Experimental quantum tomography assisted by multiply symmetric states in higher dimensions

D. Martínez,[1,2] M. A. Solís-Prosser,[1,2,*] G. Cañas,[3] O. Jiménez,[4] A. Delgado,[1,2] and G. Lima[1,2]

[1]*Departamento de Física, Universidad de Concepción, 160-C Concepción, Chile*

[2]*Millennium Institute for Research in Optics, Universidad de Concepción, 160-C Concepción, Chile*

[3]*Departamento de Física, Universidad del Bío-Bío, Collao 1202, Casilla 5C, Concepción, Chile*

[4]*Centro de Óptica e Información Cuántica, Facultad de Ciencias, Universidad Mayor, Camino La Pirámide N°5750, Huechuraba, Santiago, Chile*

(Received 21 August 2018; published 22 January 2019)

High-dimensional quantum information processing has become a mature field of research with several different approaches being adopted for the encoding of D-dimensional quantum systems. Such progress has fueled the search of reliable quantum tomographic methods aiming for the characterization of these systems, most of these methods being specifically designed for a given scenario. Here, we report on a tomographic method based on multiply symmetric states and on experimental investigations to study its performance in higher dimensions. Unlike other methods, it is guaranteed to exist in any dimension and provides a significant reduction in the number of measurement outcomes when compared to standard quantum tomography. Furthermore, in the case of odd dimensions, the method requires the least possible number of measurement outcomes. In our experiment we adopt the technique where high-dimensional quantum states are encoded using the linear transverse momentum of single photons and are controlled by spatial light modulators. Our results show that fidelities of 0.984 ± 0.009 with ensemble sizes of only 1.5×10^5 photons in dimension $D = 15$ can be obtained in typical laboratory conditions, thus showing its practicability in higher dimensions.

DOI: 10.1103/PhysRevA.99.012336

I. INTRODUCTION

The generation, manipulation, and measurement of high-dimensional quantum systems (qudits) are important theoretical and experimental research subjects in quantum information science. This is motivated, in part, because certain fundamental features of quantum mechanics such as, for instance, quantum contextuality [1–3], cannot be tested with two-dimensional quantum systems. The use of high-dimensional quantum systems also leads to improvements in several entangled based quantum information protocols since, in this case, some Bell inequalities exhibit increased robustness against noise [4,5], and tolerate lower detection efficiencies for closing the detection loophole [6]. Last, due to the larger amount of information that can be encoded in single qudits, the performance of several protocols in quantum communications [7–12] and quantum computation [13–17] is enhanced when they are employed. Typically, photonic platforms are used as testbed experiments to study quantum information processing in higher dimensions because different degrees of freedom of single photons can be efficiently used to encode the qudits. For instance, one can resort to the orbital angular momentum [18–22], frequency [23–25], time bin [26], path [27], and the transverse position or momentum [28,29] encoding methods.

Quantum tomography (QT) is a collection of methods that makes possible the estimation of unknown quantum states [30]. Today, QT has become a standard tool for the quality assessment of the generation of quantum states [31,32], the implementation of quantum processes [33–35], and the performance of quantum devices [36,37]. Quantum tomographic methods provide an estimate of the unknown state from the outcomes of measurements carried out on an ensemble of identically, independently prepared systems. Finite statistics effects and unavoidable experimental errors require the post-processing of the experimentally acquired data by means of statistical inference methods such as, for instance, maximum likelihood estimation [38–41] or Bayesian inference [42–50]. Traditionally, the total number of measurement outcomes is considered as a resource. Thus, there is a search for QT methods relying on a smaller number of measurement outcomes [51–56]. Standard quantum tomography for a single qudit is based on the measurement of a D-dimensional representation of the $D^2 - 1$ generators of the SU(D) group, which leads to a total number of measurement outcomes of $2D^2 - D$ [57]. This number can be reduced to $D^2 + D$ with quantum tomography based on mutually unbiased bases (MUBs) [58]. The existence of MUBs has been proven when the dimension D is an integer power of a prime number [59,60]. Otherwise, the existence of mutually unbiased bases is still an open problem. A further reduction can be achieved with quantum tomography based on a symmetric informationally complete (SIC) positive-operator valued measure (POVM), which consists of D^2 subnormalized projectors [61]. This is the smallest number of measurement outcomes to estimate unknown quantum states. Numerical studies [62–64] have indicated the existence of this class of measurements in all dimensions up to $D = 1155$ and exact analytical solutions are available in dimensions $D = 2$–16, 19, 24, 28, 35, 48, 120, 124, and 323

*msolisp@udec.cl

[62,65–68]. Unfortunately, a dimension-independent demonstration is still missing.

Here, we propose and experimentally test a new quantum tomographic method, which is based on the measurement of an informationally complete POVM formed by subnormalized projectors onto multiply symmetric states [69]. These are constructed by applying products of integer powers of unitary transformations on a fixed fiducial quantum pure state. Unlike SIC-POVMs and MUBs, our tomographic method can be constructed in any finite dimension. Furthermore, in the case of odd dimensions, the POVM has D^2 subnormalized projectors and thus it requires the smallest number of measurement outcomes to estimate unknown quantum states. In the case of even dimensions, the POVM has $3D^2/2$ measurement operators, which is a significative reduction from the case of standard tomography. A first estimate of the unknown state can be obtained by linear inversion, which does not introduce bias [70]. The numerical stability of this process can be improved to a great extent by a suitable choice of the fiducial state. This also contributes to speed up the rate of convergence in the postprocessing of the experimentally acquired data. Our experimental implementation is based on the encoding of D-dimensional quantum states onto the linear transverse momentum of single photons. These are created by defining D different propagation paths available for the photon transmission at diffractive apertures addressed on spatial light modulators (SLMs) [71]. A second set of SLMs allows one to project the D-dimensional state onto any other fixed D-dimensional state [72]. The use of SLMs for preparation [71–75] and measurement of these so-called spatial qudits has been extensively explored for quantum information tasks such as quantum key distribution (QKD) [76], Bell-type nonlocality and noncontextuality tests [21,77–79], and quantum tomography [53,72,80,81], among others [82–84]. We test our tomographic method in dimensions 6 and 15 reaching fidelities of 0.998 and 0.984 with ensemble sizes of only 6×10^4 and 1.5×10^5, respectively. Experiments performed with similar optical setups have achieved lower fidelities of 0.96 for dimensions 6 and 7, 0.985 for dimension 8, and 0.887 for dimension 10 [53,72,85], while resorting to larger ensembles of detected photons. Thus, our experimental results demonstrate the practicability of our method in higher dimensions and the good performance of our experimental setup.

This article is organized as follows: In Sec. II, we introduce the theoretical background and formulate our tomographic method. In Sec. III, we introduce the experimental setup and analyze the results provided by the experimental realization of our method. In Sec. IV, we summarize, comment on possible extensions to the multipartite case, and conclude.

II. THEORY

In this section we briefly recall the general notion of multiply symmetric states. Thereafter, we study a particular family of multiply symmetric states and build the tomographic method upon it. We solve explicitly the inversion problem and provide a simple analytical expression relating the experimentally acquired data, the measurement settings, and the estimate of the unknown state.

A. Multiply symmetric states

In general, states $|\psi_{k_1,k_2,\ldots,k_M}\rangle$ are said to be multiply symmetric if they can be written as [86]

$$\left|\psi_{k_1,k_2,\ldots,k_M}\right\rangle = U_1^{k_1} U_2^{k_2} \ldots U_M^{k_M} |\psi_{0,0,\ldots,0}\rangle, \tag{1}$$

where $k_j = 0, \ldots, N_j - 1$, $|\psi_{0,0,\ldots,0}\rangle$ is the fiducial state of the set, and U_j are unitary transformations that satisfy $U_j^{N_j} = \mathbb{I}$ (for every j), where $\mathbb{I}$ is the identity operator acting on the Hilbert space of a single qudit. We will limit ourselves to the case of $M = 3$. Thus, we define the constant matrices

$$\mathcal{X} = \sum_{k=0}^{D-1} |k \oplus 1\rangle\langle k|, \tag{2}$$

$$\mathcal{Z} = \sum_{k=0}^{D-1} e^{2\pi i k/D} |k\rangle\langle k|, \tag{3}$$

$$\mathcal{V} = \sum_{k=0}^{\kappa-1} |k\rangle\langle k| - i \sum_{k=\kappa}^{D-1} |k\rangle\langle k|, \tag{4}$$

where D is the dimension of the Hilbert space and $\kappa = [\![D/2]\!]$, $[\![x]\!]$ being the operation that rounds x to the closest integer number. These matrices represent, respectively, the shift operator ($\mathcal{X}$), the clock operator ($\mathcal{Z}$), and an additional phase-only transform ($\mathcal{V}$) with diagonal entries $v_k = \langle k|\mathcal{V}|k\rangle$ that adopt values of 1 and $-i$. The symbol $\oplus$ in Eq. (2) denotes addition mod(D). By using the operators $\mathcal{X}$, $\mathcal{Z}$, and $\mathcal{V}$, we now define a set of multiply symmetric states $\{|\alpha_{\ell,m,j}\rangle\}$ given by

$$|\alpha_{\ell,m,j}\rangle = \mathcal{V}^\ell \mathcal{X}^m \mathcal{Z}^j |\alpha_0\rangle = \sum_{k=0}^{D-1} a_k v_{k\oplus m}^\ell e^{2\pi i j k/D} |k \oplus m\rangle, \tag{5}$$

where $\ell = 0, 1, 2, 3$, $m = 0, \ldots, D-1$, and $j = 0, \ldots, D-1$. The fiducial state is a pure quantum state $|\alpha_0\rangle = \sum_{k=0}^{D-1} a_k |k\rangle$, whose coefficients fulfill the normalization condition $\sum_{k=0}^{D-1} |a_k|^2 = 1$.

B. Essential subsets of states

Let us now consider a physical system described by an unknown D-dimensional quantum state ρ. In 2010, Paiva-Sanchez and co-workers [69] studied quantum state tomography assisted by a basis $\mathcal{B}_0(\alpha)$ of D equidistant states, which they denoted as $|\alpha_j\rangle$. These states are such that the inner product between them is given by $\langle\alpha_j|\alpha_{j'}\rangle = \alpha$ $(j > j')$, where α is a fixed constant. Additional $D-1$ bases $\mathcal{B}_s(\alpha)$ are constructed by applying $\mathcal{X}^s$ on the elements of $\mathcal{B}_0(\alpha)$. This amounts to a total of D^2 measurements. Additionally, they report a strange behavior that depends on the dimension of the Hilbert space to which the state belongs. In summary, odd dimensions require the aforementioned D^2 measurements only, whereas even dimensions require additional measurements attainable by applying $\mathcal{V}$ on the elements of each $\mathcal{B}_s(\alpha)$, which leads to a total of $3D^2/2$ measurements.

The form of the equidistant states used in Ref. [69] for this purpose resembles the one of Eq. (5). Nevertheless, an analysis of the computations of Ref. [69] indicates that a similar mathematical procedure allows us to accomplish such a tomographic process regardless of the fiducial state used,

TABLE I. Values of K_s depending on the values of s and dimension, where $\boldsymbol{\kappa} = [\![D/2]\!]$.

Value of s	K_s odd dimension	K_s even dimension
$0 \leqslant s \leqslant \boldsymbol{\kappa} - 1$	D	$2D$
$\boldsymbol{\kappa} \leqslant s \leqslant D - 1$	D	D
$D \leqslant s \leqslant D - 1 + \boldsymbol{\kappa}$	does not apply	$2D$

that is, states $|\alpha_j\rangle$ do not need to be equidistant. Thus, we resorted to multiply symmetric states for such a goal. If D is an odd number, quantum state tomography can be performed by measuring on projectors of the form $|\alpha_{0,m,j}\rangle\langle\alpha_{0,m,j}|$, with $\ell = 0$ and for every m and j ranging from 0 to $D-1$. For even dimensions, we must also consider $\ell = 1$, with $j = 0, \ldots, D-1$ and $m = 0, \ldots, D/2-1$ for these additional measurements.

In this context, a simpler mathematical description can be obtained by resorting to two subscripts only, regardless of the dimension. Thus, we define

$$\begin{aligned} |\alpha_{sj}\rangle &= \mathcal{V}^{\lfloor s/D \rfloor} \mathcal{X}^s \mathcal{Z}^j |\alpha_0\rangle \\ &= \sum_{k=0}^{D-1} a_k v_{k\oplus s}^{(s)} e^{2\pi i j k/D} |k \oplus s\rangle, \end{aligned} \tag{6}$$

where $v_k^{(s)} = \langle k|\mathcal{V}^{\lfloor s/D \rfloor}|k\rangle$, $j = 0, \ldots, D-1$, $s = 0, \ldots, s_{\mathsf{max}} - 1$, and

$$s_{\mathsf{max}} = \begin{cases} D, & \text{if } D \text{ is odd,} \\ \frac{3D}{2}, & \text{if } D \text{ is even.} \end{cases} \tag{7}$$

Thus, the set of states $\{|\alpha_{sj}\rangle\}$ in odd dimensions is still a complete set of multiply symmetric states under transformations $\mathcal{X}$ and $\mathcal{Z}$, as seen from Eq. (5). For even dimensions, on the other hand, this set encompasses a *subset* of the multiply symmetric states under the action of $\mathcal{X}$, $\mathcal{Z}$, and $\mathcal{V}$. Despite the different behavior exhibited by states $|\alpha_{sj}\rangle$ as defined here, they allow one to construct POVMs. Indeed, we may define

$$\Pi_{sj} = \frac{1}{K_s} |\alpha_{sj}\rangle\langle\alpha_{sj}|, \qquad \sum_{s=0}^{s_{\mathsf{max}}-1} \sum_{j=0}^{D-1} \Pi_{sj} = \mathbb{I}, \tag{8}$$

where the values of K_s are given in Table I. This POVM will be useful for tomographic and postprocessing purposes.

C. Tomography using multiply symmetric states

Let us define the matrix $\mathcal{P} = \sum_{s,j} p_{sj} |s\rangle\langle j|$, where $p_{sj} = \mathrm{tr}(\rho \Pi_{sj})$. Explicitly,

$$\begin{aligned} \mathcal{P} = \sum_{s=0}^{s_{\mathsf{max}}-1} \sum_{j=0}^{D-1} \Bigg(\sum_{l,m=0}^{D-1} \frac{a_l^* a_m}{K_s} e^{2\pi i (m-l) j/D} \\ \times v_{l\oplus s}^{(s)*} v_{m\oplus s}^{(s)} \rho_{l\oplus s, m\oplus s} \Bigg) |s\rangle\langle j|. \end{aligned} \tag{9}$$

This matrix contains the experimental probabilities that can be found by taking the completeness relation of Eq. (8) into consideration. So, if n_{sj} is the number of registered counts when Π_{sj} is measured, then every probability can be experimentally estimated as $p_{sj} = n_{sj} / \sum_{t,k} n_{tk}$. Afterwards, a right-Fourier transformed probability matrix $\widetilde{\mathcal{P}}$ can be defined as $\mathcal{P} \cdot \mathcal{F}$, where $\mathcal{F} = \frac{1}{\sqrt{D}} \sum_{l,m=0}^{D-1} e^{2\pi i l m/D} |l\rangle\langle m|$. Explicitly,

$$\begin{aligned} \widetilde{\mathcal{P}} = \sum_{k=0}^{D-1} \Bigg[\sum_{s=0}^{s_{\mathsf{max}}-1} \frac{\sqrt{D}}{K_s} |s\rangle \\ \times \Bigg(\sum_{q=0}^{D-1} a_{q\ominus s\oplus k}^* a_{q\ominus s} v_{q\oplus k}^{(s)*} v_q^{(s)} \langle q| \Bigg) |\vec{\boldsymbol{\rho}}_k\rangle \Bigg] \langle k|, \end{aligned} \tag{10}$$

where $|\vec{\boldsymbol{\rho}}_m\rangle$ denotes the mth diagonal of ρ, given by[1]

$$|\vec{\boldsymbol{\rho}}_m\rangle = \sum_{q=0}^{D-1} \rho_{q\oplus m, q} |q\rangle. \tag{11}$$

For convenience, we will define ancillary vectors

$$|\vec{\boldsymbol{\xi}}_{sk}\rangle = (\mathcal{X}^{s-k}|\alpha_0\rangle) \circ (\mathcal{X}^s |\alpha_0\rangle^*) \circ (\mathcal{X}^{-k} |\vec{\boldsymbol{v}}_s\rangle) \circ |\vec{\boldsymbol{v}}_s\rangle^*, \tag{12}$$

where "$\circ$" denotes the Hadamard product between matrices, and

$$|\vec{\boldsymbol{v}}_s\rangle = \sum_{r=0}^{D-1} v_r^{(s)} |r\rangle = \mathrm{diag}(\mathcal{V}^{\lfloor s/D \rfloor}). \tag{13}$$

Consequently, the right-transformed probability matrix can be compactly written as

$$\widetilde{\mathcal{P}} = \sum_{k=0}^{D-1} \mathscr{G}_k |\vec{\boldsymbol{\rho}}_k\rangle\langle k|, \tag{14}$$

where the matrix $\mathscr{G}_k$ is given by

$$\mathscr{G}_k = \sum_{s=0}^{s_{\mathsf{max}}-1} \frac{\sqrt{D}}{K_s} |s\rangle\langle\vec{\boldsymbol{\xi}}_{sk}|. \tag{15}$$

Now, it is possible to construct the density operator ρ by rearranging its components in a vector $\vec{\rho} = \mathrm{vec}(\rho)$ (see Appendix A), which is computed according to

$$\vec{\boldsymbol{\Delta}}_\rho = \sum_{k=0}^{D-1} |k\rangle \otimes |\vec{\boldsymbol{\rho}}_k\rangle, \tag{16}$$

$$\vec{\rho} = \boldsymbol{\mathcal{S}}_{\mathrm{SWAP}} \boldsymbol{\mathcal{X}} \cdot \vec{\boldsymbol{\Delta}}_\rho, \tag{17}$$

where $\otimes$ represents the Kronecker product between matrices, $\vec{\boldsymbol{\Delta}}_\rho$ is a D^2-dimensional vector containing the diagonals of ρ—given by $|\vec{\boldsymbol{\rho}}_k\rangle$—stacked on top of each other, $\boldsymbol{\mathcal{S}}_{\mathrm{SWAP}}$ is a $D^2 \times D^2$ matrix that acts as $\boldsymbol{\mathcal{S}}_{\mathrm{SWAP}}(|j\rangle \otimes |k\rangle) = |k\rangle \otimes |j\rangle$, and

$$\boldsymbol{\mathcal{X}} = \left(\sum_{m=0}^{D-1} \mathcal{X}^m \otimes |m\rangle\langle m| \right). \tag{18}$$

[1]Throughout this article, the notation $|\vec{\boldsymbol{a}}\rangle$ will refer to a purely mathematical vector $\vec{\boldsymbol{a}}$ that does not represent any physical state. However, Dirac notation is used for comfortability.

Finally, after taking Eqs. (10), (12), (14), (15), and (17) into account, the components of ρ can be isolated by computing

$$\mathcal{G} = \sum_{m=0}^{D-1} |m\rangle\langle m| \otimes \mathscr{G}_m, \tag{19}$$

$$\mathcal{G}^{\dashv} = \sum_{m=0}^{D-1} |m\rangle\langle m| \otimes \mathscr{G}_m^{\dashv}, \tag{20}$$

and

$$\begin{aligned}\vec{\rho} &= \boldsymbol{\mathcal{S}}_{\text{SWAP}}\, \boldsymbol{\mathcal{X}}\, \mathcal{G}^{\dashv} \operatorname{vec}(\widetilde{\mathcal{P}}) \\ &= (\boldsymbol{\mathcal{S}}_{\text{SWAP}}\, \boldsymbol{\mathcal{X}}\, \mathcal{G}^{\dashv}) \left(\left(\mathcal{F} \otimes \mathbb{I}_{s_{\mathsf{max}}} \right) \operatorname{vec}(\mathcal{P}) \right),\end{aligned} \tag{21}$$

where $x^{\dashv}$ denotes the Moore-Penrose pseudoinverse [87] matrix of x, $\mathbb{I}_{s_{\mathsf{max}}}$ is a $s_{\mathsf{max}} \times s_{\mathsf{max}}$ identity matrix, and $\operatorname{vec}(\mathcal{P})$ is the vectorization of matrix $\mathcal{P}$ (see Appendix A). We have used Eq. (A3) in order to write $\operatorname{vec}(\mathcal{P}\mathcal{F}) = (\mathcal{F} \otimes \mathbb{I}) \operatorname{vec}(\mathcal{P})$, $\mathcal{F}$ being a symmetric matrix. We have resorted to vectorized versions of some matrices as these allow one to write efficient numerical codes. The use of sparse matrices, as explained in Sec. II D, may contribute substantially to the efficiency of the computation of $\vec{\rho}$ for quantum systems of very high dimensions. Matrix pseudoinverse has been used instead of the usual matrix inverse because $\mathcal{G}$ contains Ds_{max} rows and D^2 columns and, consequently, may not be square. Equation (21), in summary, relates the components of the reconstructed density matrix—stored in vector $\vec{\rho}$—with the experimental measurements ($\mathcal{P}$) and the measurement settings ($\mathcal{G}$) in an explicit way. The density operator ρ is obtained by just rearranging the elements of $\vec{\rho}$ in a square matrix, which can be postprocessed if required. Large parentheses in Eq. (21) indicate the recommended multiplication order with the goal of optimizing the use of memory.

D. Comparison with general linear inversion

Let us consider an arbitrary M-outcome quantum measurement described by POVM elements Π_μ, where $\mu = 1, \ldots, M$. If $\vec{p}$ is a vector containing the probabilities associated with this measurement, it can be shown [see Eq. (A5)] that $\vec{p} = \boldsymbol{\mathcal{M}}\vec{\rho}$, where

$$\boldsymbol{\mathcal{M}} = \left(\sum_{\mu=1}^{M} |\operatorname{vec}(\Pi_\mu)\rangle\langle\mu| \right)^{\dagger}. \tag{22}$$

Thus, a simple way to obtain $\vec{\rho}$ is by means of $\vec{\rho} = \boldsymbol{\mathcal{M}}^{\dashv}\vec{p}$. If the set of measurements is not informationally complete, there will be ambiguities in the state, i.e., there can exist several solutions $\vec{\rho}$ for the problem $\vec{p} = \boldsymbol{\mathcal{M}}\vec{\rho}$ and, thus, we should expect the performance of the reconstruction to be poor in terms of fidelity, as it may depend highly on the algorithm used to compute $\boldsymbol{\mathcal{M}}^{\dashv}$. The use of informationally complete measurements eliminates these ambiguities.

Moreover, the computation of $\boldsymbol{\mathcal{M}}^{\dashv}$ for the method studied in this article might be unfeasible in very high dimensions as it requires (i) being able to store in memory a highly dense matrix $\boldsymbol{\mathcal{M}}$ of size $D^2 \times Ds_{\mathsf{max}}$, and (ii) the ability of computing its pseudoinverse. For instance, matrix $\boldsymbol{\mathcal{M}}$ for an eight-qubit system ($D = 256$, $s_{\mathsf{max}} = 384$) requires 96 GB of memory for the sole purpose of being stored using complex double-precision floating-point numbers. Thus, the computation of $\boldsymbol{\mathcal{M}}^{\dashv}$ for a multiqubit or multiqudit system might be impractical in most current computers.

On the other hand, we may see from the previous section that matrices $\boldsymbol{\mathcal{S}}_{\text{SWAP}}$, $\boldsymbol{\mathcal{X}}$, $\mathcal{G}^{\dashv}$, and $(\mathcal{F} \otimes \mathbb{I}_{s_{\mathsf{max}}})$ are very sparse for high dimensions: their densities (ratio of nonzero elements) are $1/D^2$, $1/D^2$, $1/D$, and $1/s_{\mathsf{max}}$, respectively. Consequently, the aforementioned matrices for the eight-qubit case require less than 1 GB of memory each when using sparse matrices,[2] making now the computations possible in many computers.

Due to finite statistic effects and experimental error sources, the estimates generated via linear inversion might not be positive semidefinite matrices, and thus cannot be accepted as physical states. To solve this problem, several statistical inference techniques can be employed—in particular, the maximum likelihood estimation (MLE). This is formulated as an optimization process on the space of the physical states whose output is the state with the highest probability of generating the experimentally acquired data. MLE requires an initial guess. This is chosen in our case as the matrix provided by linear inversion, which contributes to speed up the convergence of MLE. Thus, both MLE and linear inversion are employed to generate a physically acceptable estimate. MLE is formulated as an optimization problem on an exponentially scaling number of variables, and thus its computational feasibility is constrained by the available computing power. Recently, efficient techniques for solving MLE in higher dimensional systems have been proposed [88]. Thereby, the slow convergence of MLE is postponed to even higher dimensions, but not eliminated. In this scenario, linear inversion becomes a viable alternative since it is less computationally demanding than MLE. Moreover, unlike estimates obtained with the help of MLE, the linear inversion process does not exhibit bias [70,89].

Other alternative approaches such as forced purity, quick and dirty [90], or searching for the closest density operator [91] are suitable to find physically acceptable states without resorting to numerically demanding optimization problems. All of them require an initial matrix to work on, which can be obtained from $\vec{\rho}$ by linear inversion.

E. Stability of the inversion

The stability of the inversion under variations of the experimentally obtainable probabilities can be studied by inspecting Eq. (21). The problem is either well or ill-conditioned depending on the condition number $\mathcal{C}$ of the matrix involved in the inversion. This, in turn, depends on the singular values of such matrix [92]. As matrices $\boldsymbol{\mathcal{S}}_{\text{SWAP}}$, $\boldsymbol{\mathcal{X}}$, and $\mathcal{F} \otimes \mathbb{I}_{s_{\mathsf{max}}}$ are all unitary, they do not modify singular values and, hence, matrix $\mathcal{G}$ suffices to analyze the robustness of the tomographic

[2]Numerical tests for the eight-qubit case show that $(\mathcal{F} \otimes \mathbb{I}_{s_{\mathsf{max}}})$ and $\mathcal{G}^{\dashv}$ require 576.75 MB each, whereas $\boldsymbol{\mathcal{S}}_{\text{SWAP}}$ and $\boldsymbol{\mathcal{X}}$ need only 1.5 MB each.

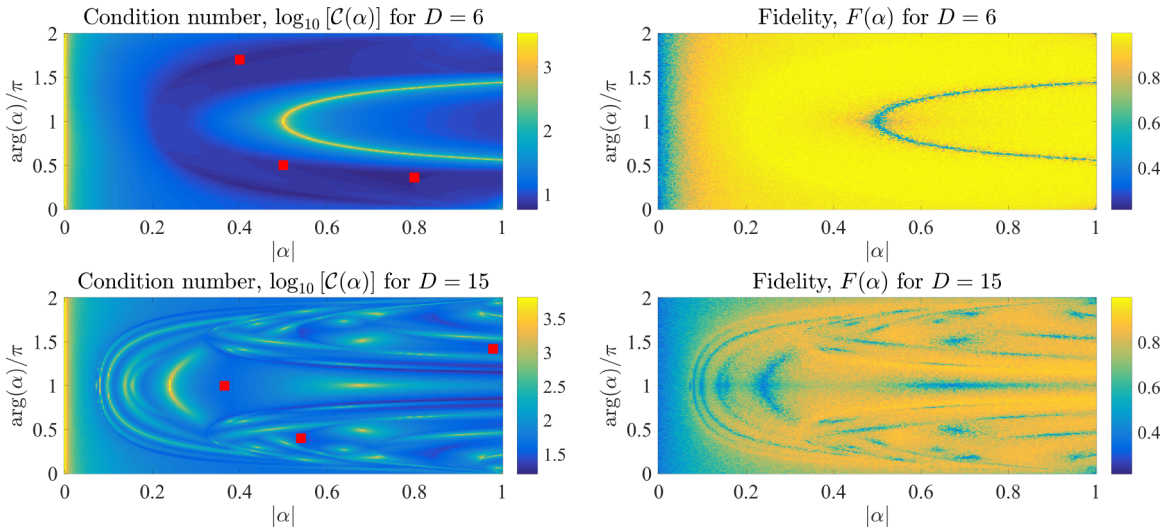

FIG. 1. Left panels: Condition number $\mathcal{C}(\mathcal{G}(\alpha))$ as function of α—given that fiducial state $|\alpha_0\rangle$ is given by Eqs. (24) and (25)—for dimensions 6 and 15. Red squares indicate the values of α that were used in the experiment reported in this article. Since $\mathcal{C}(\mathcal{G}(\alpha))$ might adopt very different values, these graphs were presented in logarithmic scale. Right panels: Fidelity of reconstructed states simulated by Monte Carlo method.

procedure under experimental noise. Indeed,

$$\mathcal{C}\left(\mathcal{S}_{\text{SWAP}}\,\mathcal{X}\,\mathcal{G}^{\dashv}\left(\mathcal{F}\otimes\mathbb{I}_{s_{\max}}\right)\right)=\mathcal{C}(\mathcal{G})=\frac{\sigma_{\max}(\mathcal{G})}{\sigma_{\min}(\mathcal{G})}, \tag{23}$$

where $\sigma_{\max}(\mathcal{G})$ and $\sigma_{\min}(\mathcal{G})$ stand for the maximal and minimal singular values of $\mathcal{G}$, respectively. It can be concluded from Eqs. (12), (19), and (23) that a study of $\mathcal{C}(\mathcal{G})$ as a function of $|\alpha_0\rangle$ allows one to predict whether a given fiducial state will be a good choice for quantum tomography. A small condition number indicates that the fiducial state is a good candidate for building the tomographic method.

As $\mathcal{C}(\mathcal{G})$ depends on D complex parameters, its optimization over the Hilbert space does not seem to be computationally easy. For the sake of simplicity, we will resort to the notation used in Ref. [69] in order to analyze $\mathcal{C}(\mathcal{G})$ in terms of a single complex parameter α. Thus, the fiducial state will be given by

$$|\alpha_0(\alpha)\rangle=\sum_{k=0}^{D-1}\sqrt{\frac{\lambda_k(\alpha)}{D}}|k\rangle, \tag{24}$$

where

$$\lambda_k(\alpha)=1-|\alpha|\frac{\sin\left(\frac{k\pi+(D-1)\arg(\alpha)}{D}\right)}{\sin\left(\frac{k\pi-\arg(\alpha)}{D}\right)}. \tag{25}$$

The left panels of Fig. 1 show the decimal logarithm of $\mathcal{C}(\mathcal{G})$ as a function of the absolute value and phase of α for dimensions 6 and 15. As can be observed, $\mathcal{C}(\mathcal{G}(\alpha))$ can adopt values ranging from $\sim 10^1$ to $\sim 10^5$. The right panels show the fidelity of the reconstruction via Monte Carlo simulations, in which the simulated number of counts by measurement had Poisson noise added. As many simulations were performed, we implemented the "quick and dirty" method [90] for post-processing the density operator in Fig. 1. Direct comparison of left and right panels indicate that numerical instability (high condition number) may introduce a significant inaccuracy in the estimate of the density matrix. This further motivated us to look for adequate fiducial states before making measurements.

In each of the left panels of Fig. 1, three red squares highlight the values of α that were used for our experiment. These values, also displayed in Table II, were chosen from regions at the figures exhibiting small condition numbers. We have dealt with the problem of numerical stability by choosing fiducial states such that $\mathcal{C}(\mathcal{G})$ adopts small values. Instead of resorting to a given parametrization, we could have generated a large set of random fiducial states and computed the value of $\mathcal{C}(\mathcal{G})$ for each one. If done so, condition numbers even lower than the ones used here could be obtained. Nonetheless, using the former procedure we were able to ensure the states in a neighborhood with small condition numbers to have a more robust reconstruction in the case of having noise due to experimental imperfections.

III. EXPERIMENT

Our setup is depicted in Fig. 2. It consists of two main blocks: the *state preparation* (SP) and *projective tomographic measurement* (PM) stages. In SP, weak coherent states are produced resorting to a 690-nm continuous-wave single-mode laser heavily attenuated with calibrated optical filters (not shown in Fig. 2 for the sake of simplicity) and modulated with

TABLE II. Values of α chosen for experimental purposes. Numbers in parentheses below each α indicate the condition number, which is extracted from data of Fig. 1.

D	α_1	α_2	α_3
6	$0.4e^{1.7\pi i}$	$0.8e^{0.36\pi i}$	$0.5e^{0.5\pi i}$
	(7.796)	(6.848)	(8.946)
15	$0.365e^{\pi i}$	$0.54e^{0.4\pi i}$	$0.98e^{1.42\pi i}$
	(40.94)	(27.32)	(33.15)

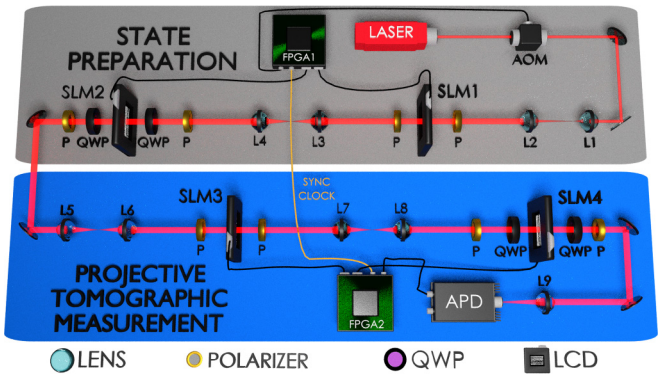

FIG. 2. Experimental setup. At the preparation state we generate an arbitrary state of a single qudit that is encoded in a single photon. A 690-nm continuous-wave laser, an AOM, and calibrated attenuators (not shown in figure for sake of clarity), are employed to generate weak coherent pulses. These illuminate spatial light modulators SLM1 and SLM2, which with the help of polarizers (P) and quarter-wave plates (QWP) are employed to modulate the incoming light in amplitude and phase, correspondingly. Electronically addressable slit patterns at SLM1 and SLM2 control the state $|\Psi^d\rangle$ of the photonic qudit. The measurement stage projects state $|\Psi^d\rangle$ onto a predefined, arbitrary single-qudit state $|\alpha_{sj}\rangle$. Spatial light modulators SLM3 and SLM4, combined with a pointlike avalanche photodiode (APD), implement the projection. In this way we are able to estimate the probability $|\langle\alpha_{sj}|\Psi^d\rangle|^2$. A set of lenses is employed to transport the images generated by the SLMs along the setup. Focal lengths are as follows: L1 = 25 mm, L2 = 200 mm, L3 = L4 = L5 = L6 = L7 = L8 = 125 mm, and L9 = 100 mm.

an acousto-optic modulator (AOM) configured at a repetition rate of 30 Hz. The mean photon number per pulse is set to $\mu = 0.9$. Consequently, this source works as an approximation to a nondeterministic single-photon source since pulses with a single photon account for 62.3% of the generated non-null pulses [93]. Contributions of multiphoton events to the recorded statistics are strongly suppressed by using a detection window much smaller than the optical pulse duration. Last, extra polarizing cubes with an overall extinction ratio greater than 10^{-7} are used to ensure a high quality of horizontal polarization of the transmitted photons. In this way, we are able to attain a high purity degree for the high-dimensional states generated with the spatial light modulators [94].

SLMs are a central part of our setup. Each pixel of a SLM is part of a twisted nematic liquid crystal display (LCD) whose birefringence can be controlled by means of standard video signals emitted by a field programable gate array (FPGA). As a result of an adequate control of the photon polarization before and after crossing the LCD, we can set the SLM to work as an amplitude-only spatial light modulator (SLM1 and SLM3) or as a phase-only modulator (SLM2 and SLM4) [95]. Arrays of D slits are displayed on SLM1, each having a width of 96 μm and transmittance coefficients t_ℓ. The centers of contiguous slits are separated by 192 μm. An imaging system projects the image of SLM1 on SLM2, where phases ϕ_ℓ are added to each slit. Thus, the state of the single photons transmitted by these SLMs is $|\Psi\rangle \propto \sum_{\ell=0}^{D-1} \sqrt{t_\ell} e^{i\phi_\ell}|\ell\rangle$—where $|\ell\rangle$ denotes the state of the photon transmitted by the ℓth slit of the SLMs—and it represents a D-dimensional quantum system encoded into the linear transverse momentum of the photons [29,71,72].

To test our tomographic method we considered three different types of states for dimension $D = 6$ and $D = 15$. The reason for choosing such dimensions are (i) to illustrate the relevance of our method while considering even and odd dimensions, and (ii) the tomographic method based on mutually unbiased bases cannot be used in these dimensions. To be more specific, the prepared states were

$$|\Psi_1^6\rangle = \frac{1}{\sqrt{6}} \sum_{j=0}^{5} |j\rangle, \tag{26a}$$

$$|\Psi_2^6\rangle = |0\rangle, \tag{26b}$$

$$|\Psi_3^6\rangle = \frac{1}{\sqrt{6}}[(|0\rangle + |2\rangle) + e^{-i\pi/4}|4\rangle + e^{-i\pi/8}(|1\rangle + |3\rangle + |5\rangle)], \tag{26c}$$

for dimension 6, and

$$|\Psi_1^{15}\rangle = \frac{1}{\sqrt{15}} \sum_{j=0}^{14} |j\rangle, \tag{27a}$$

$$|\Psi_2^{15}\rangle = |7\rangle, \tag{27b}$$

$$|\Psi_3^{15}\rangle = \frac{1}{\sqrt{15}}[(|0\rangle + |5\rangle + |8\rangle + |14\rangle) + e^{-i\pi/10}(|1\rangle + |3\rangle + |9\rangle + |12\rangle) + e^{-i\pi/9}|2\rangle + e^{-i\pi/8}(|6\rangle + |11\rangle) + e^{-i\pi/7}(|4\rangle + |10\rangle) + e^{-i\pi/6}(|7\rangle + |13\rangle)] \tag{27c}$$

for dimension 15.

The projective tomographic measurement stage contains SLM3 and SLM4, used for postselecting the state to be detected. For this purpose, a new set of transmittance coefficients τ_ℓ and phases ζ_ℓ are used on SLM3 and SLM4, respectively. Finally, detection is performed at the center of the focal plane of a lens located after SLM4 using an avalanche single-photon detector (APD) with a 10-μm-wide pinhole placed in front of it. The probability of detecting a single photon is, thus, proportional to $|\langle\Theta|\Psi\rangle|^2$ [72,76,96], where $|\Theta\rangle \propto \sum_\ell \sqrt{\tau_\ell} e^{-i\zeta_\ell}|\ell\rangle$. In our case, $|\Theta\rangle$ represents each of the states $|\alpha_{sj}\rangle$ on which the measurements are performed, so parameters τ_ℓ and ζ_ℓ are adjusted accordingly. In order to show the possibility of using different fiducial states, we used $|\alpha_0(\alpha_k)\rangle$ [see Eq. (24)] as the fiducial state for reconstructing state $|\Psi_k^D\rangle$, where the values of α_k are the ones shown in Table II.

Each projective measurement related to our tomographic method was repeated ten times, which allowed us to obtain its associated mean value of detection counts. We denote $n_{sj,r}$ as the counts obtained from measuring Π_{sj} in the rth round of measurements, where $r = 1, \ldots, 10$. Average numbers of counts $\bar{n}_{sj}$ can then be obtained by

$$\bar{n}_{sj} = \frac{1}{10} \sum_{r=1}^{10} n_{sj,r}. \tag{28}$$

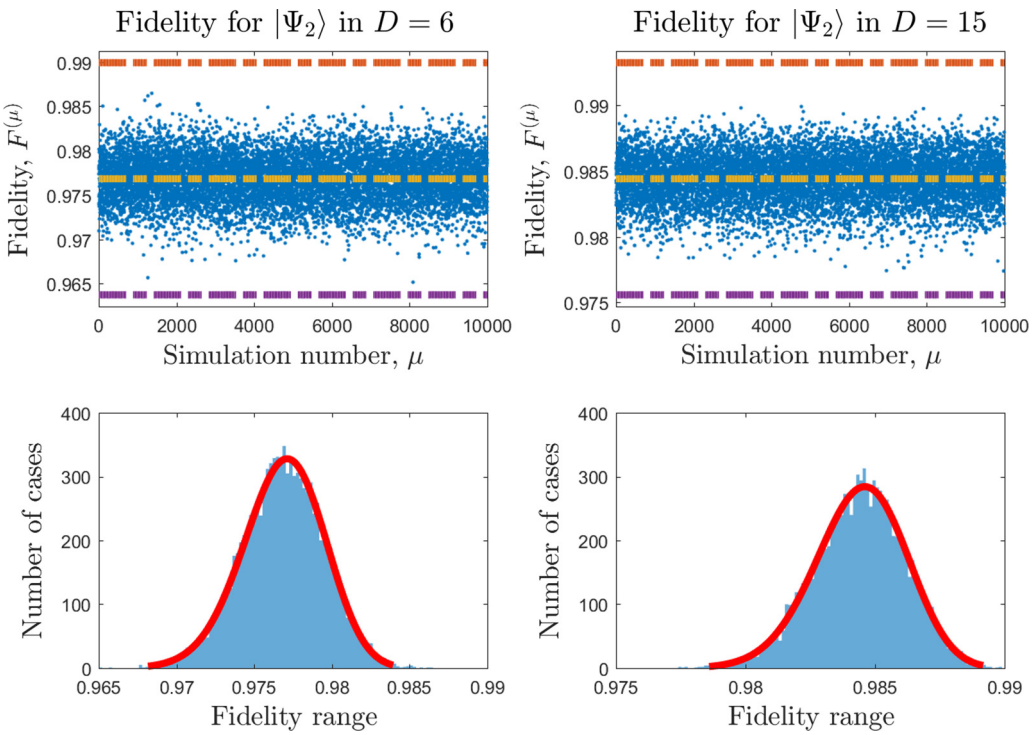

FIG. 3. Examples of the Monte Carlo simulations performed (upper panels) and their respective histograms (lower panels). Three horizontal dot-dashed lines in the upper graphs represent the mean value ($\langle\langle F\rangle\rangle$) of the simulations and the $\langle F\rangle \pm 5\sigma$ interval. The continuous line in each histogram represents a fitted beta distribution. For each fitted function, the probability of having a value outside the $\pm 5\sigma$ interval is $\sim 10^{-6}$.

Measurements were taken for times long enough to obtain around $10\,000 \times D$ counted photons, as a total over all projections Π_{sj}, for each round. Once all the detection counts were recorded and the average probabilities were computed, we proceeded to the postmeasurement processing of the data. Error margins for density matrices and its corresponding figures of merit were determined through 10 000 Monte Carlo simulations for each reconstructed state. Simulated count numbers $n_{sj}^{(\mu)}$ are obtained by adding Poisson noise to the originally averaged data, where μ denotes the number of the Monte Carlo trial and ranges from 1 to 10 000. Only in the first case there is no noise considered, i.e.,

$$n_{sj}^{(\mu)} = \begin{cases} \bar{n}_{sj}, & \text{for } \mu = 1, \\ \text{Poisson}(\bar{n}_{sj}), & \text{otherwise.} \end{cases} \tag{29}$$

Afterwards, simulated probability matrices $\mathcal{P}^{(\mu)}$ are computed for each Monte Carlo trial according to

$$\mathcal{P}^{(\mu)} = \frac{\sum_{s=0}^{s_{\max}-1} \sum_{j=0}^{D-1} n_{sj}^{(\mu)} |s\rangle\langle j|}{\sum_{t=0}^{s_{\max}-1} \sum_{k=0}^{D-1} n_{tk}^{(\mu)}}, \tag{30}$$

where Eq. (8) was taken into account. Then, matrix $\mathcal{P}^{(\mu)}$ is used in Eq. (21) in order to obtain a reconstructed density matrix $\rho^{(\mu)}$. As Eq. (21) cannot ensure its positiveness, maximum likelihood estimation (MLE) was subsequently employed (see Appendix B for details) in order to ensure matrix positiveness [38,39,57]. The fidelity $F^{(\mu)}$ between $\rho^{(\mu)}$ and the state $|\Psi\rangle$ we intended to prepare is computed as a figure of merit for each state resulting from MLE, where

$$F^{(\mu)} = \langle\Psi|\rho^{(\mu)}|\Psi\rangle. \tag{31}$$

The final result for the fidelity is expressed in terms of the mean and standard deviation of the simulated results, that is,

$$F = \langle\{F^{(\mu)}\}\rangle \pm 5\sigma(\{F^{(\mu)}\}). \tag{32}$$

Two examples of the Monte Carlo simulations are shown in Fig. 3. We have chosen $\pm 5\sigma$ as error margins since the probability of obtaining a value outside it in a new round of experiments is less than 10^{-6} in the case the values of $F^{(\mu)}$ distribute around their mean value following a normal distribution. In the worst-case scenario, such probability is less than 4%, according to the Bienaymé-Chebyshev inequality. The reconstructed density operators in dimension 6 are depicted in Fig. 4, whereas Fig. 5 illustrates the results for dimension 15. A summary of the results is shown in Table III. As can be seen, high values of fidelities were obtained. More specifically, for dimension 6 (15) an overall fidelity of 0.977 (0.957) has been recorded, while considering an ensemble of only 6×10^4 (1.5×10^5) events of photodetection for the state reconstruction procedure. In the literature there are several experiments toward quantum tomography of a single qudit: Reference [72] reported the experimental realization of tomography using mutually unbiased bases and they obtained fidelities of 0.96 ± 0.03 and 0.93 ± 0.03 for dimensions 7 and 8, respectively. Reference [53] reported 0.985 ± 0.015 for $D = 8$ using a method designed for reconstructing pure states. The experiment of Ref. [85] using SIC-POVM obtained fidelities of 0.960 ± 0.003 and 0.887 ± 0.003 for dimensions 6 and 10, respectively. Our experimental results compare favorably with these previous results, which validates the good performance of the experimental setup employed to realize the

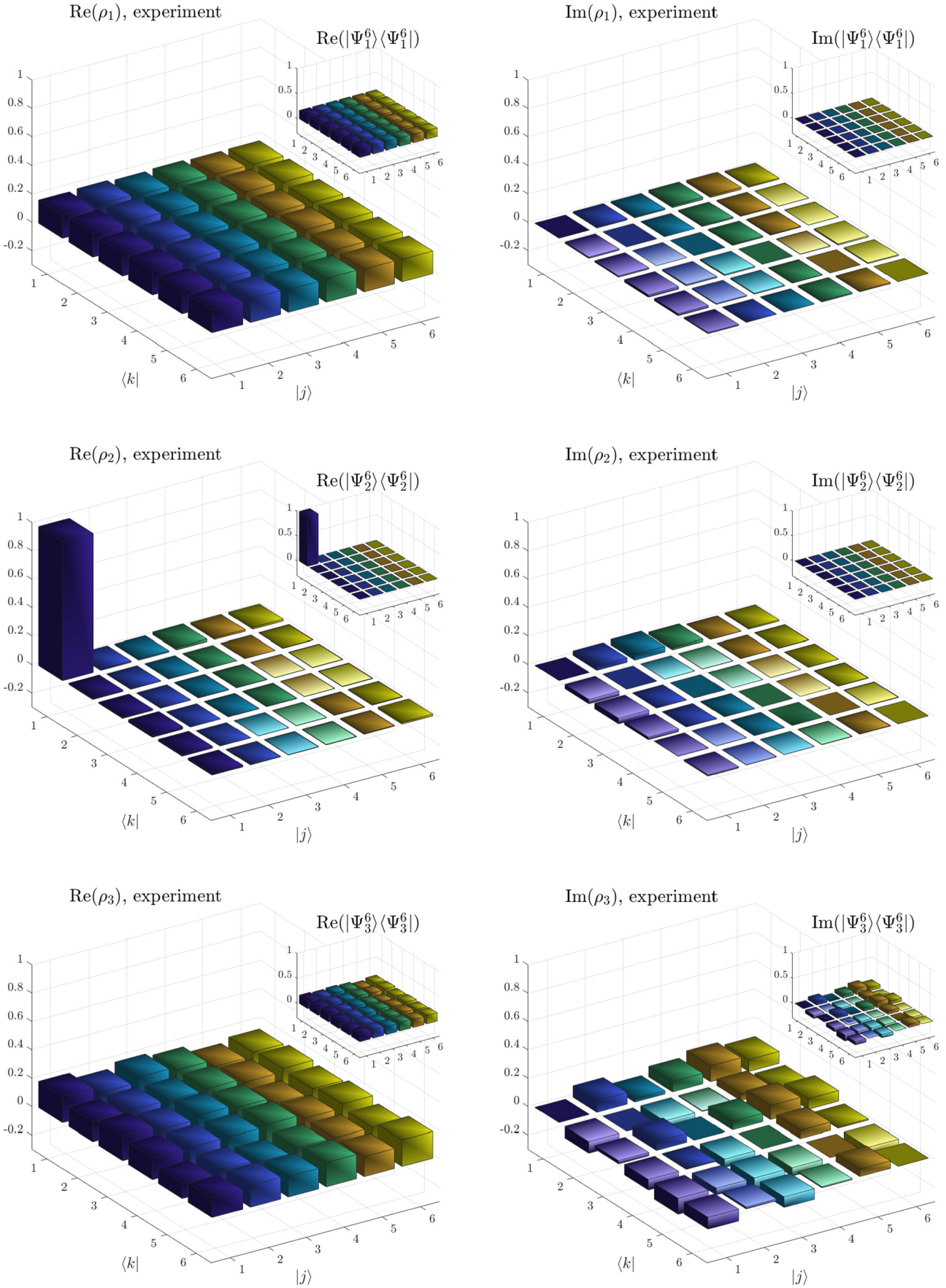

FIG. 4. Reconstructed quantum states for $D = 6$. The insets show the theoretically expected results.

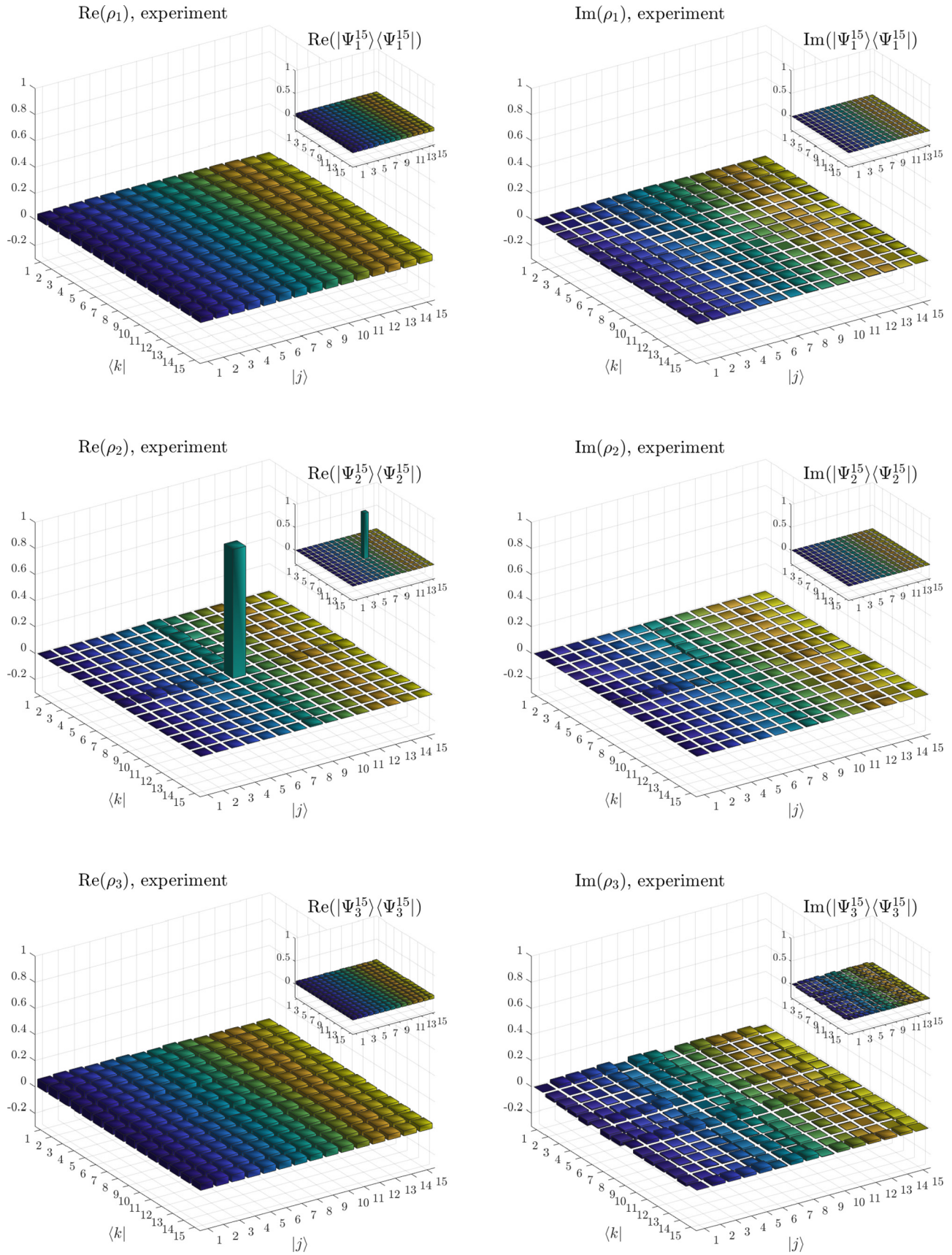

FIG. 5. Reconstructed quantum states for $D = 15$. The insets show the theoretically expected results.

TABLE III. Fidelities obtained for each of the reconstructed states, with their respective 5σ uncertainty extracted from 10 000 Monte Carlo trials. MLE was used in each trial.

D	$\|\Psi_1^D\rangle$	$\|\Psi_2^D\rangle$	$\|\Psi_3^D\rangle$
6	0.998 ± 0.001	0.977 ± 0.013	0.956 ± 0.010
15	0.965 ± 0.006	0.984 ± 0.009	0.922 ± 0.007

quantum tomography assisted by multiply symmetric states in higher dimensions.

IV. CONCLUDING REMARKS

In summary, we have reported the experimental realization of quantum state tomography assisted by multiply symmetric states for dimensions $D = 6$ and $D = 15$. Unlike MUB and SIC-POVM tomographic methods, this method is guaranteed to exist in any dimension and provides a significant reduction in the number of measurement outcomes when compared to standard quantum tomography. Furthermore, in the case of odd dimensions the method requires the least possible number of measurement outcomes. As explained in Sec. II D, this tomographic method is different from a general linear inversion in the sense that (i) multiply symmetric states constitute an informationally complete set of measurements, and (ii) it was possible to rewrite the equations in such a way that the inversion is represented now in terms of multiplying sparse matrices and computing inverses of smaller matrices rather than computing the inverse of a large matrix, which is more practical for high dimensions. The POVM elements used for this reconstruction method, as Eq. (6) shows, depend on a given fiducial state $|\alpha_0\rangle$ that can be freely chosen. Nevertheless, this fiducial state is chosen in such a way that the inversion algorithm remains numerically stable. Such stability can be analyzed in terms of the condition number of matrix $\mathcal{G}$ of Eq. (19).

Further improvements can be obtained by studying the condition number. We have reduced the complexity of this problem by studying fiducial states defined by two parameters, which led to condition numbers of the order of 10. However, Monte Carlo simulations with randomly generated fiducial states have shown that smaller condition numbers are possible. Other continuation of the current work concerns the case of multipartite systems. For instance, the state of a two-qudit system can be estimated with a minimal number of D^4 measurement outcomes. This can be achieved for D odd by conditional local estimations employing quantum tomography assisted by multiply symmetric states.

Recently, QT has been studied from the point of view of the achievable estimation accuracy. Here, the figure of merit is the Gill-Massar lower bound $\bar{I} = (D^2 - 1)(D + 1)/4N$ for the infidelity, where N is the size of the ensemble of identically, independently prepared copies of the unknown state to be estimated. This is the highest estimation accuracy for mixed states that can be achieved by means of local measurements, that is, measurements that are carried out on individual members of the ensemble. It has been demonstrated that two-stage quantum tomography for a single qudit approaches $\bar{I}$ [97]. In the first stage of this adaptive tomographic method a small ensemble is employed to obtain a first estimate via standard quantum tomography. This estimate's eigenstates are employed to represent the $D^2 - 1$ generators of SU(D), which are subsequently measured in a second stage of standard quantum tomography. For $D = 2$ two-stage quantum tomography saturates the Gill-Massar lower bound [98,99]. However, for $D > 2$ this is not the case. Furthermore, numerical evidence suggests that the estimation accuracy behaves as $\bar{I}D$. It is possible to improve the accuracy by modifying two-stage quantum tomography. Instead of projecting onto the eigenstates of the $D^2 - 1$ Gell-Mann generators, it is possible to gather enough information to estimate an unknown state by projecting onto $2D + 1$ ($2D - 1$) bases for D odd (even). The adaptive version of this method leads to an estimation accuracy that behaves as $2\bar{I}$ [100], independently of the dimension. Here arises the question whether an adaptive version of quantum tomography assisted by multiply symmetric states would lead to an estimation accuracy better than $2\bar{I}$ with the added benefit of a reduced number of total measurement outcomes.

ACKNOWLEDGMENTS

This work was supported by CONICYT FONDECYT Grants No. 1160400, No. 3170400, No. 11150324, No. 11121318, and No. 1180558, and Millennium Institute for Research in Optics (MIRO). D.M. acknowledges financial support from CONICYT Doctorado Nacional Grant No. 2116050.

APPENDIX A: MATRIX VECTORIZATION

Let us consider a general $M \times N$ matrix $A = \sum_{kl} A_{kl}|k\rangle\langle l|$ written in the computational basis. Its vectorization $\mathrm{vec}(A)$ is obtained from a linear operation such that

$$\begin{aligned} &\mathrm{vec} : \mathbb{C}^{M\times N} \to \mathbb{C}^N \otimes \mathbb{C}^M \\ &A \mapsto \mathrm{vec}(A) = \sum_{kl} A_{kl}|l\rangle \otimes |k\rangle, \end{aligned} \tag{A1}$$

where vectors $\{|j\rangle\}_{j=1}^n$ correspond to the computational basis. Vector $\mathrm{vec}(A)$ contains the columns of A stacked one on top of another. Hence, we can also write a matrix vectorization as

$$\mathrm{vec}(A) = \sum_{j=1}^{N} |j\rangle \otimes A|j\rangle. \tag{A2}$$

Additional properties of vectorization are

$$\mathrm{vec}(AB) = (\mathbb{I} \otimes A)\mathrm{vec}(B) = (B^{\intercal} \otimes \mathbb{I})\mathrm{vec}(A), \tag{A3}$$

$$\mathrm{tr}(A^{\dagger}B) = \langle \mathrm{vec}(A)|\mathrm{vec}(B)\rangle. \tag{A4}$$

Faster computations of probabilities can be performed by resorting to matrix vectorization. Indeed, if we need to compute a vector given by $\mathbf{p} = \sum_{\mu} \mathrm{tr}(\Pi_{\mu}\rho)|\mu\rangle$, where Π_{μ} are

Hermitian operators, then

$$\begin{aligned}\mathbf{p} &= \sum_{\mu} \mathrm{tr}(\Pi_\mu \rho)|\mu\rangle \\ &\overset{(A4)}{=} \sum_{\mu} \langle \mathrm{vec}(\Pi_\mu)|\mathrm{vec}(\rho)\rangle|\mu\rangle \\ &= \left(\sum_{\mu} |\mu\rangle\langle \mathrm{vec}(\Pi_\mu)|\right)|\mathrm{vec}(\rho)\rangle \\ &= \hat{\mathbf{\Pi}}^\dagger |\mathrm{vec}(\rho)\rangle, \end{aligned} \tag{A5}$$

where $\hat{\mathbf{\Pi}}$ is a matrix whose μth column is the vectorization of Π_μ. Equation (A5) is very useful when probabilities must be computed a large number of times from a constant set of probability operators.

APPENDIX B: EFFICIENT COMPUTATION OF MLE BASED ON POISSON DISTRIBUTION

We define $N_j^{\mathrm{th}}(\varrho) = \eta_j \mathrm{tr}(\Pi_j \varrho) + d_j$ as the theoretically expected number of counts for the jth detector subject to detection efficiency η_j and a mean number of dark counts given by d_j. The matrix ϱ is a positive operator representing an *unnormalized* density matrix whose trace serves as a mean ensemble size. Additionally, $\boldsymbol{n}$ and $\boldsymbol{N}_{\mathrm{th}}$ will be vectors containing the experimental and theoretical number of counts, respectively. Vector $\boldsymbol{N}_{\mathrm{th}}(\varrho)$ can be efficiently written, with the aid of Eq. (A5), as

$$\boldsymbol{N}_{\mathrm{th}}(\varrho) = \boldsymbol{\eta} \circ (\hat{\mathbf{\Pi}}^\dagger \vec{\boldsymbol{\varrho}}) + \boldsymbol{d}, \tag{B1}$$

where

$$\boldsymbol{\eta} = \sum_{j\in\mathcal{M}} \eta_j |j\rangle, \quad \boldsymbol{d} = \sum_{j\in\mathcal{M}} d_j |j\rangle, \tag{B2}$$

$$\vec{\boldsymbol{\varrho}} = |\mathrm{vec}(\varrho)\rangle, \quad \hat{\mathbf{\Pi}} = \sum_{j\in\mathcal{M}} |\mathrm{vec}(\Pi_j)\rangle\langle j|. \tag{B3}$$

In the case of composite systems, computation of $\hat{\mathbf{\Pi}}^\dagger \vec{\boldsymbol{\varrho}}$ can be performed very efficiently by following the methods used in Ref. [88].

As the number of photons provided by the source cannot be ensured to be equal for each measurement, we may resort to Poisson statistics, where the number of observed events will be given by n_j and the expected mean number of events is $N_j^{\mathrm{th}}(\varrho)$. Thus, the joint probability of having n_1 counts in detector 1 *and* n_2 counts in detector 2 *and* so and so, *given that* the ensemble state is ϱ, can be expressed as

$$\mathcal{L}_{\mathrm{P}}(\varrho) = \prod_{j\in\mathcal{M}} \frac{e^{-N_j^{\mathrm{th}}(\varrho)}\left[N_j^{\mathrm{th}}(\varrho)\right]^{n_j}}{n_j!}, \tag{B4}$$

where $\mathcal{M}$ is the set of measurements labels. Instead of working with $\mathcal{L}_{\mathrm{P}}(\varrho)$, it is highly recommended to make use of the negative log likelihood, $L_{\mathrm{P}}(\varrho) = -\ln \mathcal{L}_{\mathrm{P}}(\varrho)$, since maximization of $\mathcal{L}_{\mathrm{P}}(\varrho)$ is equivalent to minimization of $L_{\mathrm{P}}(\varrho)$ [101]. Negative log likelihood is, thus, given by

$$\begin{aligned} L_{\mathrm{P}}(\varrho) &= \sum_{j\in\mathcal{M}} \left[N_j^{\mathrm{th}}(\varrho) - n_j \ln\left(N_j^{\mathrm{th}}(\varrho)\right) + \ln\Gamma(n_j+1)\right] \\ &= \boldsymbol{u}^\intercal \boldsymbol{N}_{\mathrm{th}}(\varrho) - \boldsymbol{n}^\intercal \ln(\boldsymbol{N}_{\mathrm{th}}(\varrho)) + \boldsymbol{u}^\intercal \ln(\Gamma(\boldsymbol{n}+1)), \end{aligned} \tag{B5}$$

where $\boldsymbol{u}$ is a column vector whose entries are all equal to 1 and the logarithm of a vector is used as $[\ln(\boldsymbol{a})]_j = \ln([\boldsymbol{a}]_j)$. By using Eqs. (B1)–(B3), the negative log likelihoods of Eq. (B5) can be efficiently computed. It is noteworthy that $L_{\mathrm{P}}(\varrho)$ is a convex function defined over the convex set of positive operators. Consequently, the tools of convex optimization (particularly, we used CVX [102,103]) can be used to find its minima, $L_{\mathrm{P}}(\varrho_{\mathrm{opt}})$. The estimated density operator ρ_{opt} is obtained by normalizing ϱ_{opt} after the optimization has finished.

[1] E. Specker, Dialectica **14**, 239 (1960).
[2] S. Kochen and E. P. Specker, in *Logico-Algebraic Approach to Quantum Mechanics* (Springer, Dordrecht, 1975), pp. 293–328.
[3] G. Cañas, S. Etcheverry, E. S. Gómez, C. Saavedra, G. B. Xavier, G. Lima, and A. Cabello, Phys. Rev. A **90**, 012119 (2014).
[4] D. Kaszlikowski, P. Gnaciński, M. Żukowski, W. Miklaszewski, and A. Zeilinger, Phys. Rev. Lett. **85**, 4418 (2000).
[5] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu, Phys. Rev. Lett. **88**, 040404 (2002).
[6] T. Vértesi, S. Pironio, and N. Brunner, Phys. Rev. Lett. **104**, 060401 (2010).
[7] Č. Brukner, M. Żukowski, and A. Zeilinger, Phys. Rev. Lett. **89**, 197901 (2002).
[8] D. Bruß and C. Macchiavello, Phys. Rev. Lett. **88**, 127901 (2002).
[9] N. J. Cerf, M. Bourennane, A. Karlsson, and N. Gisin, Phys. Rev. Lett. **88**, 127902 (2002).
[10] T. Durt, N. J. Cerf, N. Gisin, and M. Żukowski, Phys. Rev. A **67**, 012311 (2003).
[11] G. M. Nikolopoulos, K. S. Ranade, and G. Alber, Phys. Rev. A **73**, 032325 (2006).
[12] I. Ali-Khan, C. J. Broadbent, and J. C. Howell, Phys. Rev. Lett. **98**, 060503 (2007).
[13] D. P. O'Leary, G. K. Brennen, and S. S. Bullock, Phys. Rev. A **74**, 032334 (2006).
[14] B. P. Lanyon, M. Barbieri, M. P. Almeida, T. Jennewein, T. C. Ralph, K. J. Resch, G. J. Pryde, J. L. O'Brien, A. Gilchrist, and A. G. White, Nat. Phys. **5**, 134 (2009).
[15] M.-X. Luo, X.-B. Chen, Y.-X. Yang, and X. Wang, Sci. Rep. **4**, 4044 (2014).
[16] Z. Gedik, I. A. Silva, B. Çakmak, G. Karpat, E. L. Vidoto, D. O. Soares-Pinto, E. R. DeAzevedo, and F. F. Fanchini, Sci. Rep. **5**, 1 (2015).
[17] H. S. Tonchev and N. V. Vitanov, Phys. Rev. A **94**, 042307 (2016).
[18] A. Mair, A. Vaziri, G. Weihs, and A. Zeilinger, Nature (London) **412**, 313 (2001).

[19] N. K. Langford, R. B. Dalton, M. D. Harvey, J. L. O'Brien, G. J. Pryde, A. Gilchrist, S. D. Bartlett, and A. G. White, Phys. Rev. Lett. **93**, 053601 (2004).

[20] B. Jack, J. Leach, H. Ritsch, S. M. Barnett, M. J. Padgett, and S. Franke-Arnold, New J. Phys. **11**, 103024 (2009).

[21] A. C. Dada, J. Leach, G. S. Buller, M. J. Padgett, and E. Andersson, Nat. Phys. **7**, 677 (2011).

[22] M. Agnew, E. Bolduc, K. J. Resch, S. Franke-Arnold, and J. Leach, Phys. Rev. Lett. **113**, 020501 (2014).

[23] C. K. Law, I. A. Walmsley, and J. H. Eberly, Phys. Rev. Lett. **84**, 5304 (2000).

[24] C. Bernhard, B. Bessire, T. Feurer, and A. Stefanov, Phys. Rev. A **88**, 032322 (2013).

[25] M. Kues, C. Reimer, P. Roztocki, L. R. Cortés, S. Sciara, B. Wetzel, Y. Zhang, A. Cino, S. T. Chu, B. E. Little, D. J. Moss, L. Caspani, J. Azaña, and R. Morandotti, Nature (London) **546**, 622 (2017).

[26] H. Takesue and Y. Noguchi, Opt. Express **17**, 10976 (2009).

[27] A. Rossi, G. Vallone, A. Chiuri, F. De Martini, and P. Mataloni, Phys. Rev. Lett. **102**, 153902 (2009).

[28] L. Neves, S. Pádua, and C. Saavedra, Phys. Rev. A **69**, 042305 (2004).

[29] L. Neves, G. Lima, J. G. Aguirre Gómez, C. H. Monken, C. Saavedra, and S. Pádua, Phys. Rev. Lett. **94**, 100501 (2005).

[30] *Quantum State Estimation*, edited by M. Paris and J. Řeháček, Lecture Notes in Physics Vol. 649 (Springer, Berlin, 2004).

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

[32] A. Ourjoumtsev, H. Jeong, R. Tualle-Brouri, and P. Grangier, Nature (London) **448**, 784 (2007).

[33] I. L. Chuang and M. A. Nielsen, J. Mod. Opt. **44**, 2455 (1997).

[34] I. L. Chuang, L. M. K. Vandersypen, X. L. Zhou, D. W. Leung, and S. Lloyd, Nature (London) **393**, 143 (1998).

[35] M. Mohseni, A. T. Rezakhani, and D. A. Lidar, Phys. Rev. A **77**, 032322 (2008).

[36] G. M. D'Ariano, M. D. Laurentis, M. G. A. Paris, A. Porzio, and S. Solimeno, J. Opt. B **4**, S127 (2002).

[37] J. G. Titchener, M. Gräfe, R. Heilmann, A. S. Solntsev, A. Szameit, and A. A. Sukhorukov, npj Quantum Info. **4**, 19 (2018).

[38] Z. Hradil, Phys. Rev. A **55**, R1561 (1997).

[39] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, Phys. Rev. A **64**, 052312 (2001).

[40] D. F. V. James, P. G. Kwiat, Z. Hradil, J. Rehacek, and A. G. White, Quantum state reconstruction: A comparison of maximum likelihood and tomographic schemes, in *Technical Digest, Summaries of Papers Presented at the Quantum Electronics and Laser Science Conference, Postconference Technical Digest (IEEE Cat. No. 01CH37172), Baltimore, MD, USA* (IEEE, Piscataway, NJ, 2001), p. 238.

[41] J. Řeháček, Z. Hradil, and M. Ježek, Phys. Rev. A **63**, 040303 (2001).

[42] K. Jones, Ann. Phys. (NY) **207**, 140 (1991).

[43] P. B. Slater, Phys. Lett. A **206**, 66 (1995).

[44] V. Bužek, R. Derka, G. Adam, and P. Knight, Ann. Phys. (NY) **266**, 454 (1998).

[45] R. Schack, T. A. Brun, and C. M. Caves, Phys. Rev. A **64**, 014305 (2001).

[46] F. Huszár and N. M. T. Houlsby, Phys. Rev. A **85**, 052120 (2012).

[47] K. S. Kravtsov, S. S. Straupe, I. V. Radchenko, N. M. T. Houlsby, F. Huszár, and S. P. Kulik, Phys. Rev. A **87**, 062122 (2013).

[48] G. I. Struchalin, I. A. Pogorelov, S. S. Straupe, K. S. Kravtsov, I. V. Radchenko, and S. P. Kulik, Phys. Rev. A **93**, 012103 (2016).

[49] C. Granade, J. Combes, and D. G. Cory, New J. Phys. **18**, 033024 (2016).

[50] D. S. Gonçalves, C. L. N. Azevedo, C. Lavor, and M. A. Gomes-Ruggiero, J. Appl. Stat. **45**, 1846 (2018).

[51] 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).

[52] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, Phys. Rev. Lett. **105**, 150401 (2010).

[53] D. Goyeneche, G. Cañas, S. Etcheverry, E. S. Gómez, G. B. Xavier, G. Lima, and A. Delgado, Phys. Rev. Lett. **115**, 090401 (2015).

[54] Q. Pears Stefano, L. Rebón, S. Ledesma, and C. Iemmi, Phys. Rev. A **96**, 062328 (2017).

[55] G. Tóth, W. Wieczorek, D. Gross, R. Krischek, C. Schwemmer, and H. Weinfurter, Phys. Rev. Lett. **105**, 250403 (2010).

[56] T. Moroder, P. Hyllus, G. Tóth, C. Schwemmer, A. Niggebaum, S. Gaile, O. Gühne, and H. Weinfurter, New J. Phys. **14**, 105001 (2012).

[57] R. T. Thew, K. Nemoto, A. G. White, and W. J. Munro, Phys. Rev. A **66**, 012303 (2002).

[58] W. K. Wootters and B. D. Fields, Ann. Phys. (NY) **191**, 363 (1989).

[59] S. Bandyopadhyay, O. Boykin, V. Roychowdhury, and F. Vatan, Algorithmica **34**, 512 (2002).

[60] T. Durt, B.-G. Englert, I. Bengtsson, and K. Życzkowski, Int. J. Quantum Inf. **08**, 535 (2010).

[61] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, J. Math. Phys. **45**, 2171 (2004).

[62] A. J. Scott and M. Grassl, J. Math. Phys. **51**, 042203 (2010).

[63] C. A. Fuchs and B. C. Stacey, Axioms **6**, 21 (2010).

[64] M. Grassl and A. J. Scott, J. Math. Phys. **58**, 122201 (2017).

[65] D. M. Appleby, J. Math. Phys. **46**, 052107 (2005).

[66] D. M. Appleby, I. Bengtsson, S. Brierley, M. Grassl, D. Gross, and J. A. Larsson, Quantum Inf. Comput. **12**, 404 (2012).

[67] D. M. Appleby, I. Bengtsson, S. Brierley, A. Ericsson, M. Grassl, and J. A. Larsson, Quantum Inf. Comput. **14**, 339 (2014).

[68] D. M. Appleby, T.-Y. Chien, S. Flammia, and S. Waldron, J. Phys. A: Math. Theor. **51**, 165302 (2018).

[69] C. Paiva-Sánchez, E. Burgos-Inostroza, O. Jiménez, and A. Delgado, Phys. Rev. A **82**, 032115 (2010).

[70] C. Schwemmer, L. Knips, D. Richart, H. Weinfurter, T. Moroder, M. Kleinmann, and O. Gühne, Phys. Rev. Lett. **114**, 080403 (2015).

[71] G. Lima, A. Vargas, L. Neves, R. Guzmán, and C. Saavedra, Opt. Express **17**, 10688 (2009).

[72] G. Lima, L. Neves, R. Guzmán, E. S. Gómez, W. A. T. Nogueira, A. Delgado, A. Vargas, and C. Saavedra, Opt. Express **19**, 3542 (2011).

[73] M. A. Solís-Prosser, A. Arias, J. J. M. Varga, L. Rebón, S. Ledesma, C. Iemmi, and L. Neves, Opt. Lett. **38**, 4762 (2013).

[74] J. J. M. Varga, L. Rebón, M. A. Solís-Prosser, L. Neves, S. Ledesma, and C. Iemmi, J. Phys. B **47**, 225504 (2014).
[75] J. J. M. Varga, S. Ledesma, C. Iemmi, and L. Rebón, Phys. Rev. A **96**, 032309 (2017).
[76] S. Etcheverry, G. Cañas, E. S. Gómez, W. A. T. Nogueira, C. Saavedra, G. B. Xavier, and G. Lima, Sci. Rep. **3**, 2316 (2013).
[77] G. Cañas, M. Arias, S. Etcheverry, E. S. Gómez, A. Cabello, G. B. Xavier, and G. Lima, Phys. Rev. Lett. **113**, 090404 (2014).
[78] M. Arias, G. Cañas, E. S. Gómez, J. F. Barra, G. B. Xavier, G. Lima, V. D'Ambrosio, F. Baccari, F. Sciarrino, and A. Cabello, Phys. Rev. A **92**, 032126 (2015).
[79] G. Cañas, E. Acuña, J. Cariñe, J. F. Barra, E. S. Gómez, G. B. Xavier, G. Lima, and A. Cabello, Phys. Rev. A **94**, 012337 (2016).
[80] W. M. Pimenta, B. Marques, M. A. Carvalho, M. R. Barros, J. G. Fonseca, J. Ferraz, M. Terra Cunha, and S. Pádua, Opt. Express **18**, 24423 (2010).
[81] W. M. Pimenta, B. Marques, T. O. Maciel, R. O. Vianna, A. Delgado, C. Saavedra, and S. Pádua, Phys. Rev. A **88**, 012112 (2013).
[82] M. A. Solís-Prosser, M. F. Fernandes, O. Jiménez, A. Delgado, and L. Neves, Phys. Rev. Lett. **118**, 100501 (2017).
[83] 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).
[84] D. Martínez, A. Tavakoli, M. Casanova, G. Cañas, B. Marques, and G. Lima, Phys. Rev. Lett. **121**, 150504 (2018).
[85] N. Bent, H. Qassim, A. A. Tahir, D. Sych, G. Leuchs, L. L. Sánchez-Soto, E. Karimi, and R. W. Boyd, Phys. Rev. X **5**, 041006 (2015).
[86] S. M. Barnett, Phys. Rev. A **64**, 030303 (2001).
[87] R. Penrose and J. A. Todd, Math. Proc. Cambridge Philos. Soc. **51**, 406 (1955).
[88] J. Shang, Z. Zhang, and H. K. Ng, Phys. Rev. A **95**, 062336 (2017).
[89] G. B. Silva, S. Glancy, and H. M. Vasconcelos, Phys. Rev. A **95**, 022107 (2017).
[90] M. S. Kaznady and D. F. V. James, Phys. Rev. A **79**, 022109 (2009).
[91] J. A. Smolin, J. M. Gambetta, and G. Smith, Phys. Rev. Lett. **108**, 070502 (2012).
[92] W. Cheney and D. Kincaid, *Numerical Mathematics and Computing*, 6th ed. (Thomson Brooks/Cole, Belmont, CA, 2008).
[93] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Rev. Mod. Phys. **74**, 145 (2002).
[94] F. A. Torres-Ruiz, G. Lima, A. Delgado, S. Pádua, and C. Saavedra, Phys. Rev. A **81**, 042104 (2010).
[95] I. Moreno, P. Velásquez, C. R. Fernández-Pousa, M. M. Sánchez-López, and F. Mateos, J. Appl. Phys. **94**, 3697 (2003).
[96] G. Taguchi, T. Dougakiuchi, N. Yoshimoto, K. Kasai, M. Iinuma, H. F. Hofmann, and Y. Kadoya, Phys. Rev. A **78**, 012307 (2008).
[97] L. Pereira, L. Zambrano, J. Cortés-Vega, S. Niklitschek, and A. Delgado, Phys. Rev. A **98**, 012339 (2018).
[98] D. H. Mahler, L. A. Rozema, A. Darabi, C. Ferrie, R. Blume-Kohout, and A. M. Steinberg, Phys. Rev. Lett. **111**, 183601 (2013).
[99] Z. Hou, H. Zhu, G. Xiang, C.-F. Li, and G.-C. Guo, NPJ Quantum Inf. **2**, 16001 (2016).
[100] L. Pereira, D. Martinez, G. Cañas, G. Lima, and A. Delgado (unpublished).
[101] K. Banaszek, G. M. D'Ariano, M. G. A. Paris, and M. F. Sacchi, Phys. Rev. A **61**, 010304 (1999).
[102] M. Grant and S. Boyd, Graph implementations for nonsmooth convex programs, in *Recent Advances in Learning and Control*, edited by V. D. Blondel, S. P. Boyd, and H. Kimura, Part of the Lecture Notes in Control and Information Sciences book series (LNCIS, volume 371) (Springer-Verlag Limited, London, 2008), pp. 95–110.
[103] M. Grant and S. Boyd, CVX: Matlab Software for Disciplined Convex Programming, version 2.1, http://cvxr.com/cvx, March 2014.