OPEN

Estimation of pure quantum states in high dimension at the limit of quantum accuracy through complex optimization and statistical inference

Leonardo Zambrano[1,2], Luciano Pereira[1,2], Sebastián Niklitschek[1,3] & Aldo Delgado[1,2✉]

Quantum tomography has become a key tool for the assessment of quantum states, processes, and devices. This drives the search for tomographic methods that achieve greater accuracy. In the case of mixed states of a single 2-dimensional quantum system adaptive methods have been recently introduced that achieve the theoretical accuracy limit deduced by Hayashi and Gill and Massar. However, accurate estimation of higher-dimensional quantum states remains poorly understood. This is mainly due to the existence of incompatible observables, which makes multiparameter estimation difficult. Here we present an adaptive tomographic method and show through numerical simulations that, after a few iterations, it is asymptotically approaching the fundamental Gill–Massar lower bound for the estimation accuracy of pure quantum states in high dimension. The method is based on a combination of stochastic optimization on the field of the complex numbers and statistical inference, exceeds the accuracy of any mixed-state tomographic method, and can be demonstrated with current experimental capabilities. The proposed method may lead to new developments in quantum metrology.

The basic principles of quantum mechanics strongly limit the processes that can be performed in nature. The linearity of quantum operations prohibits the possibility of creating in a deterministic way perfect copies of unknown quantum states[1]. Information encoded in non-orthogonal quantum states can be perfectly recovered only at the expense of allowing a probability of error[2–4]. Measurements, a fundamental tool for understanding the natural world, are also limited. In remarkable difference with the classical realm, a measurement in a quantum system alters the quantum state of the system, which prevents further information gain. In this way, an ensemble of many equally prepared systems is necessary to obtain information about a quantum system and, consequently, quantum measurements become statistical in nature[5–7]. Furthermore, quantum mechanics establishes through the quantum Cramér–Rao bound[8–10] a fundamental limit in the best accuracy that a measurement or estimation process can achieve as a function of the ensemble size.

The accuracy limit also bounds the estimation of unknown quantum processes and states. These are central problems in the theory of quantum measurements[11] and play a key role in the control of quantum systems, the benchmarking of quantum technologies[12–14], and quantum metrology[15]. Today there is a large collection of methods[11,16–25] for estimating unknown quantum states, which are collectively known as quantum tomographic methods. These are based on the post-processing of data acquired through the measurement of a set of positive operator-valued measures in an ensemble of N identically, independently prepared copies of the unknown state to be estimated. Recently, there has been a great deal of research activity aimed at designing tomographic methods with increasing accuracy, in particular methods that saturate the quantum accuracy limit. From the theoretical and experimental point of view, it has been shown that a two-stage adaptive tomographic method achieves the

¹Instituto Milenio de Investigación en Óptica, Universidad de Concepción, Concepción, Chile. ²Departamento de Física, Facultad de Ciencias Físicas y Matemáticas, Universidad de Concepción, Concepción, Chile. ³Departamento de Estadística, Facultad de Ciencias Físicas y Matemáticas, Universidad de Concepción, Concepción, Chile. ✉email: aldelgado@udec.cl

quantum accuracy limit in the case of unknown mixed states of a single qubit[26,27]. Unfortunately, this result does not hold in the higher dimensional case[28].

Here, we report an adaptive tomographic method that asymptotically approaches the quantum accuracy limit in the important case of estimating unknown pure quantum states in high dimensions. This is an instance of multi-parameter estimation, where due to the information trade-off among incompatible observables the progress has been slow. In the case of pure states the quantum accuracy is limited by the Gill–Massar lower bound[29]. The method is based on the concatenation of the Complex simultaneous perturbation stochastic approximation (CSPSA), a recently proposed iterative stochastic optimization method on the field of the complex numbers[30], and Maximum likelihood estimation (MLE), a well known statistical inference method[31]. The method here proposed reaches the quantum accuracy limit after a small number of iterations, typically of the order of 8, for all inspected dimensions. Thereby, the method makes an optimal use of the ensemble size and surpasses the estimation accuracy of known methods for pure-state tomography[32–35]. Moreover, the method also surpasses the estimation accuracy of any tomographic method designed to estimate mixed states via separable measurements on the ensemble of equally prepared copies.

Method

The accuracy achieved in the estimation of the parameters of a quantum state can be studied by means of three fundamental inequalities. These are the Cramér–Rao inequality[36] $\mathcal{C} \leq \mathcal{I}^{-1}$, the quantum Cramér–Rao inequality[36,37] $\mathcal{C} \leq \mathcal{J}^{-1}$, and the Gill–Massar inequality[29] $Tr(\mathcal{I}\mathcal{J}^{-1}) \leq d - 1$, where $d, \mathcal{C}, \mathcal{I}$, and $\mathcal{J}$ are the dimension of the Hilbert space, the covariance matrix, the classical Fisher information matrix, and the quantum Fisher information matrix, respectively. These inequalities allows one to deduce lower bounds for several metrics of accuracy, such as infidelity or mean square error, considering the impact of the ensemble size. As accuracy metric we employ the infidelity[38,39] $I(|\tilde{\psi}\rangle, |\hat{\psi}\rangle) = 1 - |\langle \hat{\psi} | \tilde{\psi} \rangle|^2$ between an unknown state $|\tilde{\psi}\rangle$ and its estimate $|\hat{\psi}\rangle$. The estimation accuracy is given by the expectation or mean value $\bar{I}(|\tilde{\psi}\rangle)$ of the infidelity with respect to all possible estimates $|\hat{\psi}\rangle$ of a fixed unknown state $|\tilde{\psi}\rangle$, that is,

$$\bar{I}(|\tilde{\psi}\rangle) = \mathbb{E}[I(|\tilde{\psi}\rangle, |\hat{\psi}\rangle) | |\hat{\psi}\rangle] = \int I(|\tilde{\psi}\rangle, |\hat{\psi}\rangle) p(|\hat{\psi}\rangle) d\hat{\psi}, \tag{1}$$

where $p(|\hat{\psi}\rangle)$ is the probability density function of obtaining the estimate $|\hat{\psi}\rangle$. Depending on the general characteristics of the estimation process, the fundamental inequalities lead to various lower bounds for the estimation accuracy. For the estimation of pure states, which have $2(d - 1)$ independent parameters, the ultimate quantum estimation accuracy is given by the inequality $\bar{I}(|\tilde{\psi}\rangle) \geq \bar{I}_p$, with $\bar{I}_p = (d - 1)/N$[29,40,41] the Gill–Massar lower bound. This state-independent bound can be used as a benchmark to assess tomographic methods.

Our main goal is designing a tomographic method for pure quantum states in high dimensions that achieves an estimation accuracy equal to $\bar{I}_p$. The design of the proposed tomographic method originates from observing[42–45] that an unknown and fixed pure quantum state $|\tilde{\psi}\rangle$ can be characterized as the minimizer of the infidelity in the Hilbert space of the estimate $|\hat{\psi}\rangle$, that is, $I = 0$ when $|\hat{\psi}\rangle = |\tilde{\psi}\rangle$. Then, we can envision the use of an optimization method to iteratively drive a sequence of estimates toward decreasing values of the infidelity.

The choice of the optimization method requires certain consideration. Traditional optimization methods are based on the evaluation of higher derivatives of the function to be optimized. In the case of the infidelity this is not possible, since the derivatives depend on the state to be estimated, which is unknown. We also require a fast convergence rate from the optimization method. Furthermore, the implementation of the method should be as economical as possible from the point of view of computing and physical resources.

To deal with these situations we resort to the Complex simultaneous perturbation stochastic approximation. This method can optimize real valued functions of complex variables that also depend on unknown complex parameters. The infidelity is a non-holomorphic function, that is, it violates the Cauchy-Schwarz conditions, which are necessary for the existence of a complex derivative. In this case, the usual approach is to optimize with respect to a real parametrization of the complex variables entering in the infidelity, that is, the complex probability amplitudes. This approach has some unwanted side effects. The elements of the real-valued gradient of the infidelity are in general more convoluted than would be those of a complex gradient formed by first order derivatives with respect to the initial complex variables. Additionally, any inherent structures present in the complex derivatives of the infidelity, which could be exploited to enhance the performance of optimization methods, are unused[46]. The CSPSA method has been designed to avoid these unwanted features. The main tool in the formulation of CSPSA is the Wirtinger complex calculus. For a target function $f(\mathbf{z}, \mathbf{z}^*) : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{R}$ the Wirtinger derivatives are defined by[47]

$$\partial_{z_i} = \frac{1}{2}(\partial_{x_i} - i\partial_{y_i}) \text{ and } \partial_{z_i^*} = \frac{1}{2}(\partial_{x_i} + i\partial_{y_i}), \tag{2}$$

where x_i and y_i are the real and imaginary parts of z_i, respectively. These derivatives exist even if $f(\mathbf{z}, \mathbf{z}^*)$ is non-holomorphic. Extremal points of $f(\mathbf{z}, \mathbf{z}^*)$ are completely characterized by the conditions $\partial_{z_i^*}f = 0 \,\forall\, i = 1, \ldots, d$ or, equivalently, $\partial_{z_i}f = 0 \,\forall\, i = 1, \ldots, d$[48–50]. Thereby, the complex gradient is defined by $\mathbf{g} = \partial_{\mathbf{z}^*}f$ with $\partial_{\mathbf{z}^*} = (\partial_{z_1^*}, \ldots, \partial_{z_d^*})$. The CSPSA method is defined by the iterative rule[30]

$$\hat{\mathbf{z}}_{k+1} = \hat{\mathbf{z}}_k - a_k \hat{\mathbf{g}}_k(\hat{\mathbf{z}}_k, \hat{\mathbf{z}}_k^*), \tag{3}$$

where a_k is a positive gain coefficient and $\hat{\mathbf{z}}_k$ is the estimate of the minimizer $\tilde{\mathbf{z}}$ of $f(\mathbf{z}, \mathbf{z}^*)$ at the k-th iteration. The iteration starts from an initial guess $\hat{\mathbf{z}}_0$, which is randomly chosen. Instead of employing the complex gradient

g, CSPSA resorts to an estimator $\hat{g}_k(\hat{z}_k, \hat{z}_k^*)$ for the Wirtinger gradient g of $f(z, z^*)$ the components of which are defined by

$$\hat{g}_{k,i} = \frac{f(\hat{z}_{k+}, \hat{z}_{k+}^*) + \epsilon_{k,+} - (f(\hat{z}_{k-}, \hat{z}_{k-}^*) + \epsilon_{k,-})}{2 c_k \Delta_{k,i}^*}, \tag{4}$$

with $\hat{z}_{k\pm} = \hat{z}_k \pm c_k \boldsymbol{\Delta}_k$, c_k a positive gain coefficient, and $\epsilon_{k,\pm}$ describes the presence of noise in the values of $f(\hat{z}_{k\pm}, \hat{z}_{k\pm}^*)$. The components of the vector $\boldsymbol{\Delta}_k \in \mathbb{C}^n$ are identically and independently distributed random variables in the set $\{\pm 1, \pm i\}$. The gain coefficients a_k and c_k control the convergence of CSPSA and are chosen as

$$a_k = \frac{a}{(10k + 1 + A)^s}, \quad c_k = \frac{b}{(10k + 1)^r}, \tag{5}$$

where the values of a, A, s, b and r can be adjusted to increase the rate of convergence. The estimates $\hat{z}_k$ provided by CSPSA converge asymptotically in mean to the minimizer of $\tilde{z}$ of f and $\hat{g}_k$ is an asymptotically unbiased estimator of the Wirtinger gradient.

The use of an estimator $\hat{g}_k(\hat{z}_k, \hat{z}_k^*)$ for the Wirtinger gradient g allows CSPSA to optimize functions with unknown parameters, provided that the values of $f(\hat{z}_{k\pm}, \hat{z}_{k\pm}^*)$ can be obtained and the unknown parameters remain constant along the optimization procedure. This is precisely the case of the infidelity. Considering an unknown pure quantum state $|\psi(\tilde{z})\rangle$ of a qudit given by

$$|\psi(\tilde{z})\rangle = \frac{1}{\sqrt{N}} \sum_{i=0}^{d-1} \tilde{z}_i |i\rangle, \tag{6}$$

with $N = \sum_{i=0}^{d-1} |\tilde{z}_i|^2$, and an estimate $|\psi(\hat{z})\rangle$ of $|\psi(\tilde{z})\rangle$ given by

$$|\psi(\hat{z})\rangle = \frac{1}{\sqrt{K}} \sum_{i=0}^{d-1} \hat{z}_i |i\rangle, \tag{7}$$

with $K = \sum_{i=0}^{d-1} |\hat{z}_i|^2$, the infidelity becomes

$$I|\psi(\tilde{z})\rangle, (|\psi(\hat{z})\rangle) = 1 - \frac{|\hat{z} \cdot \tilde{z}^*|^2}{NK}. \tag{8}$$

This function quantifies the deviation of the estimate $|\psi(\hat{z})\rangle$ from the true unknwon state $|\psi(\tilde{z})\rangle$. In this function $\hat{z}$ is a set of complex variables and $\tilde{z}$ plays the role of a set of fixed unknown complex parameters. According to Eq. (4) CSPSA evaluates at each iteration k the infidelity $I(|\psi(\tilde{z})\rangle), |\psi(\hat{z})\rangle$ at $\hat{z} = \hat{z}_{k\pm}$. These values can be obtained by projecting the unknown state $|\psi(\tilde{z})\rangle$ onto a d-dimensional orthonormal base $B_{k\pm} = \{|\psi_{i,k\pm}\rangle\}$ (with $i = 0, \ldots, d - 1$) that contains the state $|\psi(\hat{z}_{k\pm})\rangle$. This procedure generates the detection statistics $n_{i,k\pm}$ that lead to the probability distributions $p_{i,k\pm} = n_{i,k\pm} / \sum_j n_{j,k\pm}$, where $N_{est} = \sum_j n_{j,k\pm}$ is the total number of copies of the unknown states employed in the projective measurements. These probability distributions are employed to obtain the estimates

$$I(|\psi(\tilde{z})\rangle, |\psi(\hat{z}_{k\pm})\rangle) = 1 - p_{0,k\pm} \tag{9}$$

for the infidelity, where we have assumed the convention $|\psi_{0,k\pm}\rangle = |\psi(\hat{z}_{k\pm})\rangle$. These estimates together with Eq. (3) generate the next estimate $|\psi(\hat{z}_{k+1})\rangle$ for the unknown quantum state $|\psi(\tilde{z})\rangle$. The CSPSA method can be understood as a generalization of the Simultaneous perturbation stochastic approach (SPSA)[51,52], a well known stochastic gradient-free optimization method working in the field of the real numbers. SPSA has been proposed[42] and experimentally demonstrated[44] as tomographic method for pure quantum states. In this context, it has been show[30] that CSPSA achieves a higher convergence rate than SPSA.

Algorithm 1 CSPSA-MLE tomographic method

1: Consider an known pure state $|\psi(\tilde{z})\rangle$
2: Set initial guess $\hat{z}_1$, and gain coefficients a, A, s, b and r
3: **for** $k = 1, \ldots, k_{max}$ **do**
4:      Set

$$a_k = \frac{a}{(10k+1+A)^s}, \quad c_k = \frac{b}{(10k+1)^r}.$$

5:      Choose $\Delta_{k,i}$ randomly in the set $\{\pm 1, \pm i\}$.
6:      Calculate $|\psi_{k\pm}\rangle = |\psi(\hat{z}_{k\pm})\rangle$, with $\hat{z}_{k\pm} = \hat{z}_k \pm c_k \Delta_k$.
7:      Choose two bases $B_{k\pm} = \{|\psi_{i,k\pm}\rangle\}$ such as $|\psi_{0,k\pm}\rangle = |\psi_{k\pm}\rangle$.
8:      Estimate experimentally the probability distributions $p_{i,k\pm} = |\langle\psi_{i,k\pm}|\psi(\tilde{z})\rangle|^2$ on a sample of size N_{est}.
9:      Calculate the Infidelities $I(|\psi(\tilde{z})\rangle, |\psi_{k\pm}\rangle) = 1 - p_{0,k\pm}$.
10:     Estimate the gradient as

$$\hat{g}_{k,i} = \frac{I(|\psi(\tilde{z})\rangle, |\psi_{k+}\rangle) - I(|\psi(\tilde{z})\rangle, |\psi_{k-}\rangle)}{2c_k \Delta_{k,i}^*}.$$

11:     Actualize the guess $\hat{z}_{k+1}^0 = \hat{z}_k - a_k \hat{g}_k$.
12:     Refine the guess maximizing the acumulative Likelihood function using $|\psi(z_{k+1}^0)\rangle$ as starting point,

$$\hat{z}_{k+1} = \arg\max_{|\psi\rangle} L_k(|\psi\rangle), \quad \text{s. t.} \quad \langle\psi|\psi\rangle = 1.$$

13: **end for**

CSPSA generates 2 probability distributions at each iteration, that is, a total of $2d$ different probabilities. However, only two of them are employed to estimate the required values of the infidelity. The remaining $2d - 2$ probabilities are not occupied by the algorithm. Thus, CSPSA generates a large amount of accumulated data that is simply discarded. Here, we show that precisely this information can be used to improve the convergence rate of CSPSA in such a way that the estimation accuracy reaches the Gill–Massar lower bound for the estimation of pure quantum states. This is accomplished by resorting to Maximum likelihood estimation, a well known statistical inference method that is extensively employed as a post-processing stage in quantum tomographic methods. MLE is a method in statistical inference aimed at estimating unknown parameters of a population from observed data. The underlying idea is to choose as estimator the maximizer of the probability of obtaining the observed data[53,54]. MLE was introduced in quantum tomography as a post-processing method[31] to obtain physically acceptable quantum states. A quantum system that undergoes a measurement process described by the set of projectors $\{|a_i\rangle\langle a_i|\}$ has a likelihood function $L(\rho)$ given by

$$L(\rho) = \frac{N!}{\Pi_j n_j!} \Pi_i Tr(\rho|a_i\rangle\langle a_i|)^{n_i}, \tag{10}$$

where the detection statistic of each projector is given by the number of counts n_i and the total number of counts is given by $N = \sum_i n_i$. If the quantum state prior to the measurement is ρ, then $L(\rho)$ is the total joint probability of registering data $\{n_i\}$. MLE is defined by the convex optimization problem

$$\arg\max_\rho L(\rho), \, s.t. \, Tr(\rho) = 1, \, \rho \geq 0. \tag{11}$$

At this point we link together CSPSA and MLE. Employing the accumulated data $\{n_{i,m\pm}\}$ between iterations $m = 1$ until $m = k$ we define the accumulated likelihood $L_k(\rho)$ by the expression

$$L_k(\rho) = \Pi_{m=1}^k \Pi_{\lambda=\pm} \Pi_{i=0}^{d-1} Tr(\rho|\psi_{i,m\lambda}\rangle\langle\psi_{i,m\lambda}|)^{n_{i,m\lambda}}, \tag{12}$$

which is maximized in the set of pure states employing as starting guess the estimate $|\psi(\hat{z}_{k+1})\rangle$ provided by CSPSA. The refined estimate provided by MLE is then employed as starting guess for the next iteration with CSPSA. We refer to this procedure as the CSPSA-MLE tomographic method. The main steps of the CSPSA-MLE tomographic method have been summarized as pseudocode in Algorithm 1 above.

Results

For a fixed unknown state, the CSPSA-MLE method exhibits three sources of randomness. Since there is no a priori information about the unknown state, the initial guess is chosen according to a Haar-uniform distribution. At each iteration, the vector Δ_k is also randomly chosen and two measurements are performed on an ensemble of size N_{est}, which leads to finite statistics effects. Thus, each time the CSPSA-MLE method is employed to estimate a fixed unknown state $|\psi(\tilde{z})\rangle$, a different estimate $|\psi(\hat{z})\rangle$ is generated. In this scenario, the accuracy of the estimation procedure for a fixed unknown state $|\psi(\tilde{z})\rangle$ is given by the expectation value $\bar{I}(|\psi(\tilde{z})\rangle)$ of Eq. (1).

To study the performance of the CSPSA-MLE method several Monte Carlo experiments in the regime of a small number of iterations were carried out. A set Ω_d with 2×10^2 pure quantum states $|\psi(\tilde{z})\rangle$ of a single
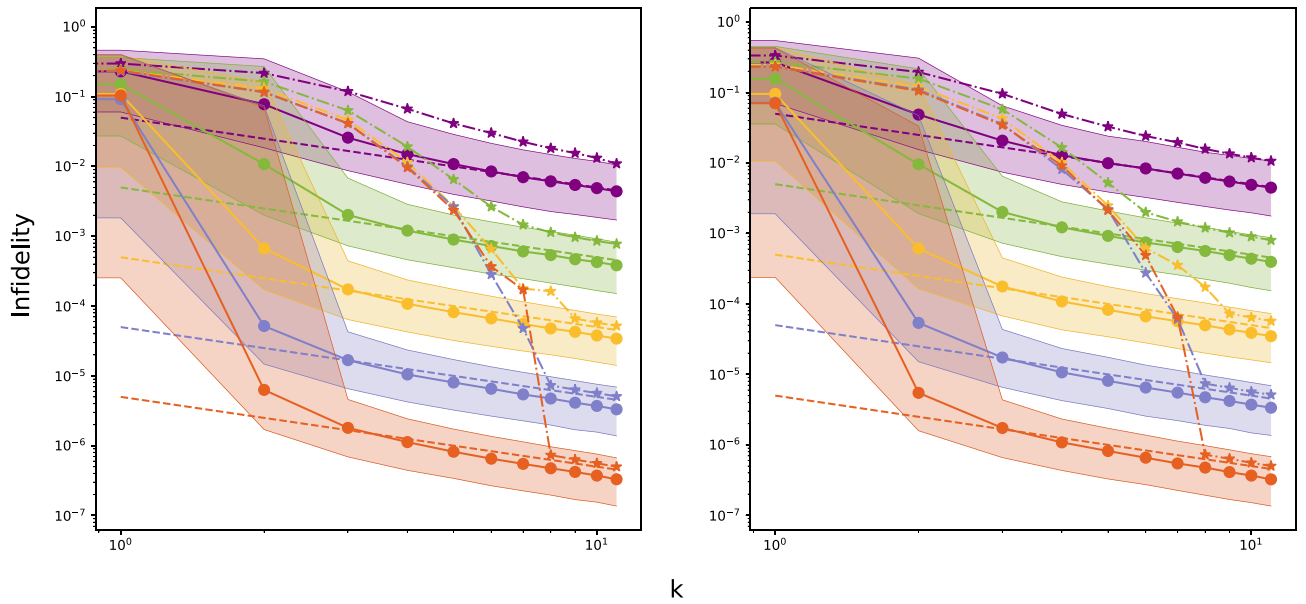
Figure 1. Mean and median infidelity $\bar{I}$ (stars) and $\bar{M}$ (circles), respectively, as functions of the number k of iterations achieved by CSPSA-MLE method for two randomly selected pure states of a single qubit in Ω_2 with with $N_{est} = 10$ (violet), 10^2 (green), 10^3 (yellow), 10^4 (light blue) and 10^5 (purple) from top to bottom. Dashed lines indicate the Gill–Massar lower bound $\bar{I}_p$ for $N = 2N_{est}k$. Shaded areas represent interquartile range. Values of gain coefficients are: $s = 1, r = 1/6, a = 3, A = 0$, and $b = 0.35, 0.3, 0.07, 0.06, 0.03$ for $N_{est} = 10, 10^2, 10^3, 10^4, 10^5$, respectively

d-dimensional quantum system (qudit), uniformly distributed on the unit hypersphere, was generated. Each state in Ω_d was reconstructed via the CSPSA-MLE- method considering a number G of initial guesses, also uniformly distributed on the unit hypersphere, and R independent simulations for each fixed pair of unknown state and initial guess. At each iteration the values of $I(|\psi(\tilde{z})\rangle, |\psi(\hat{z}_{k,\pm})\rangle)$ were estimated considering a multinomial distribution on an ensemble of size N_{est}. For each state in Ω_d; mean, variance, median and interquartile range for the infidelity as functions of the number of iterations k for several values of ensemble size N_{est} were estimated. Similar numerical experiments were performed via CSPSA without MLE for comparison purposes. Since the optimization of the gains is a computational costly problem, we have resorted to the gains $s = 1$ and $r = 0.166$. These lead to a high rate of convergence for CSPSA in the regime of few iterations. The resting coefficients have been set to $A = 0, a = 3$, and $b = 0.35, 0.3, 0.07, 0.06, 0.03$ for $N_{est} = 10, 10^2, 10^3, 10^4, 10^5$, correspondingly.

Figure 1 displays a log-log graphic of the mean infidelity $\bar{I}(|\psi(\tilde{z})\rangle)$ (stars), for two randomly chosen states $|\psi(\tilde{z})\rangle$ in Ω_2, that is, for a single qubit, as a function of the number k of iterations and for several values of the ensemble size N_{est}. This mean infidelity is estimated as

$$\bar{I}(|\psi(\tilde{z})\rangle) = \frac{1}{RG} \sum_{\hat{z}} I(|\psi(\hat{z})\rangle, |\psi(\tilde{z})\rangle), \tag{13}$$

with $G = 500$ and $R = 20$. Within the first 10 iterations the mean infidelity $\bar{I}(|\psi(\tilde{z})\rangle)$ exhibits a fast decrease that is followed by an asymptotic linear behavior. The decrease of $\bar{I}(|\psi(\tilde{z})\rangle)$ becomes more pronounced as N_{est} increases. After 10 iterations the CSPSA-MLE method leads to a mean infidelity $\bar{I}(|\psi(\tilde{z})\rangle)$ approximately equal to $10^{-2}, 7 \times 10^{-4}, 5 \times 10^{-5}, 5 \times 10^{-6}$ and 5×10^{-7}, for increasing N_{est}. For the same states CSPSA without MLE yields after 10 iterations a mean infidelity $\bar{I}(|\psi(\tilde{z})\rangle)$ approximately equal to $2.1 \times 10^{-1}, 5.2 \times 10^{-2}, 4.1 \times 10^{-2}, 4.0 \times 10^{-2}$ and 3.9×10^{-2} (see Fig. 1 in[30]), for increasing N_{est}. Thus, the concatenation of CSPSA to MLE yields a mean infidelity that is 10^{-1} to 10^{-5} times closer to the true minimum than the one provided by CSPSA alone. Let us note that this comparison considers the same type and amount of physical resources for both methods, that is, ensemble size N_{est} for the estimation of the infidelity and total number $2dk$ of measurement outcomes. The estimation via CSPSA can reach similar mean infidelity values to the CSPSA-MLE method but at the expense of more iterations. For instance, with $N_{est} = 10^4$ and after 10 iterations the CSPSA-MLE method delivers a mean infidelity of 5×10^{-6}. A similar value can be achieved via estimation with CSPSA with $N_{est} = 10^4$ after 100 iterations, which represents an increase of one order of magnitude in the total ensemble size N as well as in the total number of measurements. Thus, the concatenation of MLE to CSPSA provides a significative improvement in the rate of convergence and a large reduction of the required physical resources.

After 7 iterations and for $N_{est} = 10, 10^2, 10^3, 10^4, 10^5$ the mean infidelity $\bar{I}(|\psi(\tilde{z})\rangle)$ asymptotically approximates linear behavior. This resembles the Gill–Massar lower limit $\bar{I}_p$ for the average infidelity reached in the estimation of pure states, where now $N = 2kN_{est}$ is the total ensemble size employed after k iterations. This bound imposes a fundamental precision limit on the achievable mean infidelity: no method for the estimation of pure
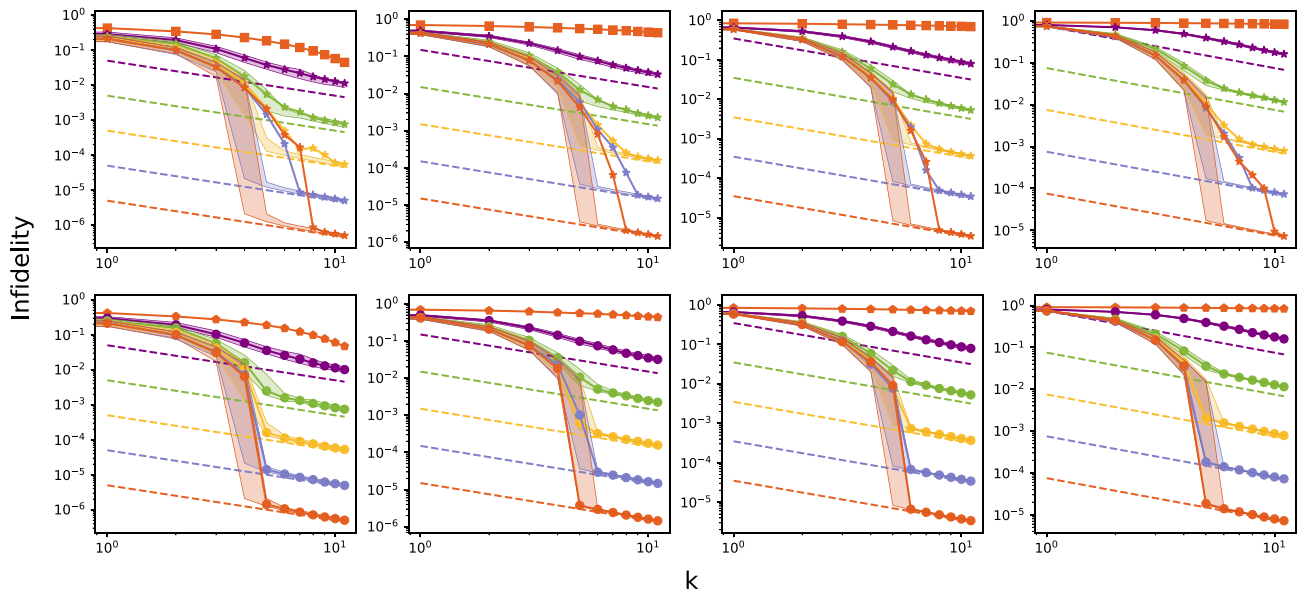
Figure 2. Mean and median infidelity $\bar{\mathbb{I}}$ (stars, upper row) and $\bar{\mathbb{M}}$ (solid circles, lower row) as functions of the number k of iterations obtained via the CSPSA-MLE method for the estimation of single d-dimensional pure quantum states in Ω_d for $d = 2, 4, 8, 16$, from left to right, with $N_{est} = 10$ (violet), 10^2 (green), 10^3 (yellow), 10^4 (light blue) and 10^5 (purple), from top to bottom. Shaded areas describe interquartile range. Dashed lines indicate the value of the Gill–Massar lower bound $\bar{I}_p$ for $N = 2kN_{est}$. Mean $\bar{\mathbb{I}}$ (solid orange squares, upper row) and median $\bar{\mathbb{M}}$ (solid orange rombos, lower row) as functions of the number k of iterations obtained via the CSPSA method for the estimation of single d-dimensional pure quantum states in Ω_d for $d = 2, 4, 8, 16$ with $N_{est} = 10^5$. Gain coefficients as in Fig. 1

states attains a mean infidelity $\bar{I}(|\psi(\tilde{z})\rangle)$ lower than $\bar{I}_p$. Dashed lines in Fig. 1 correspond to $\bar{I}_p$ as function of k for several values of N_{est}. Clearly, the CSPSA-MLE method delivers a mean infidelity of the same order of magnitude than $\bar{I}_p$, albeit slightly higher. However, the gap between $\bar{I}(|\psi(\tilde{z})\rangle)$ and $\bar{I}_p$ tends to close asymptotically as k increases. Thus, the infidelity provided by the CSPSA-MLE method tends to approach the Gill–Massar lower bound $\bar{I}_p$ with a convergence rate that increases with N_{est}.

Figure 1 also displays the median infidelity $\bar{M}(|\psi(\tilde{z})\rangle)$ of $I(|\psi(\hat{z})\rangle, |\psi(\tilde{z})\rangle)$ as a function of k for both randomly chosen states $\psi(\hat{z})\rangle$ and for several values of N_{est}. This exhibits a much faster decrease and an earlier onset of the asymptotic linear behavior. The shaded areas in Fig. 1 correspond to the interquartile range, which is divided into two areas above and below the median infidelity $\bar{M}(|\psi(\tilde{z})\rangle)$. In the lineal regime, the Gill–Massar lower bound $\bar{I}_p$ lays in the upper half of the interquartile range. This indicates that more than 50% of the reconstruction attempts lead to an infidelity lower than $\bar{I}_p$. Nevertheless, the mean infidelity $\bar{I}(|\psi(\tilde{z})\rangle)$ is above the median $\bar{M}(|\psi(\tilde{z})\rangle)$. This points to the existence of a small fraction of realizations with high values of the infidelity, which increases the value of the mean infidelity above the median infidelity. As the number of iterations increases the impact of these realizations on the value of the mean infidelity decreases and mean and median infidelity tend to reach similar values.

The main features exhibited in Fig. 1 are typical, that is, all randomly chosen states in Ω_d display a similar behavior. This is shown in Fig. 2 where mean $\bar{\mathbb{I}}$ (stars), median $\bar{\mathbb{M}}$ (dots) and interquartile range (shaded areas) of the mean infidelity $\bar{I}(|\psi(\tilde{z})\rangle)$ over all states in Ω_d as a function of k for several values of N_{est} are depicted and compared to $\bar{I}_p$ (dashed lines). The mean $\bar{\mathbb{I}}$ (stars) correspond to the expectation value of $\bar{I}(|\psi(\tilde{z})\rangle)$ on the Hilbert space, that is,

$$\bar{\mathbb{I}} = \mathbb{E}[\bar{I}(|\tilde{\psi}\rangle)||\tilde{\psi}\rangle] = \int \bar{I}(|\tilde{\psi}\rangle)f(|\tilde{\psi}\rangle)d\tilde{\psi}, \tag{14}$$

where $f(|\tilde{\psi}\rangle)$ is the probability density function of randomly and uniformly generating the unknown state $|\tilde{\psi}\rangle$. The mean $\bar{\mathbb{I}}$ is estimated as

$$\bar{\mathbb{I}} = \frac{1}{200}\sum_{|\psi(\tilde{z})\rangle \in \Omega} \bar{I}(|\psi(\tilde{z})\rangle). \tag{15}$$

Figure 2 exhibits a fast decrease of $\bar{\mathbb{I}}$ until approximately iteration 7, which is followed by asymptotic linear behavior. The median $\bar{\mathbb{M}}$ shows a similar behavior, although it enters an asymptotic linear behavior faster than the mean $\bar{\mathbb{I}}$ in iteration 5 approximately. As the number of iterations increases mean $\bar{\mathbb{I}}$ and median $\bar{\mathbb{M}}$ overlap almost perfectly in the linear regime. The interquartile range is very narrow and nearly indistinguishable from these two quantities. This indicates that in the linear regime the mean infidelities $\bar{I}(|\psi(\tilde{z})\rangle)$ of the states in Ω_d are concentrated in an extremely narrow interval around the mean $\bar{\mathbb{I}}$ and the median $\bar{\mathbb{M}}$. Thus, all states in Ω_d

are estimated by the CSPSA-MLE method with an accuracy that is very close to $\bar{\bar{\mathbb{I}}}$. Furthermore, the estimation accuracy $\bar{\mathbb{I}}$ provided by the CSPSA-MLE method tends to converge from above to $\bar{I}_p$. Thereby, the CSPSA-MLE method produces a mean infidelity $\bar{\mathbb{I}}$ for any unknown pure state that approaches asymptotically the best possible estimation accuracy of pure states $\bar{I}_p$ allowed by the laws of quantum mechanics. Figure 2 also displays the behavior of $\bar{\mathbb{I}}$ (squares) and median $\bar{\mathbb{M}}$ (rombos) obtained with the use of CSPSA only, that is, without employing MLE to refine the guesses provided by CSPSA, for the case of $N_{est} = 10^5$. Clearly, the CSPSA-MLE tomographic method outperforms the CSPSA tomographic method by at least 5 orders of magnitude. Therefore, the concatenation of MLE to CSPSA plays a key role in improving the accuracy of the estimation and bringing it closer to the lower bound of Gill–Massar.

Once the CSPSA-MLE method enters in a lineal regime, after approximately 10 iterations in the inspected dimensions, delivers an estimation accuracy $\bar{I}(|\tilde{\psi}\rangle)$ close to $\bar{I}_p$. Thereby, we can write

$$N \approx \frac{d-1}{\bar{I}(|\tilde{\psi}\rangle)}, \tag{16}$$

or since $N = 2N_{est}k$ also

$$N_{est}k \approx \frac{d-1}{2\bar{I}(|\tilde{\psi}\rangle)} \quad \text{for k} \geq 10. \tag{17}$$

Thus, in a given dimension d the CSPSA-MLE tomographic method achieves a predefined estimation accuracy with an ensemble size N that can be divided into $2k$ ensembles of size N_{est} each one. Thereby, we can employ a small value of N_{est} and a large number of iterations or a large value of N_{est} and a small number of iterations. The last alternative provides a faster convergence rate. In this case, the estimation of the complex gradient is closer to the exact gradient and the convergence of CSPSA becomes similar to a complex formulation of a deterministic first-order iterative optimization algorithm. However, whether the first or second alternative is more appropriate also depends largely on the characteristics of the experimental platform where the estimation is realized.

We can also compare the mean infidelity $\bar{\mathbb{I}}$ achieved by the CSPSA-MLE method with the Gill–Massar lower bound $\bar{I}_m$ for the mean infidelity achieved in the estimation of full-rank mixed states via separable measurements on the ensemble of equally prepared copies, which is given by $\bar{I}_m = [(d+1)/2]^2\bar{I}_p$. Tomographic methods designed to estimate unknown mixed states cannot achieve a better accuracy than $\bar{I}_m$ as long as they resort to separable measurements on the ensemble of equally prepared copies. This departs quadratically from $\bar{I}_p$ and $\bar{\mathbb{I}}$ as the dimensions increases. Thus, tomographic methods that do not employ the a priori information about the purity of the unknown state cannot estimate pure states with an accuracy better than $\bar{I}_m$. Thereby, the CSPSA-MLE method provides an advantage for pure state estimation over standard quantum tomography, two-stage standard quantum tomographic, and methods such as the ones based on mutually unbiased bases, symmetric informationally complete positive-operator-valued measures, and equidistant states.

Discussion

The accurate estimation of quantum states with a limited ensemble size is a difficult task. In the case of 2-dimensional pure quantum states, the best measurement strategy, that is, the one that leads to the (local unbiased) estimator that saturates the quantum Cramér–Rao bound, is generally a function of the parameters of the unknown state itself. Thereby, the use of the optimal estimation strategy is unfeasible[55]. It is possible, however, to employ an adaptive strategy that approaches the quantum Cramér–Rao bound. This strategy consists of a sequence of measurements where each one is optimal for a given guess of the unknown state[56,57]. The multi-parameter estimation of an unknown d-dimensional quantum state, which is defined by $2d - 2$ independent real numbers, cannot be carried out following a similar strategy since the optimal measurement strategy is unknown.

Instead, we have approached the estimation of unknown d-dimensional quantum states from the optimization of the metric used to characterize the accuracy of the estimation process. The optimization is solved by a method based on the concatenation of CSPSA to MLE. The proposed a method allows for estimating quantum pure states with high accuracy. CSPSA drives a sequence of projective measurements toward the infidelity minimizer. MLE provides at each iteration a refinement of the estimates considering the accumulated data generated by all previous measurements. Monte Carlo experiments for dimensions $d = 2, 4, 8$ and 16 indicate that the mean infidelity for a fixed arbitrary unknown state exhibits a fast decrease within few iterations followed by an asymptotic linear trend. In the linear regime the attained mean infidelity closely approaches the fundamental limit on the accuracy established by the Gill–Massar lower bound for the mean infidelity of the estimation of pure states. Hence, the CSPSA-MLE method surpasses the accuracy of known tomographic methods for pure quantum states. The mean infidelity is also lower than the Gill–Massar lower bound for the infidelity of the estimation of mixed quantum states. Consequently, no tomographic method for mixed states can achieve a mean infidelity lower than the one attained by the CSPSA-MLE method. The median infidelity is also below the Gill–Massar lower bound for pure states. Therefore, more than 50% of the estimation attempts leads to lower infidelities than the Gill–Massar lower bound.

The CSPSA-MLE method exhibits a clear trade-off. The concatenation of CSPSA to MLE increases the rate of convergence of the mean infidelity, which leads to a decrease in the number of iterations and in the ensemble size. However, there is an increase in the computational complexity of the algorithm because in each iteration it is now necessary to solve the optimization problem corresponding to MLE. Recently, optimized methods for MLE in high dimension have been proposed[58,59].

An experimental realization of the CSPSA-MLE method applied to state of a single 2-dimensional quantum system can be carried out with current experimental techniques[44,57] for generating and measuring single-photon

polarization states. The higher dimensional case can be demonstrated by means of experimental setups based on single photons and concatenated spatial light modulators[32,60,61] or via integrated quantum photonics[62,63]. These two experimental platforms offer the possibility of performing electronically controlled adaptive measurements.

Code availability

The source codes that support the results of this study are available from the corresponding author upon reasonable request.

References

1. Wootters, W. K. & Zurek, W. H. A single quantum cannot be cloned. *Nature* **299**, 802 (1982).
2. Ivanovic, I. D. How to differentiate between non-orthogonal states. *Phys. Lett. A* **123**, 257 (1987).
3. Dieks, D. Overlap and distinguishability of quantum states. *Phys. Lett. A* **126**, 303 (1988).
4. Peres, A. How to differentiate between non-orthogonal states. *Phys. Lett. A* **128**, 19 (1988).
5. Von Neumann, J. *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, New Jersey, 1983).
6. Holevo, A. S. *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).
7. Peres, A. *Quantum Theory: Concepts and Methods* (Kluwer, Dordrecht, 1993).
8. Helstrom, C. W. *Quantum Detection and Estimation Theory* Vol. 84 (Academic Press, New York, 1976).
9. Helstrom, C. W. The minimum variance of estimates in quantum signal detection. *IEEE Trans. Inf. Theory* **14**, 234 (1968).
10. Braunstein, S. L. & Caves, C. M. Statistical distance and the geometry of quantum states. *Phys. Rev. Lett.* **72**, 3439. https://doi.org/10.1103/PhysRevLett.72.3439 (1994).
11. Paris, M. & Řeháček, J. (eds) *Quantum State Estimation, Lecture Notes in Physics* Vol. 649 (Springer, Berlin, 2004).
12. DiVincenzo, D. P. The physical implementation of quantum computation. *Fortschr. Phys.* **48**, 771 (2000).
13. Kok, P. *et al.* Linear optical quantum computing with photonic qubits. *Rev. Mod. Phys.* **79**, 135. https://doi.org/10.1103/RevModPhys.79.135 (2007).
14. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, New York, 2010).
15. Giovannetti, V., Lloyd, S. & Maccone, L. Quantum-enhanced measurements: beating the standard quantum limit. *Science* **306**, 1330. https://doi.org/10.1126/science.1104149 (2004).
16. Banaszek, K., Cramer, M. & Gross, D. Focus on quantum tomography. *New J. Phys.* **15**, 125020 (2013).
17. Wootters, W. K. & Fields, B. D. Optimal state-determination by mutually unbiased measurements. *Ann. Phys. (NY)* **191**, 363 (1989).
18. James, D. F. V., Kwiat, P. G., Munro, W. J. & White, A. G. Measurement of qubits. *Phys. Rev. A* **64**, 052312. https://doi.org/10.1103/PhysRevA.64.052312 (2001).
19. Renes, J. M., Blume-Kohout, R., Scott, A. J. & Caves, C. M. Symmetric informationally complete quantum measurements. *J. Math. Phys.* **45**, 2171. https://doi.org/10.1063/1.1737053 (2004).
20. Gross, D., Liu, Y.-K., Flammia, S. T., Becker, S. & Eisert, J. Quantum state tomography via compressed sensing. *Phys. Rev. Lett.* **105**, 150401. https://doi.org/10.1103/PhysRevLett.105.150401 (2010).
21. Cramer, M. *et al.* Efficient quantum state tomography.. *Nat. Commun.* **1**, 149. https://doi.org/10.1038/ncomms1147 (2010).
22. Ahn, D. *et al.* Adaptive compressive tomography with no a priori information. *Phys. Rev. Lett.* **122**, 100404. https://doi.org/10.1103/PhysRevLett.122.100404 (2019).
23. Salazar, R. & Delgado, A. Quantum tomography via unambiguous state discrimination. *Phys. Rev. A* **86**, 012118. https://doi.org/10.1103/PhysRevA.86.012118 (2012).
24. Paiva-Sanchez, C., Burgos-Inostroza, E., Jiménez, O. & Delgado, A. Quantum tomography via equidistant states. *Phys. Rev. A* **82**, 032115. https://doi.org/10.1103/PhysRevA.82.032115 (2010).
25. Martínez, D., Solís-Prosser, M. A., Cañas, G., Jiménez, O., Delgado, A. & G. Lima. Experimental quantum tomography assisted by multiply symmetric states in higher dimensions *Phys. Rev. A* **99**, 012336 (2019). https://doi.org/10.1103/PhysRevA.99.012336
26. Mahler, D. H. *et al.* Adaptive quantum state tomography improves accuracy quadratically. *Phys. Rev. Lett.* **111**, 183601. https://doi.org/10.1103/PhysRevLett.111.183601 (2013).
27. Hou, Z., Zhu, H., Xiang, G., Li, C.-F. & Guo, G.-C. Achieving quantum precision limit in adaptive qubit state tomography. *npj Quantum Inf.* **2**, 16001 (2016). https://www.nature.com/articles/npjqi20161
28. Pereira, L., Zambrano, L., Cortés-Vega, J., Niklitschek, S. & Delgado, A. Adaptive quantum tomography in high dimensions. *Phys. Rev. A* **98**, 012339. https://doi.org/10.1103/PhysRevA.98.012339 (2018).
29. Gill, R. D. & Massar, S. State estimation for large ensembles. *Phys. Rev. A* **61**, 042312. https://doi.org/10.1103/PhysRevA.61.042312 (2000).
30. Utreras-Alarcón, A., Rivera-Tapia, M., Niklitschek, S. & Delgado, A. Stochastic optimization on complex variables and pure-state quantum tomography. *Sci. Rep.* **9**, 16143 (2019).
31. Hradil, Z. Quantum-state estimation. *Phys. Rev. A* **55**, R1561–R1564. https://doi.org/10.1103/PhysRevA.55.R1561 (1997).
32. Goyeneche, D. *et al.* Five measurement bases determine pure quantum states on any dimension. *Phys. Rev. Lett.* **115**, 090401. https://doi.org/10.1103/PhysRevLett.115.090401 (2015).
33. Carmeli, C., Heinosaari, T., Kech, M., Schultz, J. & Toigo, A. Stable pure state quantum tomography from five orthonormal bases. *EPL* **115**, 30001. https://doi.org/10.1209/0295-5075/115/30001/pdf (2016).
34. Sosa-Martinez, H. *et al.* Experimental study of optimal measurements for quantum state tomography. *Phys. Rev. Lett.* **119**, 150401. https://doi.org/10.1103/PhysRevLett.119.150401 (2017).
35. Zambrano, L., Pereira, L. & Delgado, A. Improved estimation accuracy of the 5-bases-based tomographic method. *Phys. Rev. A* **100**, 022340. https://doi.org/10.1103/PhysRevA.100.022340 (2019).
36. Paris, M. G. A. Quantum estimation for quantum technology. *Int. J. Quantum. Inf.* **07**, 125. https://doi.org/10.1142/S0219749909004839 (2009).
37. Helstrom, C. W. Minimum mean-squared error of estimates in quantum statistics. *Phys. Lett. A* **25**, 101 (1967).
38. Fuchs, C. & van de Graaf, J. Cryptographic distinguishability measures for quantum-mechanical states. *IEEE Trans. Inf. Theory* **45**, 1216. https://doi.org/10.1109/18.761271 (1999).
39. Wootters, W. K. Statistical distance and Hilbert space. *Phys. Rev. D* **23**, 357. https://doi.org/10.1103/PhysRevD.23.357 (1981).
40. Zhu, H. *Quantum State Estimation and Symmetric Informationally Complete POMs*. Ph.D. thesis, National University of Singapore (2012).
41. Hayashi, M. in *Quantum Communication, Computing and Measurement* (eds Hirota, O. *et al.*) 99–108 (Plenum, Berlin, 1997).
42. Ferrie, C. Self-guided quantum tomography. *Phys. Rev. Lett.* **113**, 190404. https://doi.org/10.1103/PhysRevLett.113.190404 (2014).

43. Granade, C., Ferrie, C. & Flammia, S. T. Practical adaptive quantum tomography. *New J. Phys.* **19**, 113017. https://doi.org/10.1088/1367-2630/aa8fe6/meta (2017).
44. Chapman, R. J., Ferrie, C. & Peruzzo, A. Experimental demonstration of self-guided quantum tomography. *Phys. Rev. Lett.* **117**, 040402. https://doi.org/10.1103/PhysRevLett.117.040402 (2016).
45. Hou, Z. *et al.* Experimental realization of self-guided quantum process tomography. *Phys. Rev. A* **101**, 022317. https://doi.org/10.1103/PhysRevA.101.022317 (2020).
46. Sorber, L., van Barel, M. & de Lathauwer, L. Unconstrained optimization of real functions in complex variables. *SIAM J. Optim.* **22**, 879–898. https://doi.org/10.1137/110832124 (2012).
47. Wirtinger, W. Zur formalen Theorie der Funktionen von mehr komplexen Veränderlichen. *Math. Ann.* **97**, 357. https://doi.org/10.1007/BF01447872 (1927).
48. Brandwood, D. H. A complex gradient operator and its application in adaptive array theory. *Proc. IEE-H* **130**, 11–16 (1983).
49. Nehari, Z. *Introduction to Complex Analysis* (Allyn & Bacon, Boston, 1961).
50. Remmert, R. *Theory of Complex Functions* (Springer, New York, 1991).
51. Spall, J. C. *Introduction to Stochastic Search and Optimization* (Wiley, Hoboken, 2003).
52. Spall, J. C. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Trans. Autom. Control* **37**, 332–341 (1992).
53. Cox, D. R. *Principles of Statistical Inference* (Cambridge University Press, New York, 2006).
54. Lehmann, E. & Casella, G. *Theory of Point Estimation* (Springer, New York, 1998).
55. Cochran, W. G. Experiments for Nonlinear Functions. *J. Am. Stat. Assoc.* **68**, 771. https://doi.org/10.2307/2284499 (1973).
56. Nagaoka, H. in *Asymptotic Theory of Quantum Statistical Inference*, edited by M. Hayashi (World Scientific, Singapore, 2005).
57. Okamoto, R. *et al.* Experimental demonstration of adaptive quantum state estimation. *Phys. Rev. Lett.* **109**, 130404. https://doi.org/10.1103/PhysRevLett.109.130404 (2012).
58. Shang, J., Zhang, Z. & Ng, H. K. Superfast maximum-likelihood reconstruction for quantum tomography. *Phys. Rev. A* **95**, 062336. https://doi.org/10.1103/PhysRevA.95.062336 (2017).
59. Bolduc, E., Knee, G. C., Gauger, E. M. & Leach, J. Projected gradient descent algorithms for quantum state tomography. *npj Quantum Inf.* **3**, 44 (2017).
60. Solís-Prosser, M. A., Fernandes, M. F., Jiménez, O., Delgado, A. & Neves, L. Experimental minimum-error quantum-state discrimination in high dimensions. *Phys. Rev. Lett.* **118**, 100501. https://doi.org/10.1103/PhysRevLett.118.100501 (2017).
61. Martínez, D. *et al.* Experimental quantum tomography assisted by multiply symmetric states in higher dimensions. *Phys. Rev. A* **99**, 012336. https://doi.org/10.1103/PhysRevA.99.012336 (2019).
62. Wang, J., Sciarrino, F., Laing, A. & Thompson, M. G. Integrated photonic quantum technologies. *Nat. Photon.* **14**, 273. https://doi.org/10.1038/s41566-019-0532-1 (2019).
63. Cariñe, J. *et al.* Multi-core fiber integrated multi-port beam splitters for quantum information processing. *Optica* **7**, 542 (2020).

Acknowledgements

Author Contributions

A. D. conceived the stochastic optimization method on the field of complex numbers and L. Z. proposed the concatenation of statistical inference to CSPSA. L. Z. and L. P. performed numerical experiments. A. D. and S. N. analyzed the results of numerical experiments. A. D. wrote the article with input from all authors. All authors reviewed and contributed to the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to A.D.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.