Efficient Bayesian credible-region certification for quantum-state tomography

C. Oh, Y. S. Teo, and H. Jeong

1 Center for Macroscopic Quantum Control, Department of Physics and Astronomy, Seoul National University, 08826 Seoul, South Korea
2 Frontier Physics Research Division, Department of Physics and Astronomy, Seoul National University, 08826 Seoul, South Korea

Standard Bayesian credible-region theory for constructing an error region on the unique estimator of an unknown state in general quantum-state tomography to calculate its size and credibility relies on heavy Monte Carlo sampling of the state space followed by sample rejection. This conventional method typically gives negligible yield for very small error regions originating from large datasets. We propose an operational reformulated theory to compute both size and credibility from region-average quantities that in principle convey information about behavior of these two properties as the credible-region changes. We next suggest the accelerated hit-and-run Monte Carlo sampling, customized to the construction of Bayesian error-regions, to efficiently compute region-average quantities, and provide its complexity estimates for quantum states. Finally by understanding size as the region-average distance between two states in the region (measured for instance with either the Hilbert-Schmidt, trace-class or Bures distance), we derive approximation formulas to analytically estimate both distance-induced size and credibility under the pseudo-Bloch parametrization without resorting to any Monte Carlo computation.

I. INTRODUCTION

All physical-quantity estimates obtained from collected data should be accompanied by “error-bars” to accurately convey all properties of the physical system of interest. This applies to quantum-state tomography [1–5], which is an important preliminary step for implementing all quantum cryptography and computation protocols [6–8] reliably.

Bootstrapping procedures [9, 10] are amongst some of the most widely-used techniques for assigning “error-bars” to reconstructed quantum states. Recently, it was pointed out in [11] that such assignments lack rigorous statistical foundations and may produce “error-bars” that are too small for reliable conclusions. The rather more justified approach falls under the study of hypothesis testing [12]. Two grand schools of thought exist for this purpose. In the context of quantum-state reconstruction of an unknown state \( \rho \), one may treat \( \rho \) as “absolute” (the frequentist school) and attempt to extract this knowledge from collected data. This suggests the constructions of confidence regions [13–15], which are error-regions for the state estimator \( \hat{\rho} \) from all plausible datasets, including those unmeasured in the experiment. An accurate \( \hat{\rho} \) for \( \rho \) would then entail a collection of typically small confidence regions with high probability that \( \rho \) for each plausible dataset lies in the corresponding region.

Given that only one dataset (the measured one) is really available to the observer, we shall focus on the apt Bayesian school of thought that instead regards this dataset as “absolute” and constructs credible regions [16, 17] as the error regions for \( \hat{\rho} \) beginning with some prior distribution \( p(\rho) \) of \( \rho \). A fairly accurate estimator \( \hat{\rho} \) for some unknown \( \rho \) naturally implies a credible region (generated from the measured dataset) of a small size with a large probability that this \( \rho \) is inside the region—a high credibility [18]. In order to obtain a reasonably small error region (be it that of credible- or confidence-type), one may either resort to adaptive strategies [19] and optimize additional properties of the region, or simply increase the dataset collected in quantum tomography.

The complicated quantum state-space boundary [20, 21] renders any analytical attempt at calculating size and credibility for any credible region futile, leaving numerical Monte Carlo (MC) methods as the only viable option. As the size of the credible region is defined as its volume fraction with the quantum state space, one needs an extremely large sample of the state space to finally end up having a reasonable sample for the region—the strategy of MC [22, 23] followed by another rejection sampling. Despite the optimistic advantages that some of these MC schemes may have in generating samples of arbitrary distribution, one deleterious issue for such a MC-rejection strategy becomes apparent when the dataset is large, which is the common situation in any tomography experiment. The resulting credible region eventually becomes too tiny relative to the quantum state space for any MC method to generate a sufficient sample so that rejection sampling can produce any effective yield to properly compute the size and credibility. Therefore, an alternative strategy for Bayesian error certification is in order.

In our reformulated credible-region theory, we prove the central lemma stating that the size (and credibility) of any credible region are related to a class of region-average quantities through a first-order differential equation that is solvable numerically. We then propose the accelerated hit-and-run algorithm to efficiently compute region-average quantities and estimate its complexity from the geometry of the quantum state space. The region-average formalism encourages the creative perspective that size may equivalently be regarded as the average relative distance between two points in the region. This region-average distance may be induced by any of the common measures used in quantum information, and we shall explicitly consider the Hilbert-Schmidt, trace-class, and Bures distance measures as popular examples. It turns out that this perspective gives rise to closed-form analytical approximation formulas for rapid approximated Bayesian error certification with large datasets with no MC methods necessary.

This article is organized as follows: A preliminary introduction to the basic notions of standard Bayesian credible-
region theory shall ensue in Sec. II, and the stage for discussions with large data is set in Sec. III. Next, we present our region-average theory for size and credibility that works for any kind of data and prior in Sec. IV. Afterwards, we describe how region-average quantities can be numerically computed and estimate the computational complexities in Sec. V. Section VI then proceeds to redefine size in terms of region-average distances induced by all three measures. Finally, for fast analytical Bayesian error estimates, we derive asymptotic formulas for all important region-average quantities in Sec. VII based on the perspective of distance-induced size. All results are then demonstrated and discussed in Sec. VIII with explicit examples in quantum tomography.

II. STANDARD BAYESIAN CREDIBLE-REGION THEORY

Before a quantum-state tomography experiment commences, the observer might have some (justifiable) preconception about the unknown quantum state $\rho \geq 0$ (tr$(\rho) = 1$) of Hilbert-space dimension $D$. Such preconception is usually not uniquely privileged, and therefore weighted with some prior probability distribution $p(\rho)$. After the experiment, the observer collects a set of data $\mathcal{D}$ that are informationally complete (IC) such that a unique estimator $\hat{\rho}$ for $\rho$ is acquired. In quantum theory, the measurements are modeled as a probability-operator measurement (POM) consisting a set of $M$ positive operators $\Pi_j \geq 0$ that sum to the identity. Associated to every such experiment is the likelihood function $L = L(\mathcal{D}|\hat{\rho})$, with which the observer obtains a posterior probability distribution (knowledge after-the-fact) that is a function of $L$.

It was formerly established in Ref. [16] that for this measured dataset $\mathcal{D}$, if $\hat{\rho}$ is taken to be the estimator that maximizes $L$—the maximum-likelihood (ML) estimator—, then a Bayesian credible region (CR) $\mathcal{R}$ can be constructed around $\hat{\rho}_{\text{ML}}$, which turns out to have a constant likelihood boundary $\partial \mathcal{R}$ within the quantum state space $\mathcal{R}_0$. For this CR, which is a subregion of $\mathcal{R}_0$, we can specify its size and credibility, the latter which is the probability that $\rho \in \mathcal{R}$. Such a region is optimal in the sense that it gives the largest credibility for a given size, or equivalently possesses the smallest size for a given credibility.

In this article, we shall be interested in reconstructing the $(d = D^2 - 1)$-dimensional real vectorial parameter $r \leftrightarrow \rho$ that characterizes $\rho$. More technically, this equivalent parametrization is achieved with a Hermitian operator basis $\{1/\sqrt{D}, \Omega_j\}_{j=1}^D$ that contains $d$ trace-orthonormal traceless operators $\Omega_j$ (tr$(\Omega_j \Omega_k) = \delta_{jk}$), by which $r = \text{tr} (\rho \Omega_j)$ is defined from the column $\Omega_j$ of $\Omega_j$. Formally, in terms of the multivariate parameter $r$, the size and credibility of $\mathcal{R} = \mathcal{R}_\lambda$ for some $0 \leq \lambda \leq 1$ are respectively given by Ref. [16]

$$S_\lambda = \int_{\mathcal{R}_\lambda} (\text{d}r') = \int_{\mathcal{R}_0} (\text{d}r') \eta(L - \lambda L_{\text{max}}),$$

$$C_\lambda = \frac{1}{L(\mathcal{D})} \int_{\mathcal{R}_\lambda} (\text{d}r') L = \frac{1}{L(\mathcal{D})} \int_{\mathcal{R}_0} (\text{d}r') \eta(L - \lambda L_{\text{max}}) L,$$

where the volume measure $(\text{d}r)$ incorporates the prior distribution $p(r)$ for $r$, $\eta$ is the Heaviside function, $L(\mathcal{D}) = \int_{\mathcal{R}_0} (\text{d}r') L(\mathcal{D}|r')$. The important variable $0 \leq \lambda \leq 1$ specifies the shape and size of $\mathcal{R}_\lambda$, from which the limits $\mathcal{R}_{\lambda = 0} = \mathcal{R}_0$ and $\mathcal{R}_{\lambda = 1} = \{\hat{\rho}_{\text{ML}}\}$ are immediate. Here, we note that the probability parametrization was adopted in Ref. [16]. Upon the condition that each datum measurement, corresponding to an outcome $\Pi_j$, the inherent statistics of $\mathcal{D}$ is therefore multinomial and the log-likelihood reads $\log L = \sum_j n_j \log p_j$, with the collected relative frequencies $\sum_j n_j = N$ that make up $N$ measured data copies, and $p_j = \text{tr} (\rho_j \Pi_j)$ for any state $\rho_j$.

We can gain a clear physical picture of both size and credibility: they respectively quantify the prior and posterior content $\mathcal{R}$, hence the symbol $S_\lambda$ for the former. Owing to the dual nature of size $S_\lambda$ and credibility $C_\lambda$, it is shown Ref. [16] that

$$C_\lambda = \frac{\lambda S_\lambda + \int_0^1 \text{d} \lambda' S_{\lambda'}}{\int_0^1 \text{d} \lambda' S_{\lambda'}}.$$

(2)

Put differently, $C_\lambda$ may be straightforwardly computable through single-parameter integrations in $\lambda$ so long as $S_\lambda$ is known up to some arbitrary constant multiple.

Nonetheless, the complicated boundary $\partial \mathcal{R}_0$ of the quantum state space makes the computation of $S_\lambda$ extremely challenging even numerically. The innate definition of $S_\lambda$, namely the volume fraction of $\mathcal{R}_\lambda$ to the state space $\mathcal{R}_0$, requires, first, the generation of a sufficiently large sample of $\mathcal{R}_0$, followed by the rejection of all its sampled points that lie outside $\mathcal{R}_\lambda$ for any $\lambda$. There exist various MC methods to sample $\mathcal{R}_0$ [22]. Ultimately, MC-rejection strategy exhibits one major disadvantage: in the limit of large data sample ($N \gg 1$), $\mathcal{R}_\lambda$ would become so small relative to $\mathcal{R}_0$ that the MC-rejection strategy needs a sufficiently large number of random MC sample points from $\mathcal{R}_0$ to produce any useful yield. The scaling of MC sample size needed to maintain a fixed yield, which was estimated to be $O(N^{d/2})$ Ref. [18], thus outgrows the feasible computational yield-rate very quickly. The bottom-line: a much
III. THE LARGE-DATA CONDITION

Before presenting an alternative operational theory, unless otherwise stated, we shall consider $N > 1$ as the putative limit in pragmatic tomography experiments. We emphasize here that $N$ only has to be sufficiently large for the statistical central limit theorem to dictate a Gaussian form for $L$. In this limit, there can only be one of two cases: either $\mathcal{R}$ is completely inside $\mathcal{R}_0$ that contains a full-rank $\hat{\rho}_{\text{ML}}$ (Case A) or partially truncated by the state-space boundary $\partial \mathcal{R}_0$ that houses a rank-deficient $\hat{\rho}_{\text{ML}}$ (Case B) (see Fig. 1).

Case A arises when the unknown state $\rho$ is away from $\partial \mathcal{R}_0$, so that a sufficiently large $N$ would produce truncated regions for $\lambda$, values corresponding to desirably large $C_\lambda < 1$. This case offers a simple geometrical description for $\mathcal{R}$. Upon invoking the Taylor expansion of

$$\log L(\mathcal{D}|r') \approx \log L_{\text{max}} - \frac{1}{2} (r' - \hat{r}_{\text{ML}}) \cdot F_{\text{ML}} \cdot (r' - \hat{r}_{\text{ML}})$$

about the interior $\hat{r}_{\text{ML}}$ up to the second order, we find that the likelihood $L$ is essentially a Gaussian function centered at $\hat{r}_{\text{ML}}$ of height $L_{\text{max}}$, with its covariance profile shaped by $F_{\text{ML}}$, that is the Fisher information evaluated at $\hat{r}_{\text{ML}}$. The CRs $\mathcal{R}_\lambda$ that go with this Gaussian likelihood are, hence, simple hyperellipsoids $\mathcal{E}_\lambda$ described by the inequality $(r' - \hat{r}_{\text{ML}}) \cdot F_{\text{ML}} \cdot (r' - \hat{r}_{\text{ML}}) \leq -2 \log \lambda$.

If $r$ is located in $\partial \mathcal{R}_0$, then as $N$ increases, the ML estimator $\hat{r}_{\text{ML}}$ would eventually approach $r$ and there is a high probability that $\hat{r}_{\text{ML}} \in \partial \mathcal{R}_0$ before this happens. For sufficiently large $N$, we have Case B where $\partial \mathcal{R} \cap \partial \mathcal{R}_0$ is not disjointed and falls on the side of $\hat{r}_{\text{ML}}$. To asymptotically cope with such a situation, we may again expand

$$\log L(\mathcal{D}|r') \approx \log L_{\text{max}} + (r' - \hat{r}_{\text{ML}}) \cdot g_{\text{ML}} - \frac{1}{2} (r' - \hat{r}_{\text{ML}}) \cdot F_{\text{ML}} \cdot (r' - \hat{r}_{\text{ML}})$$

about the boundary $\hat{r}_{\text{ML}}$, where this time $L$ is a Gaussian function centered at $r_c = \hat{r}_{\text{ML}} + F_{\text{ML}}^{-1} \cdot g_{\text{ML}}$, with $g_{\text{ML}} = \partial \log L(\mathcal{D}|r')/\partial r|_{r=\hat{r}_{\text{ML}}}$, and possesses a height $L_{\text{max}} = L_{\text{max}} \exp(g_{\text{ML}} \cdot F_{\text{ML}}^{-1} \cdot g_{\text{ML}}/2) > L_{\text{max}}$. The covariance profile of this Gaussian function is still governed by $F_{\text{ML}}$, which produces hyperellipsoids $\mathcal{E}_\lambda$ described according to $(r' - r_c) \cdot F_{\text{ML}} \cdot (r' - r_c) \leq -2 \log \lambda$ for an “effective $\lambda$” $\lambda'$ defined by $2 \log (\lambda/\lambda') = g_{\text{ML}}^T \cdot F_{\text{ML}}^{-1} \cdot g_{\text{ML}}$. The CR $\mathcal{R}_\lambda$ is then asymptotically $\mathcal{E}_{\lambda'} \cap \partial \mathcal{R}_0$.

We point out that there is an intermediate case in which $\mathcal{R}_\lambda = \mathcal{E}_\lambda$, centered at $\hat{r}_{\text{ML}} \notin \partial \mathcal{R}_0$, is truncated by $\partial \mathcal{R}_0$. Such a situation can happen when $N$ is not sufficiently large, and tends to either Case A or B as $N$ grows. On a separate note, Ref. [18, 19] explicitly studies also this intermediate case.

IV. REGION-AVERAGE THEORY FOR SIZE AND CREDIBILITY

Suppose we have a CR $\mathcal{R}_\lambda$, with which we define the average quantity

$$u_{\lambda} = \frac{\int_{\mathcal{R}_\lambda} (d r') q_{\lambda}(r')}{\int_{\mathcal{R}_\lambda} (d r')} = \frac{1}{K_{\text{supp}}} \sum_{l=1}^{K_{\text{supp}}} q_{\lambda,l}$$

for some function $q_{\lambda}$, which is approximately equivalent to the discrete-sum average of $q_{\lambda,l}$ values over a sufficiently large number $K_{\text{supp}}$ of region points. If we probe the response of $u_{\lambda}$ with an incremental change $\lambda \rightarrow \lambda + \delta \lambda$ in $\lambda$ as in Fig. 2, the result is the total change

$$\delta u_{\lambda} = \left( \frac{1}{\delta \lambda} - \frac{\delta S_{\lambda}}{S_{\lambda}^2} \right) \int_{\mathcal{R}_\lambda} (d r') q_{\lambda} - \frac{\delta S_{\lambda}}{S_{\lambda}} \int_{\mathcal{R}_\lambda} (d r') q_{\lambda}$$

after limiting all small changes to the first order, which reveals that a small increment $\delta \lambda$ can be explained by a change $\delta \lambda$ in size that is accompanied by the (in)exclusion of the annular sum $\int_{\mathcal{R}_{\lambda} \setminus \mathcal{R}_{\lambda + \delta \lambda}} (d r') q_{\lambda}$. Put simply, tracking the change in $u_{\lambda}$ allows us to infer how much $S_{\lambda}$ has changed.

To better utilize this intuition, we first take the derivative of $u_{\lambda} S_{\lambda}$, which gives

$$\frac{\partial u_{\lambda} S_{\lambda}}{\partial \lambda} = -L_{\text{max}} \int_{\mathcal{R}_0} (d r') \left( \frac{\partial}{\partial \lambda} \mathcal{L}(L - \lambda L_{\text{max}}) \right) q_{\lambda}(r')$$

and

$$+ \int_{\mathcal{R}_\lambda} (d r') \frac{\partial q_{\lambda}(r')}{\partial \lambda}$$

after invoking the derivative identity $d \eta(x)/dx = \delta(x)$ between $\eta(x)$ and the Dirac delta function $\delta(x)$. Next, we impose the following two functional properties $q_{\lambda}(r' \in \partial \mathcal{R}_\lambda \cap \text{int}(\mathcal{R}_0)) = 1$ and $\partial q_{\lambda}/\partial \lambda = f(\lambda) q_{\lambda}$ for $q_{\lambda}$, where $f(\lambda)$ is some arbitrary function of $\lambda$ only. These properties simplify Eq. (7) to

$$\frac{\partial}{\partial \lambda} [(u_{\lambda} - 1) S_{\lambda}] = f(\lambda) u_{\lambda} S_{\lambda}.$$
region-average $u_3$. With the initial condition $s_{3=0} = 1$, the entire functional form of $s_3$ can then be recovered with Eq. (8). This completes the constructive proof of our so-called

Region-average computation (RAC) lemma: For any prior $(d\rho')$ and measurement data $\mathbb{D}$, both $S_3$ and $C_3$ are inferable from $u_3$ defined in Eq. (5), with $q_3$ obeying the relations $q_3 (r' \in \partial B_3 \cap \text{int}(\mathbb{D}_0)) = 1$ and $\partial q_3 / \partial \lambda = f(\lambda) q_3$.

To proceed, we first perform the substitution $y_\lambda = (u_\lambda - 1) S_\lambda$ to yield another differential equation

$$\frac{\partial y_\lambda}{\partial \lambda} = f(\lambda) \frac{u_\lambda}{u_\lambda - 1} y_\lambda. \quad (9)$$

The solution to $y_\lambda$ can then be obtained numerically through Euler’s method [24], which is the iteration

$$y_{\lambda_{j+1}} = y_{\lambda_j} + f(\lambda_j) \frac{u_{\lambda_j}}{u_{\lambda_j} - 1} y_{\lambda_j} \quad (10)$$

for a sequence of discretized $\lambda \rightarrow \lambda_j$ values. For practically straightforward computation of $u_3$ without running into underflow problems, we may choose $q_3 = \log L / \log (\lambda L_{\text{max}})$, where the corresponding $f(\lambda) = -1 / (\lambda \log (\lambda L_{\text{max}}))$.

V. REGION-AVERAGE NUMERICAL COMPUTATION

A. The hit-and-run algorithm

The hit-and-run algorithm is a direct convex-body MC sampling scheme that generates random sample points in the body according to some predefined distribution. This algorithm is thus suited for sampling $B_\lambda$ according to some prior distribution $\rho(\mathbf{r})$ for the unknown $\mathbf{r}$.

The sampling principles behind an efficient hit-and-run computation begin with defining the smallest possible convex set $B \supseteq B_\lambda$ that houses $B_\lambda$ and has an easy-to-access geometry. Starting from a known point in $B_\lambda$, say the ML estimator $\hat{F}_{\text{ML}}$, a random line segment passing through this point is generated, with its endpoints fixed at $\partial B$ that are quickly computable because of its simple geometry. Following which, sampling commences by repeatedly picking a random point along this segment until it lies in $B_\lambda$. This point is next taken to be the new reference point through which another line segment is generated to find a new random point in $B_\lambda$, until a set of $K_{\text{smp}}$ points is gathered.

We can make use of the straightforward hyperellipsoidal characteristics inherent from the central limit theorem to construct $B$. For Case A, where $B_\lambda = \mathbb{E}_3$, $B$ can just be taken to be $\mathbb{E}_3$ characterized by $\hat{F}_{\text{ML}} = F_{\text{ML}} / (-2 \log \lambda)$ from the earlier discussions in Sec. III. We now turn to the more interesting and practically ubiquitous Case B, where the large-$N$ arguments of Sec. III imply that we may fix $B = \mathbb{E}_\lambda$, the profile of which is governed by $\hat{F}_{\text{ML}} / (-2 \log \lambda)$.

Accelerated hit-and-run for uniform sampling of $B_\lambda$

Beginning with $k = 1$ and $\mathbf{r}_{\text{ref}} = \hat{F}_{\text{ML}}$, of $N \gg 1$:

1. Generate a random line segment characterized by $\mathbf{y} = \mathbf{r}_{\text{ref}} + \mu \mathbf{e}_v$, where $\mathbf{e}_v = v / |v|$ and $v$ follows the standard Gaussian distribution (mean 0 and variance 1 for each column entry). Its endpoints are parametrized by $\mu = [-b \pm \sqrt{b^2 - a(c-1)}/a$, where $\Delta = r_{\text{ref}} - r_c, a = e_v^T A e_v, b = \Delta^T A e_v, c = \Delta^T A \Delta, A = \hat{F}_{\text{ML}}$, or $\hat{F}_{\text{ML}}$.  

2. Define $\beta_1 \equiv \mu_{\text{min}} = \min \{\mu_+, \mu_-\}$ and $\beta_2 \equiv \mu_{\text{max}} = \max \{\mu_+, \mu_-\}$.

3. Pick a random number $\beta_1 \leq \beta \leq \beta_2$ uniformly and obtain $\mathbf{r}_{\text{test}} = \mathbf{r}_{\text{ref}} + \beta \mathbf{e}_v$.

4. Check whether $\rho_{\text{test}} + \mathbf{r}_{\text{test}}$ is positive via the Cholesky decomposition.

- If so, define $\mathbf{r}_{\text{ref}} = \mathbf{r}_{\text{test}}$, raise $k$ by 1, and go to step 1.

- If not, set $\beta_1 = \beta$ if $\beta < 0$ or $\beta_2 = \beta$ if $\beta > 0$, and repeat steps 3 and 4.

5. End routine if $k > K_{\text{smp}}$, the total number of sample points desired.

To further speed up the algorithm for Case B, one can assign $B$ to be the hyperellipsoidal cap composed by a hyperplane that is tangent to the isoGaussian level curve of $\mathbb{E}_\lambda$ at $\hat{F}_{\text{ML}}$ and the part of $\mathbb{E}_\lambda$ below it (refer to Sec. VII). Numerical experience shows that this speed up is negligible in the presence of the endpoint adaptation mechanism of accelerated hit-and-run.
the longest one of its edges that is almost flat (Type II) in all its dimensions, with least one of its dimensions (Type I) in whichever orientation, or (c) on one of its edges that is almost flat (Type II) in all its dimensions, with the longest $\delta_1^+$-axis oriented along the flat surface.

B. Numerical complexity estimations

After suppressing dependences on logarithmic factors and error parameters, it was argued that the number of hit-and-run steps needed to gather enough sample points and form an ensemble described by $p(r)$ in hit-and-run is $O(D^2 R_{out}^2 / R_{in}^2) = O(D^4 R_{out}^2 / R_{in}^3)$ [28] in the limit $D \gg 2$, where $R_{out}$ is the radius of the smallest outer sphere that contains $\mathcal{R}$, and $R_{in}$ is that of the largest inner sphere that can be inscribed in $\mathcal{R}$. Together with the floating-point-operations complexity $O(D^3)$ in a typical Cholesky decomposition algorithm [27], we have an estimate for the complexity $\text{cmpl} = O(D^4 R_{out}^2 / R_{in}^3)$ for the entire hit-and-run scheme.

The treatment of Case A is straightforward as we have the complete information about $\mathcal{R}_0$ in the large-$N$ limit. If we denote $\sigma_>$ and $\sigma_<$ to respectively be the largest and smallest eigenvalue of $\tilde{F}_{ML}^{-1/2}$, then the corresponding outer and inner radii are $R_{out} = \sigma_>$ and $R_{in} = \sigma_<$ [see Fig. 3(a)], so that $\text{cmpl}_A = O(D^4 \text{cond}(\tilde{F}_{ML}))$.

The analysis for Case B requires extra care given the complicated state-space boundary $\partial \mathcal{R}_0$. While complete and precise details of $\mathcal{R}_0$ are absent so far, from [21], we know that in the Euclidean space, the largest inner sphere inscribable in $\mathcal{R}_0$ has a radius that approaches $1/D$ for $D \gg 2$, and that the smallest outer sphere that contains $\mathcal{R}_0$ has a radius going to $1$ in the same dimension limit. The overall shape of $\mathcal{R}_0$ is therefore a "squashed" convex body for large $D$, such that at least one of its dimensions drops rapidly to zero. To estimate the complexity for Case B, we consider CRs of two tractable types: a Type I CR is located at an extremely sharp corner of $\mathcal{R}_0$ that is made from at least one of its rapidly shrinking dimensions, as shown in Fig. 3(b), whereas a Type II CR is situated at an extremely flat boundary of $\mathcal{R}_0$ where all of its dimensions remain approximately constant within the CR as in Fig. 3(c). For a conservative estimate of cmpl, we consider an $\mathcal{R}$ such that the longest axis of $\delta_1^+$ is aligned with the flat surface. All other types of Case-B CRs may be viewed as intermediate situations of these two and have no analytical complexity estimates known to us. The data-copy number $N \gg 1$ is assumed to be sufficiently large such that $g_{ml} \approx 0$ and $r_c \approx \tilde{r}_{ml}$.

To estimate cmpl for a Type I CR, we assume that the corner is extremely sharp in one particular dimension such that the curvature of $\partial \mathcal{R}_0$ extending out from $\tilde{r}_{ml}$ is almost flat. Then following Fig. 4(a), the concept of similar figures give $R_{out}/R_{in} \approx D$, which is independent of $\tilde{F}_{ML}$ for extremely sharp corners, and $\text{cmpl}_{B,I} = O(D^3)$. The complexity for Type II CRs may be estimated with the help of Fig. 4(b), where $R_{out}/R_{in} \approx 2 \text{cond}(\tilde{F}_{ML}^{-1/2})$ is now independent of $\partial \mathcal{R}_0$ due to its extremely mild edge features, leading us to $\text{cmpl}_{B,II} = O(D^3 \text{cond}(\tilde{F}_{ML}^{-1/2})) = \text{cmpl}_A$.
VI. DISTANCE-INDUCED SIZE

A. The operational definition

The notion of region averages paves the way to other creative ways of defining size. Doing so permits us to talk about size of a CR without referencing to \( \partial \mathcal{R}_0 \) entirely. To begin, one could think of size as a measurement of the average distance between any two points inside \( \mathcal{R} \). Intuitively, the smaller this average distance, the smaller the region and vice versa. Using this simple notion, we propose the region-average quantity

\[
S_{\varphi, \lambda} \equiv \mathcal{D}(\mathcal{R}', \tilde{\mathcal{R}}_{\text{ML}})_{\lambda} = \frac{\int_{\mathcal{R}_\lambda} (\mathrm{d}\mathbf{r}') \mathcal{D}(\mathbf{r}', \tilde{\mathbf{r}}_{\text{ML}})}{\int_{\mathcal{R}_\lambda} (\mathrm{d}\mathbf{r}')}
\]

(11)

to measure the size of \( \mathcal{R}_\lambda \), where \( \mathcal{D}(\mathbf{r}', \tilde{\mathcal{R}}_{\text{ML}}) \) is some pre-chosen distance metric. Notice that the ML estimator \( \tilde{\mathcal{R}}_{\text{ML}} \) is selected to be the reference point to which distances are measured without loss of generality.

To concretize all results, we shall look at three distance measures for states that enjoy a good reputation in quantum-information studies. We first mention the Hilbert-Schmidt (HS) distance

\[
\mathcal{D}_{\text{HS}} = \text{tr}\{ (\rho' - \tilde{\rho}_{\text{ML}})^2 \} = (\mathbf{r}' - \tilde{\mathbf{r}}_{\text{ML}})^2,
\]

(12)

which is equivalent to the squared \( l_2 \)-norm of \( \mathbf{r}' - \tilde{\mathbf{r}}_{\text{ML}} \). Closely related to the HS distance is the trace-class distance

\[
\mathcal{D}_{\text{tr}} = \text{tr}\{ |\rho' - \tilde{\rho}_{\text{ML}}| \}\]

(13)

defined by the operator absolute value \( |A| = \sqrt{A^\dagger A} \). To introduce the third measure, we start by quoting the expression of quantum fidelity [29]

\[
\mathcal{F} = \text{tr}\left\{ \sqrt{\tilde{\rho}_{\text{ML}} \rho' \tilde{\rho}_{\text{ML}}} \right\}^2
\]

(14)

between \( \rho' \) and \( \tilde{\rho}_{\text{ML}} \), to which we can define the Bures distance [30, 31] \( \mathcal{D}_B = 2 \left( 1 - \sqrt{\mathcal{F}} \right) \). In the limit of large \( N \), where the fidelity \( \mathcal{F} \approx 1 - \epsilon \) differs from 1 by a small amount, \( \mathcal{D}_B \) is also approximately the infidelity \( 1 - \mathcal{F} \).

B. Monotonic behavior of \( S_{\varphi, \lambda} \) for \( N \gg 1 \)

Here, we show that, at least for sufficiently large \( N \), \( S_{\varphi, \lambda} \), defined by any of these three distance measures, behave correctly as a size function in the sense that \( S_{\varphi, \lambda} \) should not increase as \( \lambda \) increases. We first look at the more complicated Case B, and argue that since the state space \( \mathcal{R}_0 \) generally has only corners and edges with no other mathematically pathological features, a set of hyperplanes can then be used to model any particular boundary feature on which \( \tilde{\mathcal{R}}_{\text{ML}} \) resides.

At this stage, we shall consider the asymptotic expressions of the distance measures. The HS distance \( \mathcal{D}_{\text{HS}} \) takes on the simplest (quadratic) form out of all three, which very straightforwardly gives the asymptotic dependence \( S_{\text{HS}, \lambda} \to -\log \lambda \) provided the sufficient condition \( (\mathrm{d} \alpha') = g(\alpha) (\mathrm{d} \alpha') \), which includes the uniform primitive prior \( (\mathrm{d} \mathbf{r}') = (\mathrm{d} \mathbf{r}')_{\text{unit}} \equiv \prod_j (\mathrm{d} r_j) \). It is not difficult to see that the same \( \lambda \) dependence applies to Case A by taking \( w_j = 0 \), so that \( S_{\text{HS}, \lambda} \) is monotonically decreasing with increasing \( \lambda \). Next, according to Appendix A, in the limit of large \( D \), the trace-class distance \( S_B \sim \sqrt{\text{tr} \mathcal{D}} \), which is also clearly monotonic as well owing to \( S_{\text{HS}} \)'s monotonicity. For the Bures distance \( S_B \), one can perform a Taylor expansion on them about \( \tilde{\mathcal{R}}_{\text{ML}} \) (see Sec. VII) and realize that it asymptotically depends on the dyadic \( (\mathbf{r}' - \tilde{\mathbf{r}}_{\text{ML}})(\mathbf{r}' - \tilde{\mathbf{r}}_{\text{ML}})^\dagger \), and is hence also monotonically decreasing with \( \lambda \).
VII. APPROXIMATION FORMULAS FOR $S_{k}$ AND $\alpha_k$

The prior content $S_{\lambda}$ discussed alongside $C_{\lambda}$ in Secs. II–V quantifies the size of $\mathcal{A}_{\lambda}$ relative to $\mathcal{R}_{0}$. In our earlier article [18], analytical approximation formulas for $S_{\lambda}$ were proposed in the large-$N$ limit, all of which are scaled with the volume $V_{\mathcal{A}_{\lambda}}$ of the quantum-state space $\mathcal{A}_{\lambda}$. As is also shown later in the section, this volume dependence is associated with the extension of every $\mathcal{R}_{0}$ integral

$$\int_{\mathcal{A}_{\lambda}} (d\mathbf{r}) \cdots \to \frac{1}{V_{\mathcal{A}_{\lambda}}} \int_{\text{all space}} \prod_{j} d\mathbf{r}_{j} \cdots$$

(16)

to the entire $\mathbf{r}$ space ascribed with the uniform primitive prior, which is a reasonable step to obtain analytical results under the central limit theorem since $L$ is narrow enough to reside within the confines of $\mathcal{A}_{\lambda}$ under this limit. Therefore, the valid usage of these theoretical expressions hinges on the availability of $V_{\mathcal{A}_{\lambda}}$. In quantum-state tomography where we have no complete theoretical information about $\mathcal{R}_{0}$, $V_{\mathcal{A}_{\lambda}}$ is known only for certain priors and state parametrizations [20, 21, 32, 33].

On the other hand, it is obvious that $V_{\mathcal{A}_{\lambda}}$ is canceled out for any region-average quantity after such integral extensions. This allows one to derive operational asymptotic formulas for averages like $S_{\lambda}$ and $\alpha_k$ regardless of $\mathcal{R}_{0}$ in whichever parametrization. As a calculable standard in this section, we continue to derive expressions in terms of the uniform primitive prior and $\mathbf{r}$, although the subsequent instructions may also work for other manageable priors with which $S_{\lambda}$ behaves as a proper size function. We first address the different $\mathcal{D}$ measures in the large-$N$ limit.

A. The various $\mathcal{D}$ measures

1. Hilbert-Schmidt and trace-class measures

The HS distance measure $\mathcal{D}_{\text{HS}}(\mathbf{r}', \hat{\mathbf{r}}_{\text{ML}})$ takes the very simple quadratic form in (12) under any circumstance, whereas the trace-class distance $\mathcal{D}_{\text{tr}}$ has no easy functional form in terms of $\mathbf{r}'$ for $D > 2$. Nevertheless in the limits $N \gg 1$ and $D \gg 2$, based on the principles of random matrix theory detailed in Appendix A, it is deduced that the asymptotic expression

$$S_{\text{tr}} \approx \frac{8 \sqrt{D S_{\text{HS}}}}{3 \pi}$$

(17)

relating the final $\mathcal{D}$-averages $S_{\text{HS}}$ and $S_{\text{tr}}$ is approximately valid for both Case A and B, which incidentally takes the same form found in [34] that was calculated for statistical-fluctuation studies.

2. Bures measure

The Bures distance measure $\mathcal{D}_{\text{b}}$ also has no tractable functional form in $\mathbf{r}'$ for general $D$. To find the asymptotic link with $\mathbf{r}'$ this time, it is technically more convenient to inspect the behavior of $\mathcal{F}$ around $\hat{\mathbf{r}}_{\text{ML}} \leftrightarrow \hat{\mathbf{r}}_{\text{ML}}$ as $N \gg 1$.

A Taylor expansion about $\hat{\mathbf{r}}_{\text{ML}}$ as guided in Appendix B, we have

$$\mathcal{F}_{\lambda} \approx 1 - \frac{1}{2} (\mathbf{r}' - \hat{\mathbf{r}}_{\text{ML}})^{T} \mathbf{Q}_{D} (\mathbf{r}' - \hat{\mathbf{r}}_{\text{ML}})$$

(18)

for Case A and

$$\mathcal{F}_{\lambda} \approx 1 + (\mathbf{r}' - \hat{\mathbf{r}}_{\text{ML}})^{T} \mathbf{P} \Omega$$

$$+ \frac{1}{2} (\mathbf{r}' - \hat{\mathbf{r}}_{\text{ML}})^{T} \left( \frac{1}{2} \text{tr} \{ \mathbf{P} \Omega \} \text{tr} \{ \mathbf{P} \Omega^{T} \} - \mathbf{Q}_{\lambda} \right) (\mathbf{r}' - \hat{\mathbf{r}}_{\text{ML}})$$

(19)

for Case B, where $\mathbf{P}$ is the projector onto the support of $\hat{\mathbf{r}}_{\text{ML}}$, and

$$\mathbf{Q}_{\lambda} = \sum_{j=1}^{r} \sum_{k=1}^{r} \frac{\langle \lambda_{j} | \Omega | \lambda_{k} \rangle \langle \lambda_{k} | \Omega^{T} | \lambda_{j} \rangle}{\lambda_{j} + \lambda_{k}}.$$  

(20)

B. Case A: hyperellipsoidal theory

The presentation in Sec. VII A reduces the necessary ingredients for large-$N/D$ analytical estimations of $S_{\lambda}$ to just the scalar $\int_{\mathcal{A}_{\lambda}} (d\mathbf{r}')$, column $\int_{\mathcal{A}_{\lambda}} (d\mathbf{r}') \Delta_{\text{ML}}$, and dyadic $\int_{\mathcal{A}_{\lambda}} (d\mathbf{r}') \Delta_{\text{ML}}^{T} \Delta_{\text{ML}}^{T}$, where $\Delta_{\text{ML}} = \mathbf{r}' - \hat{\mathbf{r}}_{\text{ML}}$.

When $\mathcal{A}_{\lambda} \approx \delta_{\lambda}$, these three integrals takes on simple analytical forms. We start with

$$\int_{\mathcal{A}_{\lambda}} (d\mathbf{r}') = \int_{\mathcal{A}_{0}} (d\mathbf{r}') \eta(1 - \Delta_{\text{ML}}^{T} F_{\text{ML}} \Delta_{\text{ML}} / (-2 \log \lambda))$$

(21)

and transform $\mathbf{r}' \to \mathbf{r}'' = D^{1/2} \mathbf{O}^{-1/2} \Delta_{\text{ML}}$ to the translated diagonal coordinate variables of $F_{\text{ML}} / (-2 \log \lambda) = OD^{1/2}O^{T}$, so that in the large-$N$ limit and uniform primitive prior, we may relax the boundary of $\mathcal{R}_{0}$ and write

$$\int_{\mathcal{A}_{\lambda}} (d\mathbf{r}') \to \frac{\det \{ D^{-1/2} \}}{V_{\mathcal{A}_{0}}} \int (d\mathbf{r}'') \eta(1 - \rho^{2})$$

$$= \frac{V_{d}}{V_{\mathcal{A}_{0}}} (-2 \log \lambda)^{d/2} \det \{ F_{\text{ML}} \}^{-1/2},$$

(22)

which is a function of the volume $V_{d} = \pi^{d/2} / (d/2)!$ of the $d$-dimensional unit hyperball, the inverse of $F_{\text{ML}}$ that characterizes $\delta_{\lambda}$ together with the logarithm of $\lambda$.

In this case, the integral column is zero since the integrand after variable transformation becomes odd in $\mathbf{r}''$, and we are thus left with

$$\int_{\mathcal{A}_{\lambda}} (d\mathbf{r}') \Delta_{\text{ML}}^{T} \Delta_{\text{ML}}^{T} \to \frac{\det \{ D^{-1/2} \}}{V_{\mathcal{A}_{0}}} O D^{-1/2} I D^{-1/2} O^{T},$$

(23)

and

$$I = \int (d\mathbf{r}'') \eta(1 - \rho^{2}) \rho'' \rho'^{T}$$
\[= \int_{\text{unit sphere}} (d^\prime r^\prime) r^\prime r^{\prime \top} \]
\[= \int_0^1 dr^\prime r^{\prime d+1} (d\{\text{solid angle}\}) e^\prime e^{\prime \top} = \frac{V_d}{d+2} \mathbf{1}, \]

where the last equality is explained by the orthogonally invariant of the \((d-1)\)-dimensional solid-angle measure over the unit columns \(e^\prime\), and so

\[
\int_{\partial R_0} (d^\prime r^\prime) \Delta_{ML}^\prime \tau_{ML} \rightarrow \frac{V_d}{V_{\partial R_0}} (-2 \log \lambda)^{d/2+1} \det\{F_{ML}\}^{-1/2} F_{ML}^\prime. \tag{25}
\]

With all these components, the relevant asymptotic formulas concerning all three distance measures

\[
S_{\text{HS}, \lambda}^{(A)} \approx \text{Tr}\{F^\prime_{ML}\} \frac{-\log \lambda}{d/2 + 1},
\]

\[
S_{u, \lambda}^{(A)} \quad \text{as in (17),}
\]

\[
S_{b, \lambda}^{(A)} \approx \text{Tr}\{F^\prime_{ML} Q_D\} \frac{-\log \lambda}{d/2 + 1}. \tag{26}
\]

Here \(\text{Tr}\) now addresses the dyadic character, as opposed to \(tr\), and we witness the manifestation of logarithmic divergences from both the relaxation of \(\partial R_0\) and Gaussian approximation of \(L\).

Next, to analytically calculate \(u_{\lambda} \equiv \frac{\log L}{\partial \lambda} / (\log(L_{\max}))\) with which \(C_3\) can thereafter be found, we note that due to the Gaussian form of \(L\),

\[
u_{\lambda} = \frac{1}{\log (\lambda L_{\max})} \left( \log L_{\max} - \int_{\partial R_0} (d^\prime r^\prime) \Delta_{ML}^{\prime \top} F_{ML} \Delta_{ML}^\prime \right) \frac{2}{\int_{\partial R_0} (d^\prime r^\prime)} \]

is a dyadic trace function of \(\int_{\partial R_0} (d^\prime r^\prime) \Delta_{ML}^{\prime \top} F_{ML} \Delta_{ML}^\prime\), so that we may use the right-hand side of (25) and put down

\[
u_{A, \lambda} \approx \frac{\log \left( \lambda d/(d+2) L_{\max} \right)}{\log (\lambda L_{\max})} \]

after some basic trace and logarithmic manipulations. It is clear that \(d/(d+2) \leq u_{A, \lambda} \leq 1\) is bounded.

### C. Case B: hyperellipsoidal-cap theory

In Case B, although the geometry of \(R_\lambda \approx \delta^\prime_+ \cap \partial R_0\) is now much trickier to deal with, the central limit theorem proposed in Sec. III allows us to approximate \(R_\lambda\) by a regular analytical region.

As shown in Fig. 6, one can introduce a hyperplane \(R\), described by \(\mathbf{n} \cdot (r^\prime - \hat{r}_{ML}) = 0\ (\mathbf{n} \sim g_{ML})\) that is tangent to the level curve of the Gaussian function in (4) at \(\hat{r}_{ML}\). The hyperspherical cap formed by \(R\) and \(\delta^\prime_+\) hence asymptotically contains \(R_\lambda\), where we have essentially modeled the highly nontrivial \(\partial R_\lambda \cap \partial R_0\) as \(R\). This model implies the estimated assignment

\[
\int_{\partial R_\lambda} (d^\prime r^\prime) \cdots \]

\[
= \frac{1}{V_{\partial R_0}} \int (d^\prime r^\prime) \eta (1 - (r^\prime - \hat{r}_{ML}) \cdot F_{ML} (r^\prime - \hat{r}_{ML}) / (-2 \log \lambda)) \times \eta (n \cdot (\hat{r}_{ML} - r^\prime)) \cdots. \tag{29}
\]

The change of variable \(r^\prime \rightarrow r^\prime = D^{1/2} O^\top (r^\prime - \hat{r}_{ML})\) with respect to the diagonal coordinates of \(F_{ML}/(-2 \log \lambda) = O^\prime D^\prime O^\top\) leads to

\[
\int_{\partial R_\lambda} (d^\prime r^\prime) q_{\lambda}(r^\prime) \quad \frac{\det\{D^{1/2}\}}{V_{\partial R_0}} \int (d^\prime r^\prime) q_{\lambda}(O^\prime D^{1/2} r^\prime) \]

for any function \(q\), which is parametrized by the cap element \((d^\prime r^\prime)_{\text{cap}} = (d^\prime r^\prime)_{\text{unit}} \eta (1 - (r^\prime - \hat{r}_{ML}) \cdot a - \mathbf{b}^\top r^\prime), a = n^\top (\hat{r}_{ML} - \hat{r}_{c}),\) and \(b = D^{1/2} O^\top \mathbf{n}\). One can check that

\[
I \equiv \frac{\mathbf{a} \cdot (\hat{r}_{ML} - \hat{r}_{c})}{|b|} = \frac{|g_{ML} \cdot (\hat{r}_{ML} - \hat{r}_{c})|}{|D^{1/2} O^\top g_{ML}|} \leq 1. \tag{31}
\]

In other words, we have

\[
\frac{a^\prime}{d_{\lambda}} (r^\prime)_{\partial R_\lambda} \approx \frac{\int (d^\prime r^\prime)_{\text{cap}} q_{\lambda}(O^\prime D^{1/2} r^\prime)}{\int (d^\prime r^\prime)_{\text{cap}}}, \tag{32}
\]

and that for any \(q_{\lambda}\) belonging to either one of the three distance measures or \(\log \lambda / \log (L_{\max})\), as reasoned in Sec. VII C, the building blocks of \(q_{\lambda}(r^\prime)_{\partial R_\lambda}\) are only \(\int (d^\prime r^\prime)_{\text{cap}}, \int (d^\prime r^\prime)_{\text{cap}} r^\prime r^\prime_{\text{cap}}\) and \(\int (d^\prime r^\prime)_{\text{cap}} r^\prime r^{\prime \top}\). These integrations are all carried out in Appendix C.

In combining all results gathered from Appendices B and C, we denote \(N_{d,l,x} = V_d l^{(1-l)/2}((d + x)/2, (d + x)/2),\) which
VIII. RESULTS AND DISCUSSIONS

We first present the computation results of $S_A$ and $C_A$ from $u_{\lambda}$ in Figs. 7 and 8 for quantum systems of various dimensions $D$. To be more technically precise about our use of Euler's method described in Sec. IV, we first solve (9) for $y_2$ by iterating (10) starting with a numerically small $\lambda$ value, say $10^{-10}$, to $\lambda \approx 1$ using the function $q_\lambda = \log L / \log (\lambda L_{max})$.

The behavior of $S_A$ shows the expected decreasing trend not only in $\lambda$, but also in overall magnitude as $D$ increases. This is partnered with the corresponding sharpening of $u_{\lambda}$ around $\lambda = 0$ as $D/N$ increases, which equivalently indicates that the (log-)likelihood is turning into a delta-function peak. For larger $D$ or $N$, the computational accuracy of $S_A$ and $C_A$ using Euler’s numerical method may be maintained by exploring many more $\lambda$ values near zero, as all curves possess sharp gradient changes in this $\lambda$ range.

In Fig. 9, both simulated data and theoretical curves of all three distance-induced size functions $S_{\text{his}}, S_u$, and $S_\text{M}$ are plotted against the credibility $C$ for Case A. In this case, there exists no other factors that could spoil the perfect hyperellipsoidal geometry of $\mathcal{R}_A$. As such, the analytical curves fit almost perfectly with the simulated points. We note that even the average trace-class distance $S_u$, which is approximated with (17) through the theory of random matrices, performs very well relative to the simulated data points.

In Case B, we can start to see discrepancies between theory and simulation from Fig. 10 especially for larger $D$. Such deviations are inevitable as the hyperellipsoidal-cap estimation of the actual CR $\mathcal{R}_B$ proposed in Sec. VIIIC introduces additional space outside $\mathcal{R}_B$ that is certainly not contained in $\mathcal{R}_B$. More generally, for very large $D$, if the rank-deficient ML estimator $\hat{\rho}_\text{ML}$ is located at an extremely sharp state-space

depending on the incomplete Euler’s beta function $I(\cdot, \cdot)$, and organize two new auxiliary quantities

\[
\begin{align*}
\mathbf{m} &= -\frac{V_d - 1}{(d + 1)} (1 - r^2)^{(d+1)/2} + N_{d,1,1} F_{\text{ML}}^{-1} \mathbf{G}_{\text{ML}}, \\
\mathbf{M} &= -\frac{\log \lambda'}{d+2} N_{d,1,1} F_{\text{ML}}^{-1} + \frac{1}{2} \mathbf{m} \mathbf{g}_{\text{ML}} \mathbf{F}_{\text{ML}}^{-1}.
\end{align*}
\]

This helps to clean the respective formulas

\[
\begin{align*}
S_{\text{his},A}^{(B)} &\approx \frac{\text{Tr} \{ 2 \mathbf{M} \}}{N_{d,1,1}}, \\
S_{\text{tr},A}^{(B)} &\approx \text{Tr} \{ \mathbf{M} \mathbf{Q}_{a} \} - \frac{\text{Tr} \{ \mathbf{F}_{\text{ML}} \mathbf{M} \}}{N_{d,1,1}},
\end{align*}
\]

for the distance-induced size functions and

\[
\begin{align*}
u_{B,\lambda}^{(B)} &\approx \frac{1}{\text{log}(\lambda' L_{max})} \left( \text{log} L_{max} + m^2 \mathbf{g}_{\text{ML}} - \text{Tr} \{ \mathbf{F}_{\text{ML}} \mathbf{M} \} \right). \tag{35}
\end{align*}
\]

We caution the Reader once more regarding the actions of $\text{tr}$ and $\text{Tr}$ at the right-hand side of $S_{\text{his},A}$ in (34).

For consistency, we end this section by noting that Eqs. (34) and (35) cover Eqs. (26) and (27) because Case A implies that $\lambda' = \lambda$, $g_{\text{ML}} = 0 = \mathbf{m}$, such that $l = 0$ then gives $N_{d,0,1} = V_d$ and $\mathbf{M} = (\log \lambda') F_{\text{ML}}^{-1}(d+2)$. 

\[\text{FIG. 7. (Color Online) Graphs of the conventional size function } S_A, \text{ credibility } C_A \text{ and } u_{\lambda} \text{ for Case A with } 2 \leq D \leq 5. \text{ M, the number of POM outcomes, is set to } D^3 \text{ and the POM is chosen to be a random square-root measurement as a simulation example for each } D. \text{ Here } N/M = 500 \text{ and all horizontal axes (including those of the insets) represent } \lambda. \text{ All markers are guided by dashed line segments.}\]

\[\text{FIG. 8. (Color Online) Graphs of computed } S_A \text{ and } C_A \text{ from the sample average } u_{\lambda} \text{ for Case B against } \lambda, \text{ with } 2 \leq D \leq 4, \text{ which have, otherwise, the same specifications as Fig. 7. The rank of the ML estimator for each panel is explicitly stated.}\]
VII. C

We may still be multidimensional parameter estimation. The construction of after tomography is, unfortunately, a highly nontrivial problem owing to the complex constraints inherited from the state space.

When the sample space is relatively big, as the correlations (size and credibility) quickly becomes infeasible when the dataset collected in an experiment is relatively big, the corresponding credible region would be very small with respect to the state space.

In this article, we have established and explained a reformulated Bayesian error-region theory to compute credible-region qualities by calculating a sequence of region-average quantities. This procedure transforms the general credible-region construction into a sequence of direct region sampling followed by a simple numerical solution to a single-variable differential equation. This results in no sample wastage since no points are discarded. The technique of accelerated hit-and-run is one numerical scheme that can be used to compute region averages rather efficiently, with roughly predictable numerical complexity in the context of tomography despite the complicated state-space boundary.

Furthermore, for highly complex quantum systems of extremely large dimensions, where all numerical methods eventually become practically infeasible, we derive a set of analytical formulas to perform approximate Bayesian error certification using the concept of distance-induced size measures that alternatively quantifies how large a credible region is. One advantage of these theoretical formulas can be clearly seen in error certifications of more complex objects such as quantum processes, which are quadratically larger in Hilbert-space dimension than quantum states and possess more complicated trace-preserving parameter spaces of obscure geometry.

IX. CONCLUSIONS

Quantum-state tomography is an important application of multidimensional parameter estimation. The construction of Bayesian credible regions to the reconstructed quantum states after tomography is, unfortunately, a highly nontrivial problem owing to the complex constraints inherited from the quantum-state space. Standard numerical recipe of first doing a Monte Carlo sampling of the state space and next discarding points outside the credible region to compute its region qualities (size and credibility) quickly becomes infeasible when the dataset collected in an experiment is relatively big, as the corresponding credible region would be very small with respect to the state space.

In this article, we have established and explained a reformulated Bayesian error-region theory to compute credible-region qualities by calculating a sequence of region-average quantities. This procedure transforms the general credible-region construction into a sequence of direct region sampling followed by a simple numerical solution to a single-variable differential equation. This results in no sample wastage since no points are discarded. The technique of accelerated hit-and-run is one numerical scheme that can be used to compute region averages rather efficiently, with roughly predictable numerical complexity in the context of tomography despite the complicated state-space boundary.

Furthermore, for highly complex quantum systems of extremely large dimensions, where all numerical methods eventually become practically infeasible, we derive a set of analytical formulas to perform approximate Bayesian error certification using the concept of distance-induced size measures that alternatively quantifies how large a credible region is. One advantage of these theoretical formulas can be clearly seen in error certifications of more complex objects such as quantum processes, which are quadratically larger in Hilbert-space dimension than quantum states and possess more complicated trace-preserving parameter spaces of obscure geometry.

ACKNOWLEDGMENTS

We acknowledge financial support from the BK21 Plus Program (21A201311111123) funded by the Ministry of Education (MOE, Korea) and National Research Foundation of Korea (NRF), the NRF grant funded by the Korea government (MSIP) (Grant No. 2010-0018295), and the Basic Science Research Program through the NRF funded by the Ministry of Education (No. 2018R1D1A1B07048633).
Appendix A: The relationships between $S_{HS}$ and $S_{tr}$ in the large-$N$ limit

Apart from $D_{HS}$, all other measures have no direct analogs in the $r'$ parametrization. However in certain limits, all these measures have approximate relations with $D_{HS}$.

We start with making an approximate connection between $S_{tr}$ and $S_{HS}$ by examining the Hermitian operator $\Delta \rho' = \rho' - \bar{\rho}_{ML}$ ($\rho' \in \mathcal{H}$). In Case A, the distribution of $\Delta \rho'$ in $\mathcal{H}$ has zero mean, $\Delta \rho' \overset{\text{dist}}{=} 0$. This is also approximately true for the Case B situation when $N$ is sufficiently large such that $\mathcal{H}$ to be small. Furthermore, the CR $\mathcal{F}$ is essentially a bounded set of Hermitian random operators. Here, we shall make the assumption that each matrix entry $[\Delta \rho']_{jk}$ in the computational basis is an independent random complex number up from the diagonal elements. Under this condition, the $\Delta \rho'$'s form what is now known as a Wigner ensemble [35–38] with the second moment equal to $[\Delta \rho]_{jk}^{2} = \text{tr}{(\Delta \rho')^{2}} = S_{HS}$. Moreover, they are known to have an i.i.d. eigenvalue spectrum that follows the Wigner semicircle law

$$\sigma{(\Delta \rho' / \sqrt{\mathcal{D}})} \sim \frac{1}{2\pi S_{HS}} \sqrt{4S_{HS} - x^{2}} \quad \text{for} \quad -2 \sqrt{S_{HS}} \leq x \leq 2 \sqrt{S_{HS}}$$

(A1)

in the large-$D$ limit. The trace-class distance $D_{tr}$ can thus be calculated with the integral

$$D_{tr} \approx \frac{\sqrt{\mathcal{D}}}{2\pi S_{HS}} \int_{-2\sqrt{S_{HS}}}^{2\sqrt{S_{HS}}} dx |x| \sqrt{4S_{HS} - x^{2}} = \frac{8}{3\pi} \sqrt{DS_{HS}} \quad \text{so that we end up with (17).}$$

For Case B, that $\Delta \rho' \overset{\text{dist}}{=} 0$ is obvious, but as we have no means of analytically estimate $\Delta \rho' \overset{\text{dist}}{=} 0$, we make a further approximation that as long as $D_{tr}$ is sufficiently small, the offset to $\Delta \rho' \overset{\text{dist}}{=} 0$ will proportionately be small, so that (17) remains a reasonable asymptotic approximation.

Appendix B: Fidelity in the large-$N$ limit

A Taylor expansion of $\mathcal{F}$ about $\tilde{\rho}_{ML}$, or

$$\mathcal{F} \approx 1 + (r' - \tilde{\rho}_{ML})^{T} \frac{\partial \mathcal{F}_{ML}}{\partial \tilde{\rho}_{ML}} + \frac{1}{2} (r' - \tilde{\rho}_{ML})^{T} \frac{\partial^{2} \mathcal{F}_{ML}}{\partial \tilde{\rho}_{ML}^{2}} (r' - \tilde{\rho}_{ML}),$$

(B1)

reveals the large-$N$ characteristics that is needed for analysis. The structure of (14), however, demands the operator variation of $\sqrt{A}$ for a positive (semidefinite) $A$. An integral representation of $\sqrt{A}$ exists [39] and can be written as

$$\sqrt{A} = \lim_{\epsilon \to 0^{+}} \int_{0}^{\infty} \frac{dt}{\pi \sqrt{t}} \frac{A}{t + \epsilon},$$

(B2)

where the limit is understood to be applied at the very end of all calculations so that Eq. (B2) is valid even for $A$ with zero eigenvalues.

The first-order variation of $\text{tr}\{\sqrt{A}\}^{2}$ produces

$$\delta \text{tr}\{\sqrt{A}\}^{2} = 2 \text{tr}\{\sqrt{A}\} \lim_{\epsilon \to 0} \int_{0}^{\infty} \frac{dt}{\pi \sqrt{t}} \text{tr}\left\{ \delta \frac{A}{t + \epsilon} \right\}$$

$$= 2 \text{tr}\{\sqrt{A}\} \lim_{\epsilon \to 0} \int_{0}^{\infty} \frac{dt}{\pi \sqrt{t}} \left\{ \delta A \frac{1}{t + \epsilon} - A \frac{\delta A}{(t + \epsilon)^{2}} \right\} \delta A \frac{1}{t + \epsilon}.$$ (B3)

In terms of $\mathcal{F}$, we substitute $A = \tilde{\rho}_{ML}^{1/2} \rho' \tilde{\rho}_{ML}^{1/2}$, and evaluate the above result with $\rho' = \bar{\rho}_{ML}$, or $A \to \Lambda_{ML} = \tilde{\rho}_{ML}^{1/2} \Lambda \tilde{\rho}_{ML}^{1/2}$, then with $\delta \Lambda_{ML} = \tilde{\rho}_{ML}^{-1/2} \delta \tilde{\rho}_{ML} \cdot \Omega \tilde{\rho}_{ML}^{1/2},$

$$\frac{\partial \mathcal{F}_{ML}}{\partial \tilde{\rho}_{ML}} = 2 \text{tr}\left\{ \rho_{ML} \frac{\tilde{\rho}_{ML}^{-1/2} \Lambda + 2 \epsilon \Omega}{(\rho_{ML} + \epsilon)^{1/2} \Omega} \right\},$$

(B4)

where we remind the Reader that tr acts on operators only, not on the vectorial character. For Case B in which $\rho_{ML} = \sum_{j=1}^{M} |\lambda_{j}\rangle \langle \lambda_{j}|$ is rank-deficient, we get, after taking the trace,

$$\frac{\partial \mathcal{F}_{ML}}{\partial \tilde{\rho}_{ML}} = \text{tr}\{P_{r} \Omega\},$$

(B5)

where $P_{r} = \sum_{j=1}^{M} |\lambda_{j}\rangle \langle \lambda_{j}|$. It is then trivial to realize that this first-order derivative is zero for Case A. Qualitatively, this confirms the fact that when $\tilde{\rho}_{ML}$ is an interior point, $\mathcal{F}$ has a local maximum at this point as it should, while a boundary estimator evaluates to a nonzero $\mathcal{F}$ slope.

Upon denoting $W_{ML} = \tilde{\rho}_{ML}^{1/2} \Omega \tilde{\rho}_{ML}^{1/2}$, the second-order variation follows from the second line of (B3):

$$\delta \frac{\partial \mathcal{F}_{ML}}{\partial \tilde{\rho}_{ML}} \frac{\partial \mathcal{F}_{ML}}{\partial \tilde{\rho}_{ML}} = 2 \text{tr}\left\{ \sqrt{A} \right\} \lim_{\epsilon \to 0} \int_{0}^{\infty} \frac{dt}{\pi \sqrt{t}} \text{tr}\left\{ W_{ML} \frac{t + \epsilon}{(t + A + \epsilon)^{2}} \right\}$$

$$\times \text{tr}\left\{ W_{ML} \frac{t + \epsilon}{(t + A + \epsilon)^{2}} \right\}.$$ (B6)

A product-rule dissociation of (B6) comprises a $\delta \text{tr}\{\sqrt{A}\}$ and

$$\lim_{\epsilon \to 0} \int_{0}^{\infty} \frac{dt}{\pi \sqrt{t}} \text{tr}\left\{ W_{ML} \frac{t + \epsilon}{(t + A + \epsilon)^{2}} \right\}$$

$$= - \lim_{\epsilon \to 0} \int_{0}^{\infty} \frac{dt}{\pi \sqrt{t}} \text{tr}\left\{ W_{ML} \left[ \frac{t + \epsilon}{(t + A + \epsilon)^{2}} \delta A \frac{1}{(t + A + \epsilon)^{2}} + \frac{t + \epsilon}{(t + A + \epsilon)^{3}} \delta A \frac{1}{(t + A + \epsilon)} \right] \right\}. \quad (B7)$$

After evaluating the variation at $\rho' = \bar{\rho}_{ML}$ and further undoing all integrations with the help of its spectral decomposition, Case B yields

$$\frac{\partial \mathcal{F}_{ML}}{\partial \tilde{\rho}_{ML}} \frac{\partial \mathcal{F}_{ML}}{\partial \tilde{\rho}_{ML}} = \frac{1}{2} \text{tr}\{P_{r} \Omega\} \text{tr}\{P_{r} \Omega^{T}\}.$$
The counterpart expression for Case A is evident immediately, of course.

Appendix C: Hyperellipsoidal-cap averages

Under the uniform primitive prior, calculations of the hyperellipsoidal-cap integrals

\[ I_0 = \int (d\mathbf{r'})_{\text{cap}}, \]  
\[ I_1 = \int (d\mathbf{r'})_{\text{cap}} \mathbf{r'}, \]  
\[ I_2 = \int (d\mathbf{r'})_{\text{cap}} \mathbf{r'}^{\prime \top}, \]  

specified by the uniform cap-volume element \((d\mathbf{r'})_{\text{cap}} = (d\mathbf{r'})_{\text{umf}} \eta(1 - \mathbf{r'}^2) \eta(a - \mathbf{b} \cdot \mathbf{r'})\) for \(0 \leq a \leq |\mathbf{b}|\) and some column \(\mathbf{b}\), include systematic manipulations of the double Heaviside functions. One route to take exploits the following integral representation

\[ \eta(x) = \int \frac{dt}{2\pi i t - i\varepsilon} \]  

with the implicit limit \(\varepsilon \to 0^+\). We then have, for (C1),

\[ I_0 = \frac{\pi^{d/2}}{\det(A)^{1/2}} \]  
\[ I_1 = \frac{\pi^{d/2}}{\det(A)^{1/2}} \]  
\[ I_2 = \frac{\pi^{d/2}}{\det(A)^{1/2}} \]  

upon noting the well-known \(d\)-dimensional Gaussian integral result

\[ \int (d\mathbf{r'})_{\text{umf}} e^{-\mathbf{r'}^2} = \frac{\pi^{d/2}}{\det(A)^{1/2}} e^{\mathbf{c}^2/\mathbf{A} \mathbf{c}} \]  

for any positive \(\mathbf{A}\). Let us first do the \(t'\) integration by invoking the useful transformation

\[ \frac{1}{z^m} = \frac{1}{(m - 1)!} \int_0^\infty dy y^{m-1} e^{-zy} \text{ for } m > 0: \]  
\[ \int \frac{dt'}{2\pi i t'} e^{iat'} t'^2 \]  
\[ = \int_0^\infty dy \frac{dt'}{2\pi i t'} e^{\pi t'/2} \left( e^{i(a-y)t'} - e^{-i(a-y)t'} \right) \]  
\[ = \sqrt{\frac{i\pi}{\mathbf{b}^2}} \int_0^\infty dy e^{-i(a-y)^2/\mathbf{b}^2}. \]  

As a consequence,

\[ I_0 = \frac{\pi^{d/2}}{2 |\mathbf{b}|} \int_0^\infty dy \frac{dt}{2\pi (i r')^{d/2}} e^{-i(a-y)^2/\mathbf{b}^2} \]  
\[ I_1 = \frac{\pi^{d/2}}{2 (d-1)! |\mathbf{b}|} \int_0^\infty dy \frac{dt}{2\pi (i r')^{d/2}} e^{ir(t')^2/\mathbf{b}^2} \]  
\[ I_2 = \frac{\pi^{d/2}}{2 (d-1)! |\mathbf{b}|} \int_0^\infty dy \frac{dt}{2\pi (i r')^{d/2 - 2}} e^{ir(t')^2/\mathbf{b}^2} \]  

The above integral in \(y\) represents well-known special functions and to see this, we further perform the substitutions \(\cos u = (y - a)/|\mathbf{b}|\) and \(l = a/|\mathbf{b}|\):

\[ \int_0^{a+|\mathbf{b}|} dy \left[ 1 - \frac{a^2}{|\mathbf{b}|^2} \right]^{d/2} = |\mathbf{b}| \int_0^{\cos^{-1} l} d\cos \theta \sin^d \theta \]  
\[ = |\mathbf{b}| B \left( \frac{1}{2}, \frac{1}{2} \right) \frac{d+1}{(d+1/2)} \]  

which is a product of the beta function and its normalized incomplete form

\[ I_{0 \leq a \leq 1}(b, c) = \frac{1}{B(b, c)} \int_0^a du u^b (1 - u)^c. \]  

The final answer reads

\[ I_0 = V_d I_{1/2} \left( \frac{d+1}{2}, \frac{d+1}{2} \right) . \]  

For (C2),

\[ I_1 = \frac{\pi^{d/2}}{2} \frac{b}{(i r')^{d/2}} e^{ir(t')^2/\mathbf{b}^2} \]  

where the \(r'\) integration

\[ \int (d\mathbf{r'})_{\text{umf}} e^{-ir(t')^2 - ib\cdot\mathbf{r'}} \]  
\[ = -\frac{1}{i r'} \frac{\partial}{\partial \mathbf{b}} \int (d\mathbf{r'})_{\text{umf}} e^{-ir(t')^2 - ib\cdot\mathbf{r'}} \]  
\[ = -\frac{\pi^{d/2}}{2 (i r')^{d/2 + 2}} \mathbf{b} e^{ir(t')^2/\mathbf{b}^2} \]  

is simplified after a differentiation under the integral sign. Then

\[ I_1 = \frac{\pi^{d/2}}{2} \frac{b}{(i r')^{d/2 + 2}} e^{ir(t')^2/\mathbf{b}^2} \]  

To simplify the integration, we again recall Eq. (C7) to finally get

\[ I_1 = -\frac{\pi^{d/2}}{2 (d-1)!} \frac{b}{|\mathbf{b}|} \left( 1 - l^2 \right) \frac{d-1}{2}, \quad l = \frac{a}{|\mathbf{b}|}. \]
We can at least verify the $d = 1$ for Eq. (C16) after paying attention to the convention $b \to -b$, for $b \geq 0$. This corresponds to the integral

$$I_{1,d=1} = \int dr'' \eta(1-r'^{2}) \eta(a + b r'') r'' = \int^{-1}_{-1} dr'' r'' = -\frac{1}{2}(1 - l^2). \quad (C17)$$

By the same token, we may explore the dyadic integral $I_2$ in (C1) first with (C4) to obtain

$$I_2 = \int \frac{dt}{2 \pi i} \frac{e^{i t}}{t - i \epsilon} \int \frac{dt'}{2 \pi i} \frac{e^{i a t'}}{t' - i \epsilon} \int (dr'')_{\text{unit}} r'r'^{T} e^{-i tr'^{2} - i tr'r'}, \quad (C18)$$

where the dyadic $r''$ sub-integral

$$\int (dr'')_{\text{unit}} r'r'^{T} e^{-i tr'^{2} - i tr'r'} = -\frac{\pi^{d/2}}{\delta A} \frac{2 \pi \delta}{(it)^{d+1}} \frac{1}{\det(A)^{1/2}} e^{\frac{1}{2 \pi} b^T A b} \quad \bigg|_{A=1} \quad (C19)$$

after an application of Eq. (C6) and a dyadic differentiation under the integral sign this time.

This time, we choose to perform the $t$ and $t'$ integrals before taking the derivative, inasmuch as

$$I_2 = -\delta \frac{\pi^{d/2}}{\delta A \det(A)^{1/2}} \int \frac{dt}{2 \pi} \frac{e^{it}}{(it)^{d+1/2}} \int \frac{dt'}{2 \pi i} \frac{e^{iar'}}{i t'} e^{\frac{1}{2 \pi} b^T A b} \bigg|_{A=1} \quad (C20)$$

where the usage of (C7) evaluates the $t'$ integral

$$\int \frac{dt'}{2 \pi i} e^{iar'} \frac{1}{\sqrt{2 \pi} b^T A b} = 0 \int \frac{dy}{\sqrt{2 \pi} e^{\frac{1}{2 \pi} b^T A b} + i t (a - y)} = \sqrt{\frac{1}{\pi b^T A b}} \int_{0}^{\infty} dy \exp \left(-\frac{(y - a)^2 t}{b^T A b}\right) \quad (C21)$$

into another Gaussian integral. Its convenient feature becomes clear when substituted back into (C20):

$$I_2 = -\frac{\delta}{\delta A \sqrt{\det(A) b^T A b}} \left[ \mathbf{B} \left( \frac{1}{2}, \frac{d + 3}{2} \right) I_{1/2}^{(d+3)/2} \right]_{A=1} \quad (C22)$$

The end of the tunnel becomes visible after a product-rule differentiation carried out with the basic dyadic identities

$$\delta A^{-1} = -A^{-1} \delta A A^{-1} \quad \delta \det(A) = \det(A) \text{Tr}\{A^{-1} \delta A\}, \quad (C23)$$

after which we end up with the final answer

$$I_2 = \frac{\pi^{d-1}}{2 (d+1) !} \left[ \mathbf{B} \left( \frac{1}{2}, \frac{d + 3}{2} \right) I_{1/2}^{(d+3)/2} \right]_{A=1}$$

for the 1D special case can again be extracted from Eq. (C24),

$$I_2 \big|_{d=1} = \frac{1}{2} \left[ \mathbf{B} \left( \frac{1}{2}, \frac{d + 3}{2} \right) I_{1/2}^{(d+3)/2} \right] \left[ 1 + (1 - l^2) \frac{d+1}{d+2} \right]$$

and compared with the direct calculation

$$I_{2,d=1} = \int dr'' \eta(1-r'^{2}) \eta(a + b r'') r'' = \int_{-1}^{1} dr'' r'' = \int_{-1}^{1} dr'' r'^{2} = \frac{1}{3} (1 - l^3). \quad (C26)$$
[1] D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani, Measurement of the wigner distribution and the density matrix of a light mode using optical homodyne tomography: Application to squeezed states and the vacuum, Phys. Rev. Lett. 70, 1244 (1993).
[2] I. Chuang and M. Nielsen, Quantum Computation and Quantum Information (Cambridge University Press, Cambridge, 2000).
[3] J. Řeháček, Z. Hradil, E. Knill, and A. I. Lvovsky, Diluted maximum-likelihood algorithm for quantum tomography, Phys. Rev. A 75, 042108 (2007).
[4] Y. S. Teo, H. Zhu, B.-G. Englert, J. Řeháček, and Z. Hradil, Quantum-state reconstruction by maximizing likelihood and entropy, Phys. Rev. Lett. 107, 020404 (2011).
[5] H. Zhu, Quantum state estimation with informationally overcomplete measurements, Phys. Rev. A 90, 012115 (2014).
[6] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O’Brien, Quantum computers, Nature 464, 45 (2010).
[7] E. T. Campbell, B. M. Terhal, and C. Vuillot, Roads towards fault-tolerant universal quantum computation, Nature 549, 172 (2017).
[8] B. Lekitsch, S. Weidt, A. G. Fowler, K. Mølmer, S. J. Devitt, C. Wunderlich, and W. K. Hensinger, Blueprint for a microwave trapped ion quantum computer, Sci. Adv. 3, e1601540 (2017).
[9] B. Efron and R. J. Tibshirani, An Introduction to the Bootstrap (Chapman & Hall/CRC, New York, 1993).
[10] A. C. Davison and D. V. Hinkley, Bootstrap Methods and their Application (Cambridge University Press, Cambridge, 1997).
[11] D. Suess, L. Rudnicki, T. O Maciel, and D. Gross, Error regions in quantum state tomography: computational complexity caused by geometry of quantum states, New J. Phys. 19, 093013 (2017).
[12] R. Christensen, Testing fisher, neyman, pearson, and bayes, Amer. Statist. 59, 121 (2005).
[13] M. Christandl and R. Renner, Reliable quantum state tomography, Phys. Rev. Lett. 109, 120403 (2012).
[14] R. Blume-Kohout, Robust error bars for quantum tomography, (2012), quant-ph/1202.5270.
[15] P. Faist and R. Renner, Practical and reliable error bars in quantum tomography, Phys. Rev. Lett. 117, 010404 (2016).
[16] J. Shang, H. K. Ng, A. Sehrawat, X. Li, and B.-G. Englert, Optimal error regions for quantum state estimation, New J. Phys. 15, 120326 (2013).
[17] X. Li, J. Shang, H. K. Ng, and B.-G. Englert, Optimal error intervals for properties of the quantum state, Phys. Rev. A 94, 062112 (2016).
[18] Y. S. Teo, C. Oh, and H. Jeong, Bayesian error regions in quantum estimation I: analytical reasonings, New J. Phys. 20, 093009 (2018).
[19] C. Oh, Y. S. Teo, and H. Jeong, Bayesian error regions in quantum estimation II: analytical reasonings, New J. Phys. 20, 093010 (2018).
[20] I. Bengtsson and K. Życzkowski, Geometry of Quantum States: An Introduction to Quantum Entanglement (Cambridge University Press, Cambridge, 2006).
[21] I. Bengtsson, S. Weis, and K. Życzkowski, Geometric Methods in Physics, XXX Workshop 2011, edited by P. Kielanowski, S. T. Ali, A. Odzijewicz, M. Schlichenmaier, and T. Voronov, Trends in Mathematics (Springer, Basel, 2013).
[22] J. Shang, Y.-L. Seah, H. K. Ng, D. J. Nott, and B.-G. Englert, Monte carlo sampling from the quantum state space. i, New J. Phys. 17, 043017 (2015).
[23] Y.-L. Seah, J. Shang, H. K. Ng, D. J. Nott, and B.-G. Englert, Monte carlo sampling from the quantum state space. ii, New J. Phys. 17, 043018 (2015).
[24] J. C. Butcher, Numerical Methods for Ordinary Differential Equations (John Wiley & Sons, New York, 2003).
[25] S. Katsupaibul, R. L. Smith, and Z. B. Zahbinsky, An analysis of a variation of hit-and-run for uniform sampling from general regions, ACM T. Model. Comput. S. 21, 16 (2011).
[26] R. A. Horn and C. R. Johnson, Matrix Analysis (Cambridge University Press, New York, 1985).
[27] N. J. Higham, Cholesky factorization, WIREs Comp. Stat. 1, 251 (2009).
[28] B. Cousins and S. Vempala, A practical volume algorithm, Math. Prog. Comp. 8, 133 (2016).
[29] A. Uhlmann, The “transition probability” in the state space of a *-algebra, Rep. Math. Phys. 9, 273 (1976).
[30] D. Bures, An extension of kakutani’s theorem on infinite product measures to the tensor product of semifinite *-algebra, Trans. Am. Math. Soc. 135, 199 (1969).
[31] S. L. Braunstein and C. M. Caves, Statistical distance and the geometry of quantum states, Phys. Rev. Lett. 72, 3439 (1994).
[32] K. Życzkowski and H.-J. Sommers, Hilbertschmidt volume of the set of mixed quantum states, J. Phys. A: Math. Gen. 36, 10115 (2003).
[33] A. Andai, Volume of the quantum mechanical state space, J. Phys. A: Math. Gen. 39, 13641 (2006).
[34] H. Zhu and B.-G. Englert, Quantum state tomography with fully symmetric measurements and product measurements, Phys. Rev. A 84, 022327 (2011).
[35] E. Wigner, Characteristic vectors of bordered matrices with infinite dimensions, Ann. Math. 62, 548 (1955).
[36] E. Wigner, Characteristic vectors of bordered matrices with infinite dimensions ii, Ann. Math. 65, 203 (1957).
[37] A. Pizzo, D. Renfrew, and A. Soshnikov, On finite rank deformations of wigner matrices, Ann. I. H. Poincare-Pr. 49, 64 (2013).
[38] M. L. Mehta, Random Matrices (Elsevier, Amsterdam, 2004).
[39] D. A. Bini, N. J. Higham, and B. Meini, Algorithms for the matrix pth root, Numer. Algorithms 39, 349 (2005).