L.P. was supported by CONICYT-PCHA/MagísterNacional/2016-22161371. L.Z. was supported by CONICYT-PCHA/MagísterNacional/2016-22161286.

APPENDIX A: TAYLOR SERIES EXPANSION OF INFIDELITY

In this section, we calculate the second-order approximation of the infidelity by resorting to the Fréchet derivative. The infidelity between two quantum states ρ and $\tilde{\rho}$ is defined by the expression

$$I(\rho,\tilde{\rho}) = 1 - \mathrm{Tr}(\sqrt{\sqrt{\rho}\tilde{\rho}\sqrt{\rho}})^2. \tag{A1}$$

The main advantage of resorting to the Fréchet derivative is that this does not require a particular parametrization of the space of density matrices.

Let us consider that the density matrix ρ has rank $r \leqslant d$ and that the state $\tilde{\rho}$ is an infinitesimal perturbation of ρ, that is,

$$\tilde{\rho} = \rho + \Delta, \tag{A2}$$

where Δ is a Hermitian and traceless matrix such that $||\Delta|| = \sqrt{\mathrm{Tr}(\Delta^2)} \ll 1$. The operator Δ describes the error in a tomographic process. In the case of standard quantum tomography, we have that the error operator is given by

$$\Delta = \tilde{\rho} - \rho \tag{A3}$$

$$= \frac{1}{d}\mathbb{I} + \frac{1}{2}\sum_{i=1}^{d^2-1}\tilde{S}_i\sigma_i - \left(\frac{1}{d}\mathbb{I} + \frac{1}{2}\sum_{i=1}^{d^2-1}S_i\sigma_i\right) \tag{A4}$$

$$= \frac{1}{2}\sum_{i=1}^{d^2-1}(\tilde{S}_i - S_i)\sigma_i. \tag{A5}$$

The norm of the error operator becomes

$$||\Delta|| = \sqrt{\mathrm{Tr}(\Delta^2)} \tag{A6}$$

$$= \left[\frac{1}{4}\sum_{i,j=1}^{d^2-1}(\tilde{S}_i - S_i)(\tilde{S}_j - S_j)\mathrm{Tr}\big(\sigma_i\sigma_j\big)\right]^{1/2} \tag{A7}$$

$$= \left[\frac{1}{2}\sum_{i=1}^{d^2-1}(\tilde{S}_i - S_i)^2\right]^{1/2} \tag{A8}$$

$$\sim O(1/\sqrt{N}). \tag{A9}$$

Without loss of generality, we can write ρ as a direct sum between the subspace with rank r and a $d-r$ null matrix Θ_{d-r}, that is,

$$\rho = \rho_r \oplus \Theta_{d-r} = \begin{bmatrix}\rho_r & \Theta_{r,d-r}\\ \Theta_{d-r,r} & \Theta_{d-r}\end{bmatrix}. \tag{A10}$$

Similarly,

$$\Delta = \begin{bmatrix}\Delta_r & \Delta_{r,d-r}\\ \Delta_{d-r,r} & \Delta_{d-r}\end{bmatrix}, \tag{A11}$$

where Δ_{d-r} is positive semidefinite and $\mathrm{Tr}(\Delta_r) = -\mathrm{Tr}(\Delta_{d-r})$. Thus,

$$\sqrt{\rho}\Delta\sqrt{\rho} = \begin{bmatrix}\sqrt{\rho_r} & \Theta_{r,d-r}\\ \Theta_{d-r,r} & \Theta_{d-r}\end{bmatrix}\begin{bmatrix}\Delta_r & \Delta_{r,d-r}\\ \Delta_{d-r,r} & \Delta_{d-r}\end{bmatrix} \times\begin{bmatrix}\sqrt{\rho_r} & \Theta_{r,d-r}\\ \Theta_{d-r,r} & \Theta_{d-r}\end{bmatrix} \tag{A12}$$

$$= \begin{bmatrix}\sqrt{\rho_r}\Delta_r\sqrt{\rho_r} & \Theta_{r,d-r}\\ \Theta_{d-r,r} & \Theta_{d-r}\end{bmatrix} \tag{A13}$$

$$= \sqrt{\rho_r}\Delta_r\sqrt{\rho_r} \oplus \Theta_{d-r}. \tag{A14}$$

Thereby, the infidelity becomes

$$1 - I(\rho,\tilde{\rho}) = \mathrm{Tr}(\sqrt{\rho^2 + \sqrt{\rho}\Delta\sqrt{\rho}})^2 \tag{A15}$$

$$= \mathrm{Tr}\Big(\sqrt{\rho_r^2 \oplus \Theta_{d-r} + \sqrt{\rho_r}\Delta_r\sqrt{\rho_r} \oplus \Theta_{d-r}}\Big)^2 \tag{A16}$$

$$= \mathrm{Tr}\Big(\sqrt{\rho_r^2 + \sqrt{\rho_r}\Delta_r\sqrt{\rho_r}}\Big)^2. \tag{A17}$$

The infidelity between ρ and $\rho+\Delta$ only depends on the non-null eigenspace of ρ. Let us note that $\rho_r + \Delta_r$ is not a quantum state, but this will not influence our arguments. Defining $f(A) := \sqrt{A}$, we can write the infidelity as

$$I(\rho,\tilde{\rho}) = 1 - \mathrm{Tr}\big[f\big(\rho_r^2 + \sqrt{\rho_r}\Delta_r\sqrt{\rho_r}\big)\big]^2. \tag{A18}$$

Now, we can approximate $F(\rho,\rho+\Delta)$ by performing the Taylor expansion of $f(A)$ around B,

$$f(A+B) = f(A) + \sum_{m=1}^{\infty}\frac{1}{m!}D^m f(A)([B]^m), \tag{A19}$$

where $D^m f(A)([B]^m)$ is the mth directional Fréchet derivative of $f(A)$ along B, which can be calculated as

$$Df(U)(V) = \frac{d}{dt}f(U+tV)\bigg|_{t=0}. \tag{A20}$$

In our case, $A = \rho_r^2$ and $B = \sqrt{\rho_r}\Delta_r\sqrt{\rho_r}$. It should be noted that this expansion does not depend on any parametrization of density matrices. The first and second derivatives of f can be obtained from the product rule

$$Df^2(A)(B) = D[ff](A)(B) \tag{A21}$$

$$= Df(A)(B)f(A) + f(A)Df(A)(B). \tag{A22}$$

Since $Df^2(A)(B) = B$, we obtain that

$$Df(A)(B)\sqrt{A} + \sqrt{A}Df(A)(B) = B. \tag{A23}$$

Let us now consider $\{|a_i\rangle\}$ the basis of eigenvectors of A and $\{a_i\}$ the corresponding eigenvalues of A, with $i = 1, ..., r$, where r is the dimension of A, which in this case is equal

to rank of ρ. Expanding on this basis, we have

$$\sum_{j=1}^{r}([Df(A)(B)]_{ij}\sqrt{a_j}\delta_{jk}+\sqrt{a_i}\delta_{ij}[Df(A)(B)]_{jk})=B_{ik}, \tag{A24}$$

$$\sqrt{a_k}[Df(A)(B)]_{ik}+\sqrt{a_i}[Df(A)(B)]_{ik}=B_{ik}, \tag{A25}$$

$$(\sqrt{a_i}+\sqrt{a_k})[Df(A)(B)]_{ik}=B_{ik}. \tag{A26}$$

Then,

$$[Df(A)(B)]_{ik}=\frac{B_{ik}}{\sqrt{a_i}+\sqrt{a_k}}. \tag{A27}$$

Since $A=\rho_r^2$ is full rank, there is no indetermination in the derivative. For the second derivative, we use the product rule again,

$$\begin{aligned}
&D^2f^2(A)(B)(C)\\
&=D\big[Df^2(A)(B)\big](C) &\text{(A28)}\\
&=D[Df(A)(B)f(A)+f(A)Df(A)(B)](C) &\text{(A29)}\\
&=D^2f(A)(B)(C)f(A)+Df(A)(B)Df(A)(C)\\
&\quad+Df(A)(C)Df(A)(B)+f(A)D^2f(A)(B)(C) &\text{(A30)}\\
&=D^2f(A)(B)(C)\sqrt{A}+Df(A)(B)Df(A)(C)\\
&\quad+Df(A)(C)Df(A)(B)+\sqrt{A}D^2f(A)(B)(C). &\text{(A31)}
\end{aligned}$$

Using $D^2f^2(A)(B)(C)=0$ and writing per component,

$$\begin{aligned}
0&=\sum_{j=1}^{r}\Bigg([D^2f(A)(B)(C)]_{ij}\sqrt{a_j}\delta_{jk}\\
&\quad+\frac{B_{ij}}{\sqrt{a_i}+\sqrt{a_j}}\frac{C_{jk}}{\sqrt{a_j}+\sqrt{a_k}}\\
&\quad+\frac{C_{ij}}{\sqrt{a_i}+\sqrt{a_j}}\frac{B_{jk}}{\sqrt{a_j}+\sqrt{a_k}}\\
&\quad+\sqrt{a_i}\delta_{ij}[D^2f(A)(B)(C)]_{jk}\Bigg) &\text{(A32)}\\
&=\sqrt{a_k}[D^2f(A)(B)(C)]_{ik}+\sqrt{a_i}[D^2f(A)(B)(C)]_{ik}\\
&\quad+\sum_{j=1}^{r}\Bigg(\frac{B_{ij}}{\sqrt{a_i}+\sqrt{a_j}}\frac{C_{jk}}{\sqrt{a_j}+\sqrt{a_k}}\\
&\quad+\frac{C_{ij}}{\sqrt{a_i}+\sqrt{a_j}}\frac{B_{jk}}{\sqrt{a_j}+\sqrt{a_k}}\Bigg) &\text{(A33)}\\
&=(\sqrt{a_i}+\sqrt{a_k})[D^2f(A)(B)(C)]_{ik}\\
&\quad+\sum_{j=1}^{r}\Bigg(\frac{B_{ij}}{\sqrt{a_i}+\sqrt{a_j}}\frac{C_{jk}}{\sqrt{a_j}+\sqrt{a_k}}\\
&\quad+\frac{C_{ij}}{\sqrt{a_i}+\sqrt{a_j}}\frac{B_{jk}}{\sqrt{a_j}+\sqrt{a_k}}\Bigg). &\text{(A34)}
\end{aligned}$$

Then, the second derivative is

$$\begin{aligned}
&[D^2f(A)(B)(C)]_{ik}\\
&=-\frac{1}{\sqrt{a_i}+\sqrt{a_k}}\sum_{j=1}^{r}\Bigg(\frac{B_{ij}}{\sqrt{a_i}+\sqrt{a_j}}\frac{C_{jk}}{\sqrt{a_j}+\sqrt{a_k}}\\
&\quad+\frac{C_{ij}}{\sqrt{a_i}+\sqrt{a_j}}\frac{B_{jk}}{\sqrt{a_j}+\sqrt{a_k}}\Bigg). &\text{(A35)}
\end{aligned}$$

From $A=\rho_r^2$ and $B=\sqrt{\rho_r}\Delta_r\sqrt{\rho_r}$, we get

$$a_i=\lambda_i^2,\quad B_{ij}=\sqrt{\lambda_i\lambda_j}\langle\lambda_i^r|\Delta_r|\lambda_j^r\rangle, \tag{A36}$$

where $\{\lambda_i\}$ are eigenvalues of ρ_r with eigenvectors $\{|\lambda_j^r\rangle\}$, with $i=1,\ldots,r$. Thereby, the first and second derivatives are

$$\big[Df\big(\rho_r^2\big)(\sqrt{\rho_r}\Delta_r\sqrt{\rho_r})\big]_{ij}=\frac{\sqrt{\lambda_i\lambda_j}\langle\lambda_i^r|\Delta_r|\lambda_j^r\rangle}{\lambda_i+\lambda_j} \tag{A37}$$

and

$$\begin{aligned}
&\big[D^2f\big(\rho_r^2\big)([\sqrt{\rho_r}\Delta_r\sqrt{\rho_r}]^2)\big]_{ij}\\
&=-\frac{2\sqrt{\lambda_i\lambda_j}}{\lambda_i+\lambda_j}\sum_{k=1}^{r}\frac{\lambda_k\langle\lambda_i^r|\Delta_r|\lambda_k^r\rangle\langle\lambda_k^r|\Delta_r|\lambda_j^r\rangle}{(\lambda_i+\lambda_k)(\lambda_k+\lambda_j)}.
\end{aligned} \tag{A38}$$

Thereby, the function f is up to second order given by

$$\begin{aligned}
&\big[f\big(\rho_r^2+\sqrt{\rho_r}\Delta_r\sqrt{\rho_r}\big)\big]_{ij}\\
&=\big[f\big(\rho_r^2\big)\big]_{ij}+\big[Df\big(\rho_r^2\big)(\sqrt{\rho_r}\Delta_r\sqrt{\rho_r})\big]_{ij}\\
&\quad+\frac{1}{2}\big[D^2f\big(\rho_r^2\big)([\sqrt{\rho_r}\Delta_r\sqrt{\rho_r}]^2)\big]_{ij}+O(||\Delta_r||^3) &\text{(A39)}\\
&=\lambda_i\delta_{ij}+\frac{\sqrt{\lambda_i\lambda_j}\langle\lambda_i^r|\Delta_r|\lambda_j^r\rangle}{\lambda_i+\lambda_j}\\
&\quad-\frac{\sqrt{\lambda_i\lambda_j}}{\lambda_i+\lambda_j}\sum_{k=1}^{r}\frac{\lambda_k\langle\lambda_i^r|\Delta_r|\lambda_k^r\rangle\langle\lambda_k^r|\Delta_r|\lambda_j^r\rangle}{(\lambda_i+\lambda_k)(\lambda_k+\lambda_j)}+O(||\Delta_r||^3). &\text{(A40)}
\end{aligned}$$

Taking the trace,

$$\begin{aligned}
&\mathrm{Tr}\big[f\big(\rho_r^2+\sqrt{\rho_r}\Delta_r\sqrt{\rho_r}\big)\big]\\
&=1+\frac{1}{2}\sum_{i=1}^{r}\langle\lambda_i^r|\Delta_r|\lambda_i^r\rangle\\
&\quad-\frac{1}{2}\sum_{i,k=1}^{r}\frac{\lambda_k|\langle\lambda_i^r|\Delta_r|\lambda_k^r\rangle|^2}{(\lambda_i+\lambda_k)^2}+O(||\Delta_r||^3).
\end{aligned} \tag{A41}$$

The last equation can be rewritten in term of the eigenvectors $\{|\lambda_i\rangle\}$ of ρ, where the states with index $i=1,\ldots,r$ are the non-null eigenvectors of ρ and the states with $i=r+1,\ldots,d$ are orthonormal states acting onto the kernel of ρ. Let us note that $\langle\lambda_i^r|\Delta_r|\lambda_k^r\rangle=\langle\lambda_i|\Delta|\lambda_k\rangle$. Then, the first-order term

becomes

$$\sum_{i=1}^{r}\langle\lambda_i^r|\Delta_r|\lambda_i^r\rangle = \sum_{i=1}^{r}\langle\lambda_i|\Delta|\lambda_i\rangle \tag{A42}$$

$$= \mathrm{Tr}(\Delta) - \sum_{i=r+1}^{d}\langle\lambda_i|\Delta|\lambda_i\rangle \tag{A43}$$

$$= -\sum_{i=r+1}^{d}\langle\lambda_i|\Delta|\lambda_i\rangle. \tag{A44}$$

Since Δ is Hermitian, the second-order term can be rewritten as

$$\sum_{i,k=1}^{r}\frac{\lambda_k}{(\lambda_i+\lambda_k)^2}\left|\langle\lambda_i^r|\Delta_r|\lambda_k^r\rangle\right|^2$$

$$= \sum_{i,k=1}^{r}\frac{\lambda_k|\langle\lambda_i|\Delta|\lambda_k\rangle|^2}{(\lambda_i+\lambda_k)^2} \tag{A45}$$

$$= \frac{1}{2}\sum_{i,k=1}^{r}\left[\frac{\lambda_k|\langle\lambda_i|\Delta|\lambda_k\rangle|^2}{(\lambda_i+\lambda_k)^2} + \frac{\lambda_k|\langle\lambda_i|\Delta|\lambda_k\rangle|^2}{(\lambda_i+\lambda_k)^2}\right] \tag{A46}$$

$$= \frac{1}{2}\sum_{i,k=1}^{r}\left[\frac{\lambda_k|\langle\lambda_i|\Delta|\lambda_k\rangle|^2}{(\lambda_i+\lambda_k)^2} + \frac{\lambda_i|\langle\lambda_i|\Delta|\lambda_k\rangle|^2}{(\lambda_k+\lambda_i)^2}\right] \tag{A47}$$

$$= \frac{1}{2}\sum_{i,k=1}^{r}\frac{\lambda_k+\lambda_i}{(\lambda_i+\lambda_k)^2}|\langle\lambda_i|\Delta|\lambda_k\rangle|^2 \tag{A48}$$

$$= \frac{1}{2}\sum_{i,k=1}^{r}\frac{|\langle\lambda_i|\Delta|\lambda_k\rangle|^2}{\lambda_i+\lambda_k}. \tag{A49}$$

Therefore,

$$\mathrm{Tr}\left[f\left(\rho_r^2+\sqrt{\rho_r}\Delta_r\sqrt{\rho_r}\right)\right]$$

$$= 1 - \frac{1}{2}\sum_{i=r+1}^{d}\langle\lambda_i|\Delta|\lambda_i\rangle$$

$$- \frac{1}{4}\sum_{i,k=1}^{r}\frac{|\langle\lambda_i|\Delta|\lambda_k\rangle|^2}{\lambda_i+\lambda_k} + O(||\Delta||^3). \tag{A50}$$

Thus, the infidelity is up to second order given by the expression

$$I(\rho,\tilde{\rho}) = 1 - \left[1 - \frac{1}{2}\sum_{i=r+1}^{d}\langle\lambda_i|\Delta|\lambda_i\rangle\right.$$

$$\left. - \frac{1}{4}\sum_{i,k=1}^{r}\frac{|\langle\lambda_i|\Delta|\lambda_k\rangle|^2}{\lambda_i+\lambda_k} + O(||\Delta||^3)\right]^2 \tag{A51}$$

$$= \sum_{i=r+1}^{d}\langle\lambda_i|\Delta|\lambda_i\rangle + \frac{1}{2}\sum_{i,k=1}^{r}\frac{|\langle\lambda_i|\Delta|\lambda_k\rangle|^2}{\lambda_i+\lambda_k}$$

$$- \frac{1}{4}\left[\sum_{i=r+1}^{d}\langle\lambda_i|\Delta|\lambda_i\rangle\right]^2 + O(||\Delta||^3). \tag{A52}$$

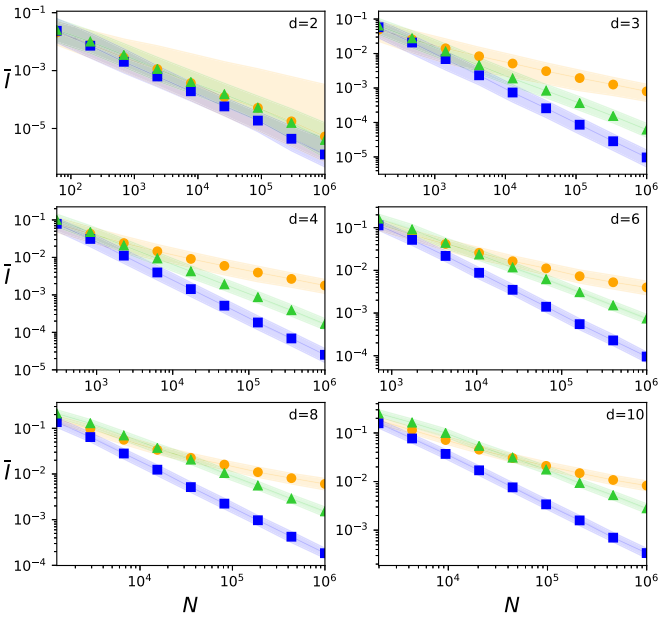

FIG. 5. Median infidelity $\bar{I}$ as a function of ensemble size N for two-stage adaptive tomography with $N_0 = N/2$ (blue squares), $N_0 = N^{2/3}$ (green triangles), and standard quantum tomography (orange circles), for several dimensions d. The median considers uniformly distributed pure states. For $d = 2, 3$ and $d = 4$ median infidelity was obtained from a sample of 5×10^3 states. For $d = 6, 8$ and $d = 10$ median infidelity was obtained from a sample of 10^3 states.

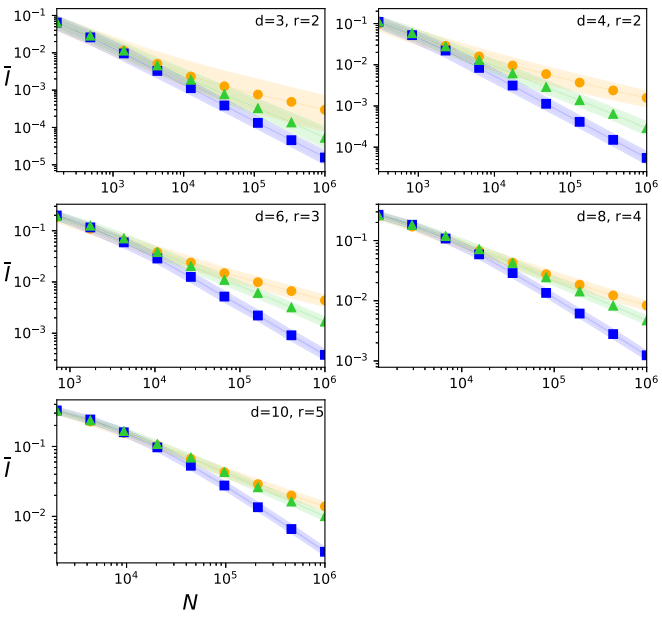

FIG. 6. Median infidelity $\bar{I}$ as a function of ensemble size N for two-stage adaptive tomography with $N_0 = N/2$ (blue squares), $N_0 = N^{2/3}$ (green triangles), and standard quantum tomography (orange circles), for several dimensions d. The median considers uniformly distributed states of rank $r = [d/2]$. For $d = 2, 3$ and $d = 4$ median infidelity was obtained from a sample of 5×10^3 states. For $d = 6, 8$ and $d = 10$ median infidelity was obtained from a sample of 10^3 states.

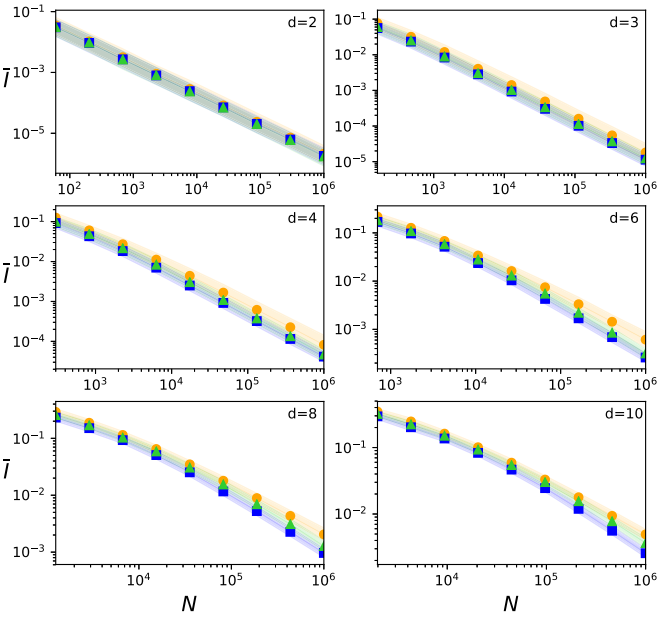

FIG. 7. Median infidelity $\bar{I}$ as a function of ensemble size N for two-stage adaptive tomography with $N_0 = N/2$ (blue squares), $N_0 = N^{2/3}$ (green triangles), and standard quantum tomography (orange circles), for several dimensions d. The median considers uniformly distributed full-rank states. For $d = 2, 3$ and $d = 4$ median infidelity was obtained from a sample of 5×10^3 states. For $d = 6, 8$ and $d = 10$ median infidelity was obtained from a sample of 10^3 states.

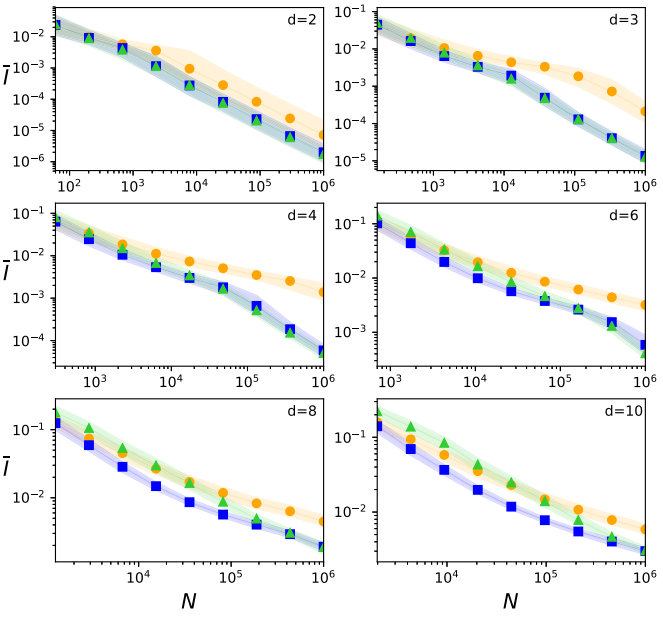

FIG. 8. Median infidelity $\bar{I}$ as a function of ensemble size N for two-stage adaptive tomography with $N_0 = N/2$ (blue squares), $N_0 = N^{2/3}$ (green triangles), and standard quantum tomography (orange circles), for several dimensions d. The median considers pure states uniformly distributed affected by white noise Eq. (43) with $\lambda = 0.99$. For $d = 2, 3$ and $d = 4$ median infidelity was obtained from a sample of 5×10^3 states. For $d = 6, 8$ and $d = 10$ median infidelity was obtained from a sample of 10^3 states.

For a low-rank state, the linear term in the expansion of the infidelity dominates,

$$I(\rho, \tilde{\rho}) = \sum_{i=r+1}^{d} \langle \lambda_i | \Delta | \lambda_i \rangle + O(||\Delta||^2). \tag{A53}$$

Otherwise, for a full-rank density matrix, the first-order term vanishes because ρ does not have a null subspace. In this case, the approximation becomes up to second order

$$I(\rho, \tilde{\rho}) = \frac{1}{2} \sum_{i,k=1}^{d} \frac{|\langle \lambda_i | \Delta | \lambda_k \rangle|^2}{\lambda_i + \lambda_k} + O(||\Delta||^3). \tag{A54}$$

APPENDIX B: MEDIAN AND INTERQUARTILE RANGE

Figures 5–8 show log-log graphics for the median infidelity obtained from simulations. Squares and triangles indicate the median infidelity obtained by means of two-stage adaptive tomography with $N_0 = N/2$ and $N_0 = N^{2/3}$, respectively. Circles indicate the median infidelity obtained by standard quantum tomography. Filled areas are the interquartile range, that is, the difference between 75th and 25th percentiles. The interquartile range is a measure of the dispersion of the data.

Mean and median have a similar behavior while interquartile range is of the same order of magnitude than mean or median infidelity. With the exception of dimension $d = 2$ and full-rank states in all dimensions, the choice $N_0 = N/2$ provides a median infidelity whose interquartile range does not overlap with the interquartile range of the median infidelity provided by standard quantum tomography. Thus, 75% of all reconstructed states are much better estimated by two-stage standard quantum tomography than standard quantum tomography.

[1] Č. Brukner, M. Zukowski, and A. Zeilinger, Phys. Rev. Lett. **89**, 197901 (2002).
[2] G. M. Nikolopoulos, K. S. Ranade, and G. Alber, Phys. Rev. A **73**, 032325 (2006).
[3] N. J. Cerf, M. Bourennane, A. Karlsson, and N. Gisin, Phys. Rev. Lett. **88**, 127902 (2002).
[4] H. S. Tonchev and N. V. Vitanov, Phys. Rev. A **94**, 042307 (2016).
[5] M. Luo, X. Chen, Y. Yang, and X. Wang, Sci. Rep. **4**, 4044 (2014).
[6] D. P. O'Leary, G. K. Brennen, and S. S. Bullock, Phys. Rev. A **74**, 032334 (2006).
[7] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu, Phys. Rev. Lett. **88**, 040404 (2002).
[8] D. Kaszlikowski, P. Gnaciński, M. Zukowski, W. Miklaszewski, and A. Zeilinger, Phys. Rev. Lett. **85**, 4418 (2000).
[9] H. Häffner, W. Hänsel, C. F. Roos, J. Benhelm, D. Chek-al-kar, M. Chwalla, T. Körber, U. D. Rapol, M. Riebe, P. O. Schmidt *et al.*, Nature (London) **438**, 643 (2005).

[10] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, Phys. Rev. A **64**, 052312 (2001).
[11] R. T. Thew, K. Nemoto, A. G. White, and W. J. Munro, Phys. Rev. A **66**, 012303 (2002).
[12] J. Schwinger, Proc. Natl. Acad. Sci. USA **46**, 570 (1960).
[13] I. D. Ivanovic, J. Phys. A **14**, 3241 (1981).
[14] W. K. Wootters and B. D. Fields, Ann. Phys. (NY) **191**, 363 (1989).
[15] A. B. Klimov, C. Muñoz, A. Fernández, and C. Saavedra, Phys. Rev. A **77**, 060303(R) (2008).
[16] S. N. Filippov and V. I. Man'ko, Phys. Scr. **2011**, 14010 (2011).
[17] R. B. A. Adamson and A. M. Steinberg, Phys. Rev. Lett. **105**, 030406 (2010).
[18] G. Lima, L. Neves, R. Guzman, E. S. Gomez, W. A. T. Nogueira, A. Delgado, A. Vargas, and C. Saavedra, Opt. Exp. **19**, 3542 (2011).
[19] E. Prugovečki, Int. J. Theor. Phys. **16**, 321 (1977).
[20] S. T. Flammia, A. Silberfarb, and C. Caves, Found. Phys. **35**, 1985 (2005).
[21] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, J. Math. Phys. **45**, 2171 (2004).
[22] T. Durt, C. Kurtsiefer, A. Lamas-Linares, and A. Ling, Phys. Rev. A **78**, 042338 (2008).
[23] Z. E. D. Medendorp, F. A. Torres-Ruiz, L. K. Shalm, G. N. M. Tabia, C. A. Fuchs, and A. M. Steinberg, Phys. Rev. A **83**, 051801(R) (2011).
[24] N. Bent, H. Qassim, A. A. Tahir, D. Sych, G. Leuchs, L. L. Sanchez-Soto, E. Karimi, and R. W. Boyd, Phys. Rev. X **5**, 041006 (2015).
[25] W. M. Pimenta, B. Marques, T. O. Maciel, R. O. Vianna, A. Delgado, C. Saavedra, and S. Padua, Phys. Rev. A **88**, 012112 (2013).
[26] R. Salazar and A. Delgado, Phys. Rev. A **86**, 012118 (2012).
[27] C. Paiva-Sanchez, E. Burgos-Inostroza, O. Jiménez, and A. Delgado, Phys. Rev. A **82**, 032115 (2010).
[28] Z. Hradil, Phys. Rev. A **55**, R1561 (1997).
[29] S. M. Tan, J. Mod. Opt. **44**, 2233 (1997).
[30] M. S. Kaznady and D. F. V. James, Phys. Rev. A **79**, 022109 (2009).
[31] R. Blume-Kohout, New J. Phys. **12**, 043034 (2010).
[32] M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, and Y.-K. Liu, Nat. Commun. **1**, 149 (2010).
[33] D. Goyeneche, G. Cañas, S. Etcheverry, E. S. Gomez, G. B. Xavier, G. Lima, and A. Delgado, Phys. Rev. Lett. **115**, 090401 (2015).
[34] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, Phys. Rev. Lett. **105**, 150401 (2010).
[35] R. D. Gill and S. Massar, Phys. Rev. A **61**, 042312 (2000).
[36] E. Bagan, M. A. Ballester, R. D. Gill, R. Muñoz-Tapia, and O. Romero-Isart, Phys. Rev. Lett. **97**, 130501 (2006).
[37] D. H. Mahler, L. A. Rozema, A. Darabi, C. Ferrie, R. Blume-Kohout, and A. M. Steinberg, Phys. Rev. Lett. **111**, 183601 (2013).
[38] S. Straupe, JETP Lett. **104**, 510 (2016).
[39] Z. Hou, H. Zhu, G. Xiang, C.-F. Li, and G.-C. Guo, NPJ Quantum Inf. **2**, 16001 (2016).
[40] M. A. Ballester, J. Phys. A: Math. Gen. **39**, 1645 (2006).
[41] R. Jozsa, J. Mod. Opt. **41**, 2315 (1994).
[42] M. Hübner, Phys. Lett. A **163**, 239 (1992).
[43] R. Bhatia, *Matrix Analysis* (Springer, New York, 1997).
[44] C. W. Helstrom, *Quantum Detection and Estimation Theory*, *Mathematics in Science and Engineering*, Vol. 123 (Academic Press, New York, 1976).
[45] W. K. Wootters, Phys. Rev. D **23**, 357 (1981).
[46] M. Bina, A. Mandarino, S. Olivares, and M. G. A. Paris, Phys. Rev. A **89**, 012305 (2014).
[47] A. Mandarino, M. Bina, C. Porto, S. Cialdi, S. Olivares, and M. G. A. Paris, Phys. Rev. A **93**, 062118 (2016).
[48] K. M. R. Audenaert, J. Calsamiglia, R. Muñoz-Tapia, E. Bagan, L. Masanes, A. Acin, and F. Verstraete, Phys. Rev. Lett. **98**, 160501 (2007).
[49] J. Calsamiglia, R. Muñoz-Tapia, L. Masanes, A. Acin, and E. Bagan, Phys. Rev. A **77**, 032311 (2008).
[50] C. Benedetti, A. P. Shurupov, M. G. A. Paris, G. Brida, and M. Genovese, Phys. Rev. A **87**, 052136 (2013).
[51] E. A. Aguilar, M. Farkas, D. Martínez, M. Alvarado, J. Cariñe, G. B. Xavier, J. F. Barra, G. Cañas, M. Pawłowski, and G. Lima, Phys. Rev. Lett. **120**, 230503 (2018).
[52] S. Etcheverry, G. Cañas, E. S. Gómez, W. A. T. Nogueira, C. Saavedra, G. B. Xavier, and G. Lima, Sci. Rep. **3**, 2316 (2013).
[53] R. Fickler, R. Lapkiewicz, W. N. Plick, M. Krenn, C. Schaeff, S. Ramelow, and A. Zeilinger, Science **338**, 640 (2012).
[54] X.-C. Yao, T.-X. Wang, P. Xu, H. Lu, G.-S. Pan, X.-H. Bao, C.-Z. Peng, C.-Y. Lu, Y.-A. Chen, and J.-W. Pan, Nat. Photon. **6**, 225 (2012).
[55] T. Monz, P. Schindler, J. T. Barreiro, M. Chwalla, D. Nigg, W. A. Coish, M. Harlander, W. Hänsel, M. Hennrich, and R. Blatt, Phys. Rev. Lett. **106**, 130506 (2011).
[56] L. Neves, G. Lima, J. G. Aguirre Gómez, C. H. Monken, C. Saavedra, and S. Padua, Phys. Rev. Lett. **94**, 100501 (2005).
[57] M. A. Solís-Prosser, M. F. Fernandes, O. Jiménez, A. Delgado, and L. Neves, Phys. Rev. Lett. **118**, 100501 (2017).
[58] G. I. Struchalin, E. V. Kovlakov, S. S. Straupe, and S. P. Kulik, arXiv:1804.05226.