

# Behavior of the Maximum Likelihood in Quantum State Tomography

Travis L Scholten and Robin Blume-Kohout

Center for Computing Research (CCR), Sandia National Labs and  
Center for Quantum Information and Control (CQuIC), University of New Mexico

(Dated: April 26, 2017)

Quantum state tomography on a  $d$ -dimensional system demands resources that grow rapidly with  $d$ . Model selection can be used to tailor the number of fit parameters to the data. However, many methods rely on comparing test statistics (such as the ratio of two models' maximum likelihoods) to their expected behavior under baseline assumptions. In state tomography, the presence of positivity constraints (e.g.  $\rho \geq 0$ ) can violate those assumptions, especially local asymptotic normality. Applying those methods to state tomography thus requires understanding and quantifying how positivity constraints affect test statistics. Here, we study the behavior of the maximum likelihood in different Hilbert space dimensions, and derive an expression for the expected value of the loglikelihood ratio statistic (roughly, the logarithm of the maximum likelihood) between state spaces of different Hilbert space dimension  $d$ . In addition to enabling reliable model selection, our results shed more light on qualitative effect of positivity constraints on state tomography.

General todo: replace  $\hat{\rho}_{MLE}$  by  $\hat{\rho}_{ML}$ .

I  
hate  
that  
phrase.  
Need  
a  
new  
one

In quantum information science, an experimentalist may wish to determine the quantum state  $\rho_0$  that is produced by a specific initialization procedure. This can be done using quantum state tomography [? ]: many copies of  $\rho_0$  are produced; they are measured in diverse ways; and finally the outcomes of those measurements (data) are collated and analyzed to produce an estimate  $\hat{\rho}$ . This is a straightforward statistical inference process [? ? ], where the data are used to fit the parameters of a statistical model. However, we don't always know what model to use. In state tomography, the parameter is  $\rho$ , and the model is the set of all possible density matrices on a Hilbert space  $\mathcal{H}$  (equipped with the Born rule). It is not always *a priori* obvious what  $\mathcal{H}$  or its dimension is; examples include optical modes [? ? ? ? ? ] and leakage levels in AMO and superconducting [? ? ] qubits. In such situations, we seek to let the data itself determine which of many candidate Hilbert spaces is best suited for reconstructing  $\rho_0$ .

Choosing an appropriate Hilbert space on the fly is an instance of a general statistical problem called *model selection*. Although model selection has been thoroughly explored in classical statistics [? ], its application to quantum tomography encounters some obstacles. They stem from the fact that quantum states (and thus, estimates of them) must satisfy a *positivity constraint* ( $\rho \geq 0$ ). A similar constraint (complete positivity) applies to process tomography. The impact of positivity constraints on state and process tomography is an active area of research [? ? ? ? ], and its implications for model selection have also been considered [? ? ? ? ? ? ? ]. In this paper, we address a specific question at the heart of this matter: *How does the loglikelihood ratio statistic used in many model selection protocols, including (but not limited to) information criteria such as Akaike's AIC, behave in the presence of the positivity constraint  $\rho \geq 0$ ?* We begin by introducing  $\lambda$ , the loglikelihood ratio statistic, and outlining how it can be used to choose a Hilbert space

ref?

dimension. Then, we show how and why the standard model of its behavior, the Wilks theorem, falls apart in the presence of positivity constraints. We propose a new theory that explicitly accounts for state space boundaries, and we derive a closed-form approximation for  $\lambda$ 's expected value under the same conditions as the Wilks theorem – but *with* the  $\rho \geq 0$  constraint. Finally, we test the validity of our approximate theory under harsh conditions by comparing it to numerical results for the realistic problem of optical heterodyne tomography.

Is there some figure we could draw here?

## I. INTRODUCTION - STATISTICAL MODEL SELECTION

Discussing model selection for state tomography requires introducing some basic statistical notions. A *model* is a parameterized family of probability distributions over some data  $D$ , usually denoted as  $\text{Pr}_{\theta}(D)$ , where  $\theta$  are the *parameters* in the model. In quantum state tomography, the data are the outcomes of the measurements of a positive operator-valued measure (POVM)  $\{E_j\}$ , the parameters are a quantum state  $\rho$ , and the probability of obtaining outcome “ $j$ ” is given by the Born rule:  $p_j = \text{Tr}(\rho E_j)$  [? ]. In this paper, a model is a set of density matrices, and  $\rho$  indicates a particular choice of the model's parameters.

Suppose we have some data obtained from an unknown state  $\rho_0$ , and two candidate models  $\mathcal{M}_1, \mathcal{M}_2$  that could be used to fit it. Many of the known methods for choosing between them (i.e., model selection) involve quantifying how well each model fits the data by its *likelihood*. The likelihood of a simple hypothesis  $\rho$  is defined as  $\mathcal{L}(\rho) = \text{Pr}(\text{Data}|\rho)$ . Models, however, are *composite* hypotheses, comprised of many possible values of  $\rho$ . A canonical way to define model  $\mathcal{M}$ 's likelihood is via the general method of *maximum likelihood* (ML), by maximizing  $\mathcal{L}(\rho)$  over  $\rho \in \mathcal{M}$ . In practice, the maximization is usually done

explicitly to find an ML estimate  $\hat{\rho}_{\text{ML},\mathcal{M}}$  [? ? ?] of  $\mathcal{M}$ 's parameters, and then  $\mathcal{L}(\mathcal{M}) = \mathcal{L}(\hat{\rho}_{\text{ML},\mathcal{M}})$ . (Note: although it is common to refer to  $\hat{\rho}_{\text{ML}}$  without specifying the model over which  $\mathcal{L}$  was maximized, we list the model explicitly in the subscript because in this paper we are constantly maximizing over different models!)

Within this framework, models can be compared using the *loglikelihood ratio* [? ? ?]:

$$\lambda(\mathcal{M}_1, \mathcal{M}_2) = -2 \log \left( \frac{\mathcal{L}(\mathcal{M}_1)}{\mathcal{L}(\mathcal{M}_2)} \right) \quad (1)$$

$$= -2 \log \left( \frac{\mathcal{L}(\hat{\rho}_{\text{ML},\mathcal{M}_1})}{\mathcal{L}(\hat{\rho}_{\text{ML},\mathcal{M}_2})} \right) \quad (2)$$

$$= -2 \log \left( \frac{\max_{\rho \in \mathcal{M}_1} \mathcal{L}(\rho)}{\max_{\rho \in \mathcal{M}_2} \mathcal{L}(\rho)} \right). \quad (3)$$

All else being equal, a positive  $\lambda$  favors  $\mathcal{M}_2$  – i.e., the model with the higher likelihood is more plausible, because it fits the data better. However, all else is rarely equal. If both models are equally valid – e.g. they both contain  $\rho_0$  – but  $\mathcal{M}_2$  has more parameters, then  $\mathcal{M}_2$  will very probably fit the data better. Models with more adjustable parameters do a better job of fitting *noise* (e.g. finite sample fluctuations) in the data. This becomes strictly true when the models are *nested*, so that  $\mathcal{M}_1 \subset \mathcal{M}_2$ . In this case, the maximum likelihood of  $\mathcal{M}_2$  is at least as high as that of  $\mathcal{M}_1$ . Not only is  $\lambda \geq 0$ , but almost surely  $\lambda > 0$ . Remarkably, the same effect also makes  $\mathcal{M}_2$ 's fit less accurate (almost surely), because that fit incorporates more of the noise in the data. These two effects constitute *overfitting*, which can be summed up as “Extra parameters make the fit look better, but perform worse.” This provides strong motivation to correct for overfitting by handicapping larger models, to prevent them from being chosen over smaller models that are no less valid, and may even yield better estimates in practice [?].

For this reason, any model selection method that relies (explicitly or implicitly) on a statistic to quantify “how well model  $\mathcal{M}$  fits the data” also relies on a *null theory* to predict how that statistic will behave if  $\rho_0 \in \mathcal{M}$ . A model selection criterion based on an invalid null theory (or a criterion used in a context where its null theory does not apply) will tend to choose the wrong model.

The null theory may be used to formulate a *decision rule* for choosing between two models. By knowing how the statistic behaves when both models are equally good, we may then compare the observed value of the statistic to the null theory. If the observed value is very improbable under the null theory, then that constitutes evidence against the smaller model, and justifies rejecting it. On the other hand, if the observed value is *consistent* with the null theory, there is no reason to reject the smaller model.

The standard null theory for  $\lambda$ , the *Wilks theorem* [?], relies on *local asymptotic normality* (LAN) [? ?]. LAN means that: (1) as  $N_{\text{samples}} \rightarrow \infty$ ,  $\hat{\rho}_{\text{ML}}$  is normally

distributed around  $\rho_0$  with covariance matrix  $\mathcal{I}^{-1}$ , and (2) the likelihood function in a neighborhood of  $\hat{\rho}_{\text{ML}}$  is locally Gaussian with Hessian  $\mathcal{I}$ , where  $\mathcal{I}$  is the classical *Fisher information matrix* associated with the POVM. It quantifies how much information the data carries about a parameter in the model. Generally speaking, the Fisher information depends strongly on  $\rho_0$  and the POVM being measured.

Most statistical models satisfy LAN. When LAN holds *and* the number of samples is large enough to reach the “asymptotic” regime, the *Wilks Theorem* states that if  $\rho_0 \in \mathcal{M}_1 \subset \mathcal{M}_2$ , where  $\mathcal{M}_1$  has  $k$  free parameters and  $\mathcal{M}_2$  has  $K + k$  free parameters, then  $\lambda$  is a  $\chi_K^2$  random variable. This is a complete null theory for  $\lambda$  (under the specified conditions), and implies that  $\langle \lambda \rangle = K$  and  $\Delta \lambda = \sqrt{2K}$ .

Therefore, in the “Wilks regime”, a simple rule for model selection would be the following: compare the observed value of  $\lambda$  to  $\lambda_{\text{thresh}} = \langle \lambda \rangle + n \Delta \lambda$ , for some  $n \approx 1$ , and reject the smaller model if  $\lambda > \lambda_{\text{thresh}}$ . While model selection rules can be more subtle and complex than this, they usually take the general form of a threshold in which  $\langle \lambda \rangle$  plays a key role. Rather than attempting to define a specific rule, our focus in this paper is to understand the behavior of  $\langle \lambda \rangle$  and derive an approximate expression for it in the context of state tomography.

The first step in doing so is to explain how, and why, the Wilks theorem can break down for state tomography.

Think about citations here, somewhat carefully

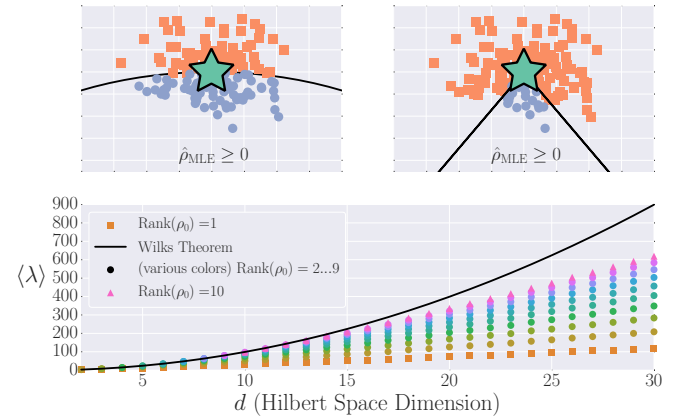


FIG. 1. Impact of the boundary on maximum likelihood tomography. **Top:** Two views through the qutrit state space. Without the positivity constraint, some estimates (orange squares) are not positive semidefinite, and do not represent valid estimates of a quantum state. The distribution of  $\hat{\rho}_{\text{MLE}}$  (blue circles) is generally non-normal, and depends on the true state  $\rho_0$  (star). **Bottom:** Comparison of the classical theory (Wilks Theorem) prediction for the loglikelihood ratio  $\langle \lambda \rangle$  to numerical data for states  $\rho_0$  with ranks  $r = 1, \dots, 10$ . The Wilks Theorem fails badly for low-rank states; our main result (Equation ??) fixes this problem (see Figure ??).

## II. QUANTUM STATE TOMOGRAPHY AND MODEL SELECTION

Quantum state tomography typically begins with  $N$  independently and identically prepared quantum systems – i.e.,  $N$  copies of an unknown state  $\rho_0$ . Each copy is measured, and without loss of generality we can assume that each measurement is described by the same positive operator-valued measure (POVM). A POVM is a collection of positive operators  $\{E_j\}$  summing to  $\mathbb{1}$ , and the probability of outcome “ $j$ ” is given by  $\text{Tr}(\rho_0 E_j)$ . The results of all  $N$  measurements constitute data: a record of the various possible outcomes’ frequencies  $\{n_j\}$ , where  $n_j$  is the number of times “ $j$ ” was observed, and  $\sum_j n_j = N$ . Finally, this data is processed through some *estimator* to yield an estimate  $\hat{\rho}$  of  $\rho_0$ . A variety of estimators have been proposed [? ? ? ? ? ? ? ?], but this is not our concern here. However, since we are concerned with computing the likelihood of a model  $\mathcal{M}$ , which is defined as likelihood of the most likely  $\rho \in \mathcal{M}$ , we will make extensive use of the *maximum likelihood* (ML) estimator  $\hat{\rho}_{\text{ML}, \mathcal{M}}$ . This should not be taken as advocacy for the ML estimate; it is only a convenient way to find the maximum of  $\mathcal{L}$  over  $\mathcal{M}$  as  $\mathcal{L}(\hat{\rho}_{\text{ML}, \mathcal{M}})$ , and once a model is chosen, a different estimator could be used.

The likelihood  $\mathcal{L}(\rho)$  is

$$\mathcal{L}(\rho) = \prod_j \text{Tr}(\rho E_j)^{n_j}, \quad (4)$$

so  $\hat{\rho}_{\text{ML}, \mathcal{M}}$  is the solution to the optimization problem

$$\hat{\rho}_{\text{ML}, \mathcal{M}} = \underset{\rho \in \mathcal{M}}{\text{argmax}} \mathcal{L}(\rho). \quad (5)$$

In state tomography,  $\mathcal{M}$  is almost always the set of all density matrices over a Hilbert space  $\mathcal{H}$ :

$$\mathcal{M}_{\mathcal{H}} = \{\rho \mid \rho \in \mathcal{B}(\mathcal{H}), \text{Tr}(\rho) = 1, \rho \geq 0\}, \quad (6)$$

where  $\mathcal{B}(\mathcal{H})$  is the space of bounded (linear) operators on  $\mathcal{H}$ . To solve the optimization problem, we can use the following facts: (a)  $\mathcal{M}_{\mathcal{H}}$  is a convex set, and (b)  $\hat{\rho}_{\text{ML}, \mathcal{M}}$  also maximizes the value of  $\log(\mathcal{L}(\rho))$ , which is a convex function. Because  $\hat{\rho}_{\text{ML}, \mathcal{M}}$  is the solution to an optimization problem involving a convex function defined over a convex set, it can be found efficiently via any of several algorithms for convex optimization [? ].

Usually,  $\mathcal{H}$  is taken for granted or chosen by fiat. In this paper, we will consider a nested family of different Hilbert spaces, indexed by their dimension  $d$ :  $\mathcal{H}_1 \subset \dots \subset \mathcal{H}_d \subset \mathcal{H}_{d+1} \subset \dots$ . The models we consider are therefore given by:

$$\mathcal{M}_d \equiv \mathcal{M}_{\mathcal{H}_d} = \{\rho \mid \rho \in \mathcal{B}(\mathcal{H}_d), \text{Tr}(\rho) = 1, \rho \geq 0\}. \quad (7)$$

To select between these models, we need to determine whether one model (say,  $\mathcal{M}_{d+1}$ ) is “better” than another (say,  $\mathcal{M}_d$ ). Clearly, the quality of a model corresponds to how well it fits the data, which means we naively would

want to choose a model with highest value of  $\mathcal{L}(\mathcal{M})$ . But to avoid overfitting, it is necessary to penalize or handicap larger models, since (all else being equal) they will *always* fit the data better. Doing so requires a null theory predicting how our test statistic –  $\lambda(\mathcal{M}_d, \mathcal{M}_{d+1})$ , the log-likelihood ratio between  $\mathcal{M}_d$  and  $\mathcal{M}_{d+1}$  – behaves when both models are valid, i.e. when  $\rho_0 \in \mathcal{M}_d, \mathcal{M}_{d+1}$ .

In classical statistics, the null theory for  $\lambda$  (the Wilks theorem) relies on local asymptotic normality (LAN). If LAN holds, the likelihood function near  $\hat{\rho}_{\text{ML}}$  is given by

$$\mathcal{L}(\rho) \propto \text{Exp}[-(\rho - \hat{\rho}_{\text{ML}})\mathcal{I}(\rho - \hat{\rho}_{\text{ML}})/2], \quad (8)$$

and the distribution of  $\hat{\rho}_{\text{ML}}$  is

$$\text{Pr}(\hat{\rho}_{\text{ML}}) \propto \text{Exp}[-(\rho_0 - \hat{\rho}_{\text{ML}})\mathcal{I}(\rho_0 - \hat{\rho}_{\text{ML}})/2], \quad (9)$$

where  $\mathcal{I}$  is the *Fisher information matrix*. (Note that in expressions involving  $\mathcal{I}$ , states  $\rho$  are treated as vectors in state space, and  $\mathcal{I}$  is a matrix or 2-index tensor acting on that state space). In classical statistics, it is common to assume that boundaries are not relevant, either because the models of interest have none, or by assuming that the true parameter values  $\rho_0$  lie far away from them. In the absence of boundaries, and in the asymptotic limit where curvature of the Fisher metric is also negligible, many statistics calculations can be simplified by changing to *Fisher-adjusted* coordinates in which the Fisher information is isotropic (i.e.,  $\mathcal{I} \propto \mathbb{1}$ ).

But the statistical models used for quantum state tomography *do* have boundaries. When  $\rho_0 \in \mathcal{M}_d$  and has rank  $r = d$ , LAN will hold – which is to say that, asymptotically, the boundary can indeed be ignored. In that case, the Wilks Theorem applies, and  $\lambda(\mathcal{M}_d, \mathcal{M}_{d+1}) \sim \chi^2_{2d+1}$ . But when  $\rho_0$  has rank  $r < d$ , it lies *on* the boundary of the model (see Figure ??). LAN does not hold, because even asymptotically, positivity constrains  $\hat{\rho}_{\text{ML}}$ , and so  $\text{Pr}(\hat{\rho}_{\text{ML}})$  is not Gaussian. The Wilks theorem does not apply in this case. Moreover, this is *the* relevant situation for our analysis, because *even if  $\rho_0$  is full-rank in  $\mathcal{M}_d$ , it must be rank-deficient in  $\mathcal{M}_{d+1}$* . So we are forced to derive a replacement for the Wilks theorem, a null theory for  $\lambda$  when  $\rho_0$  is rank-deficient.

The distribution of  $\hat{\rho}_{\text{ML}}$  is complicated, highly non-Gaussian, and singular (because estimates “pile up” on the various faces of the boundary). For this reason, we will not attempt to compute  $\text{Pr}(\lambda)$  directly. Instead, we focus on deriving a good approximation for  $\langle \lambda \rangle$ .

Furthermore, the Fisher information  $\mathcal{I}$  is generally anisotropic. And while the constraint  $\rho \geq 0$  that invalidates LAN in the first place is at least somewhat tractable in standard (Hilbert-Schmidt) coordinates, it becomes completely intractable in Fisher-adjusted coordinates. So, to obtain a semi-analytic null theory for  $\lambda$ , we will make the simplifying assumption that  $\mathcal{I}$  is proportional to the Hilbert-Schmidt metric ( $\mathcal{I} \propto \mathbb{1}$ ).

The main section of this paper is devoted to approximating  $\langle \lambda \rangle$  under the assumption  $\mathcal{I} \propto \mathbb{1}$ ; the main result is given in Equation (??). Then, in the last section of

This isn't obvious to me

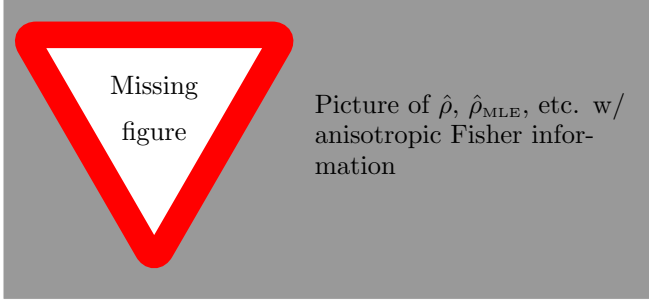
Get the references RBK is citing

fix this, and split

Write out that transformation?

See Figure ??

the paper, we check the robustness of the results that we derive with this simplifying assumption, by performing a numerical study of a tomography problem where  $\mathcal{I}$  is *extremely* anisotropic and comparing the results to the predictions derived with  $\mathcal{I} \propto \mathbb{I}$ .



### III. DERIVING A REPLACEMENT FOR THE WILKS THEOREM

In this section, we show how to derive a replacement for the Wilks Theorem which is valid for rank-deficient  $\rho_0$ . To do so, we start by utilizing the fact that, if quantum state space didn't have any boundaries (i.e, we removed the positivity constraint from our estimator), then LAN would hold. We will focus on the special and simple case where  $\mathcal{I}$  at  $\rho_0$  is proportional to the Hilbert-Schmidt metric, so the likelihood function (and the distribution of the *unconstrained* MLEs  $\hat{\rho}$ ) is given by

We've used MLE = maximum likelihood estimation. Is it confusing to do MLE = maximum likelihood estimate?

$$\mathcal{L}(\rho) \propto e^{-\text{Tr}[(\rho - \hat{\rho})^2]/2\epsilon^2} \quad (10)$$

$$\text{Pr}(\hat{\rho}) \propto e^{-\text{Tr}[(\rho_0 - \hat{\rho})^2]/2\epsilon^2} \quad (11)$$

for some  $\epsilon$  that scales as  $1/\sqrt{N_{\text{samples}}}$ . The isotropic assumption greatly simplifies our study of the problem, and permits the derivation of analytic results which capture realistic tomographic scenarios surprisingly well [?]. Because it is unconstrained,  $\hat{\rho}$  is often not positive semidefinite (see Figure ??). For each  $\hat{\rho}$ , the MLE in  $\mathcal{M}_d$ ,  $\hat{\rho}_{\text{MLE},d}$ , is the solution to the following optimization problem:

$$\hat{\rho}_{\text{MLE},d} = \underset{\rho \in \mathcal{M}_d}{\text{argmin}} \text{Tr}[(\hat{\rho} - \rho)^2]. \quad (12)$$

In turn,  $\lambda(\mathcal{M}_d, \mathcal{M}_{d+1})$  is given by

$$\lambda(\mathcal{M}_d, \mathcal{M}_{d+1}) = \frac{\text{Tr}[(\hat{\rho}_{\text{MLE},d} - \hat{\rho})^2] - \text{Tr}[(\hat{\rho}_{\text{MLE},d+1} - \hat{\rho})^2]}{\epsilon^2}. \quad (13)$$

Our derivation of  $\langle \lambda \rangle$  involves three main steps:

- Reduce calculating  $\langle \lambda(\mathcal{M}_d, \mathcal{M}_{d+1}) \rangle$  to that of computing mean-squared errors  $\langle \text{Tr}[(\hat{\rho}_{\text{MLE},d} - \rho_0)^2] \rangle$ .

- Separate out degrees of freedom in  $\hat{\rho}$  which are unaffected by the positivity constraint, and those which are.
- For the degrees of freedom of  $\hat{\rho}$  which are affected by the positivity constraint (which will turn out to be *spectral* in nature), derive an analytic approximation to the optimization problem given in Equation (??).

#### A. Relating $\langle \lambda \rangle$ to Mean-Squared Error

To simplify our derivation, we first show how we may reduce the problem of computing the expected value of  $\lambda$  given in Equation (??) to that of computing the mean-squared error between  $\hat{\rho}_{\text{MLE},d}$  and  $\rho_0$ . We start by relating  $\lambda(\mathcal{M}_d, \mathcal{M}_{d+1})$  to  $\lambda(\rho_0, \mathcal{M}_d)$  and  $\lambda(\rho_0, \mathcal{M}_{d+1})$  using the identity

$$\lambda(\mathcal{M}_d, \mathcal{M}_{d+1}) = \lambda(\rho_0, \mathcal{M}_{d+1}) - \lambda(\rho_0, \mathcal{M}_d), \quad (14)$$

where

$$\begin{aligned} \lambda(\rho_0, \mathcal{M}_d) &= -2 \log \left( \frac{\mathcal{L}(\rho_0)}{\max_{\rho \in \mathcal{M}_d} \mathcal{L}(\rho)} \right) \\ &= (\text{Tr}[(\rho_0 - \hat{\rho})^2] - \text{Tr}[(\hat{\rho}_{\text{MLE}} - \hat{\rho})^2])/\epsilon^2. \end{aligned} \quad (15)$$

This identity may be interpreted as saying that when evaluating the plausibility of two models, we may compare each to any *reference model*. Here, we take that reference model to be  $\rho_0$  itself.

Now, we derive an equivalent expression for  $\lambda(\rho_0, \mathcal{M}_d)$ . To do so, we show that replacing the local state space with a *tangent cone* at  $\rho_0$ , denoted  $C(\rho_0)$  allows us to ???

What does it allow us to do, succinctly?

Check these definitions

**Definition:** Given a state  $\rho_0 \in \mathcal{M}_d$ , the tangent cone at  $\rho_0$ , denoted  $C(\rho_0)$  is given by the closure of the set of rays emanating from  $\rho_0$  to any  $\sigma \in \mathcal{M}_d$ .

For example, if  $\rho_0$  is a pure state in  $\mathcal{M}_2$ , then  $C(\rho_0)$  is the halfspace which contains the Bloch sphere, and whose boundary touches the Bloch sphere at  $\rho_0$ . If  $\rho_0$  is a mixed state in  $\mathcal{M}_2$ ,  $C(\rho_0)$  is  $\mathbb{R}^3$ .

Notice that  $\mathcal{M}_d \subset C$ , by construction. With the tangent cone defined, we now define the maximum likelihood estimate within it.

**Definition** Given  $\hat{\rho}$ , and the tangent cone  $C(\rho_0)$ , define  $\hat{\rho}_{\text{MLE},C}$  as the solution to the following optimization problem:

$$\hat{\rho}_{\text{MLE},C} = \underset{M \in C}{\text{argmin}} \text{Tr}[(M - \hat{\rho})^2] \quad (16)$$

Note that  $\hat{\rho}_{\text{MLE},C}$  need not be positive semi-definite!



**Definition** Define the loglikelihood ratio statistic comparing  $\rho_0$  to the model  $C(\rho_0)$  as

$$\lambda(\rho_0, C) \equiv -2 \log \left( \frac{\mathcal{L}(\rho_0)}{\mathcal{L}(\hat{\rho}_{\text{MLE}, C})} \right) \\ = (\text{Tr}[(\rho_0 - \hat{\rho})^2] - \text{Tr}[(\hat{\rho}_{\text{MLE}, C} - \hat{\rho})^2]) / \epsilon^2 \quad (17)$$

Do we need a figure for this?

It turns out  $\lambda(\rho_0, C)$  has a particularly pleasing form.

**Theorem**  $\lambda(\rho_0, C) = \text{Tr}[(\rho_0 - \hat{\rho}_{\text{MLE}, C})^2]$ .

**Proof**

*Case 1:* Assume  $\hat{\rho} \notin C$ . Because  $\hat{\rho}_{\text{MLE}, C}$  is the closest point to  $\hat{\rho}$  in  $C$ , and because  $C$  contains  $\rho_0$ , the lines joining  $\rho_0$  to  $\hat{\rho}_{\text{MLE}, C}$ , and  $\hat{\rho}_{\text{MLE}, C}$  to  $\hat{\rho}$ , are at right angles to one another. (See Figure ??) Applying the Pythagorean theorem, we have

$$\text{Tr}[(\rho_0 - \hat{\rho})^2] = \text{Tr}[(\rho_0 - \hat{\rho}_{\text{MLE}, C})^2] + \text{Tr}[(\hat{\rho} - \hat{\rho}_{\text{MLE}, C})^2] \quad (18)$$

Dividing Equation (??) by  $\epsilon^2$ , and combining with Equation (??), yields the theorem statement.

*Case 2:* Assume  $\hat{\rho} \in C$ . Then,  $\hat{\rho}_{\text{MLE}, C} = \hat{\rho}$ , and Equation (??) simplifies to the theorem statement.

Finally, it turns out that the values of  $\lambda(\rho_0, \mathcal{M}_d)$  and  $\lambda(\rho_0, C)$  are close to one another.

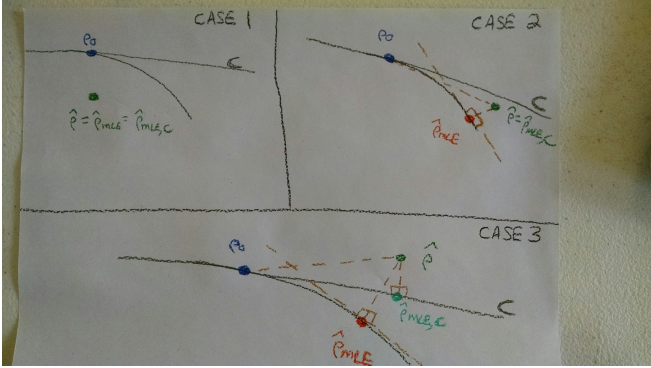


FIG. 2.

Replace this

**Theorem**  $|\langle \lambda(\rho_0, \mathcal{M}_d) \rangle - \langle \lambda(\rho_0, C) \rangle| \leq 1$

**Proof** We consider three cases:

*Case 1:*  $\hat{\rho} \in \mathcal{M}_d$ . In this case,  $\hat{\rho}_{\text{MLE}} = \hat{\rho}_{\text{MLE}, C} = \hat{\rho}$ , so  $\lambda(\rho_0, \mathcal{M}_d) - \lambda(\rho_0, C) = 0$ .

*Case 2:*  $\hat{\rho} \notin \mathcal{M}_d$ , but  $\hat{\rho} \in C$ . In this case,  $\hat{\rho}_{\text{MLE}, C} = \hat{\rho}$ , so  $\epsilon^2 \lambda(\rho_0, C) = \text{Tr}[(\rho_0 - \hat{\rho}_{\text{MLE}, C})^2]$ , and  $\epsilon^2 \lambda(\rho_0, \mathcal{M}_d) = \text{Tr}[(\rho_0 - \hat{\rho}_{\text{MLE}, C})^2] - \text{Tr}[(\hat{\rho}_{\text{MLE}} - \hat{\rho}_{\text{MLE}, C})^2]$ . So we have

$$\lambda(\rho_0, \mathcal{M}_d) - \lambda(\rho_0, C) = -\frac{1}{\epsilon^2} \text{Tr}[(\hat{\rho}_{\text{MLE}} - \hat{\rho}_{\text{MLE}, C})^2]$$

Consider the right triangle shown in Figure ??, Case 2. By the Pythagorean theorem, we have

$$\text{Tr}[(\hat{\rho}_{\text{MLE}} - \hat{\rho}_{\text{MLE}, C})^2] \leq \text{Tr}[(\hat{\rho} - \rho')^2]$$

for some  $\rho'$  which lies on the line joining  $\rho_0$  and  $\hat{\rho}_{\text{MLE}, C}$ . Further, it is clear that  $\text{Tr}[(\rho_0 - \rho')^2] \leq \text{Tr}[(\rho_0 - \hat{\rho}_{\text{MLE}, C})^2]$ . In turn, this means

$$\langle \text{Tr}[(\hat{\rho}_{\text{MLE}} - \hat{\rho}_{\text{MLE}, C})^2] \rangle \leq \langle \text{Tr}[(\rho_0 - \hat{\rho}_{\text{MLE}, C})^2] \rangle = \epsilon^2 \\ \Rightarrow -\frac{\langle \text{Tr}[(\hat{\rho}_{\text{MLE}} - \hat{\rho}_{\text{MLE}, C})^2] \rangle}{\epsilon^2} \geq 1$$

Thus,  $\langle \lambda(\rho_0, \mathcal{M}_d) \rangle - \langle \lambda(\rho_0, C) \rangle \geq -1$ .

*Case 3:*  $\hat{\rho} \notin \mathcal{M}_d, C$ . In this case,  $\epsilon^2 \lambda(\rho_0, C) = \text{Tr}[(\rho_0 - \hat{\rho}^2)] - \text{Tr}[(\hat{\rho}_{\text{MLE}, C} - \hat{\rho})^2]$  and  $\epsilon^2 \lambda(\rho_0, \mathcal{M}_d) = \text{Tr}[(\rho_0 - \hat{\rho})^2] - \text{Tr}[(\hat{\rho}_{\text{MLE}} - \hat{\rho})^2]$ , so

$$\lambda(\rho_0, \mathcal{M}_d) - \lambda(\rho_0, C) = \frac{1}{\epsilon^2} [\text{Tr}[(\hat{\rho}_{\text{MLE}, C} - \hat{\rho})^2] - \text{Tr}[(\hat{\rho}_{\text{MLE}} - \hat{\rho})^2]]$$

Do the simple thing!

$$\lambda(\rho_0, \mathcal{M}_d) - \lambda(\rho_0, C) \leq \frac{1}{\epsilon^2} [\text{Tr}[(\hat{\rho}_{\text{MLE}, C} - \hat{\rho})^2]]$$

Consider the right triangle shown in Figure ??, Case 3. By the Pythagorean theorem we have

$$\text{Tr}[(\hat{\rho}_{\text{MLE}, C} - \hat{\rho})^2] \leq \text{Tr}[(\hat{\rho} - \rho')^2]$$

For some  $\rho'$  lying on the line joining  $\rho_0$  and  $\hat{\rho}_{\text{MLE}, C}$ . Further, it is clear that  $\text{Tr}[(\rho_0 - \rho')^2] \leq \text{Tr}[(\rho_0 - \hat{\rho}_{\text{MLE}, C})^2]$ . Chaining these inequalities, and taking expected values, yields

$$\langle \lambda(\rho_0, \mathcal{M}_d) \rangle - \langle \lambda(\rho_0, C) \rangle \leq \frac{1}{\epsilon^2} \langle [\text{Tr}[(\rho_0 - \hat{\rho})^2]] \rangle \leq 1$$

Combining these three cases by taking absolute values yields the theorem statement.

To summarize, through these identities and approximations, we have reduced the problem of computing  $\lambda(\mathcal{M}_d, \mathcal{M}_{d+1})$  to that of computing the (scaled) mean-squared error  $\text{Tr}[(\hat{\rho}_{\text{MLE}} - \rho_0)^2] / \epsilon^2$  for each  $\hat{\rho}_{\text{MLE}, d}$  in both  $\mathcal{M}_d$  and  $\mathcal{M}_{d+1}$  by replacing the local state space with a tangent cone. Now, we must determine an approximate form for  $\langle \text{Tr}[(\hat{\rho}_{\text{MLE}} - \rho_0)^2] / \epsilon^2$ .

## B. Separating out Degrees of Freedom in $\hat{\rho}$

We begin by observing that  $\lambda(\rho_0, \mathcal{M}_d)$  can be written as a sum over matrix elements,

$$\lambda = \epsilon^{-2} \text{Tr}[(\hat{\rho}_{\text{MLE}} - \rho_0)^2] = \epsilon^{-2} \sum_{jk} |(\hat{\rho}_{\text{MLE}} - \rho_0)_{jk}|^2 \\ = \sum_{jk} \lambda_{jk} \quad \text{where} \quad \lambda_{jk} = \epsilon^{-2} |(\hat{\rho}_{\text{MLE}} - \rho_0)_{jk}|^2, \quad (19)$$

and therefore  $\langle \lambda \rangle = \sum_{jk} \langle \lambda_{jk} \rangle$ . Each term  $\langle \lambda_{jk} \rangle$  quantifies the average mean-squared error of a single matrix element of  $\hat{\rho}_{\text{MLE}}$ , and while the Wilks Theorem predicts  $\langle \lambda_{jk} \rangle = 1$  for all  $j, k$ , numerical simulations (see Figure

??) show that this only holds true for *some* matrix elements. A few contribute more than 1 unit (on average) while many others contribute much less, meaning that the Wilks Theorem predicts too high a value for the total  $\langle \lambda \rangle$ . (See bottom of Figure ??.) Thus motivated, we divide the parameters of  $\hat{\rho}$  into two parts (see Figure ??),

1. The “kite” comprises all diagonal elements *and* all elements on the kernel (null space) of  $\rho_0$ ,
2. The “L” comprises all off-diagonal elements on the support of  $\rho_0$  *and* between the support and the kernel,

and observe that  $\langle \lambda \rangle = \langle \lambda_L \rangle + \langle \lambda_{\text{kite}} \rangle$ . The rationale for this division is simple: small fluctuations on the “L” do not change the zero eigenvalues of  $\hat{\rho}$  to 1<sup>st</sup> order, whereas those on the “kite” do. In what follows, we study how imposing the positivity constraint  $\hat{\rho}_{\text{MLE}} \geq 0$  affects the behavior of the matrix elements in each part.

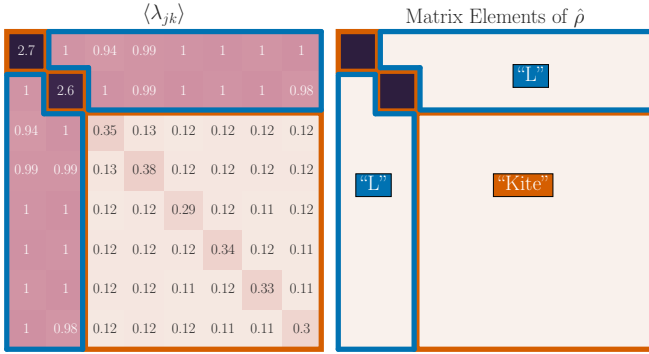


FIG. 3. When a rank-2 state is reconstructed in  $d = 8$  dimensions, the total loglikelihood ratio  $\lambda(\rho_0, \mathcal{M}_8)$  is the sum of terms  $\lambda_{jk}$  from errors in each matrix element  $(\hat{\rho}_{\text{MLE}})_{jk}$ . **Left:** Numerics show a clear division; some matrix elements have  $\langle \lambda_{jk} \rangle \sim 1$  as predicted by the Wilks Theorem, while others are either more or less. **Right:** The numerical results motivate dividing the elements of  $\hat{\rho}$  into two parts: the “kite” and the “L”.

In doing so, it is helpful to think about the error of the unconstrained estimate  $\delta \equiv \hat{\rho} - \rho_0$ , a normally-distributed *traceless* matrix. To simplify the analysis, we explicitly drop the  $\text{Tr}(\rho) = 1$  constraint and let  $\delta$  be  $\mathcal{N}(0, \epsilon^2 \mathbb{1})$  distributed over the  $d^2$ -dimensional space of Hermitian matrices (a good approximation when  $d \gg 2$ ) [?], which makes  $\delta$  proportional to an element of the Gaussian Unitary Ensemble (GUE) [?].

### 1. Computing $\langle \lambda_L \rangle$

To first order in  $\epsilon$ , elements of  $\delta$  in the “L” do not affect positivity, so they are unconstrained by the boundary, and behave exactly as expected from classical theory. The  $\delta_{jk}$  in the “L” may be seen as errors which arise due to small unitary perturbations of  $\rho_0$ . Writing  $\hat{\rho} = U^\dagger \rho_0 U$ ,

where  $U = e^{i\epsilon H}$ , we have

$$\hat{\rho} \approx \rho_0 + i\epsilon[\rho_0, H] + \mathcal{O}(\epsilon^2).$$

Then,  $\delta \approx i\epsilon[\rho_0, H]$ . If  $j = k$ , then  $\delta_{jj} = 0$ . Thus, small unitaries cannot create errors in the diagonal matrix elements, at  $\mathcal{O}(\epsilon)$ . If  $j \neq k$ , then  $\delta_{jk} \neq 0$ , in general. (Small unitaries *can* introduce errors on off-diagonal elements.)

However, if either  $j$  or  $k$  (or both) lie within the *kernel* of  $\rho_0$  (i.e.,  $\langle k|\rho_0|k \rangle$  or  $\langle j|\rho_0|j \rangle$  is 0), then the corresponding  $\delta_{jk}$  are zero. The only off-diagonal elements where small unitaries can introduce errors are those which are coherent between the kernel of  $\rho_0$  and its support. These off-diagonal elements are precisely the “L”, and are the set  $\{\delta_{jk} \mid \langle j|\rho_0|j \rangle \neq 0, j \neq k, 0 \leq j, k \leq d-1\}$ . Each  $\delta_{jk}$  in the “L” is a  $\mathcal{N}(0, \epsilon^2)$  random variable, and crucially, is identical to the error  $(\hat{\rho}_{\text{MLE}} - \rho_0)_{jk}$ . This is because these  $\delta_{jk}$  are *unaffected by the boundary*, so when we impose the positivity constraint (i.e., compute  $\hat{\rho}_{\text{MLE}}$ ), their values remain the same. Therefore,  $\langle \lambda_{jk} \rangle = \langle \delta_{jk}^2 \rangle / \epsilon^2 = 1$ . As there are  $2rd - r(r+1)$  of them,  $\langle \lambda_L \rangle = 2rd - r(r+1)$ .

### 2. Setting the Stage to Compute $\langle \lambda_{\text{kite}} \rangle$

We next need a procedure to compute  $\hat{\rho}_{\text{MLE}}$  given  $\hat{\rho}$  – i.e., to solve the optimization problem in Eq. (??). Fortunately, an algorithm for doing so was presented in Ref. [?]:

1. Subtract  $q\mathbb{1}$  from the unconstrained  $\hat{\rho}$ , for a particular real scalar  $q$ ,
2. “Truncate”  $\hat{\rho} - q\mathbb{1}$ , by replacing each of its negative eigenvalues with zero.

Here,  $q$  is defined implicitly such that  $\text{Tr}[\text{Trunc}(\hat{\rho} - q\mathbb{1})] = 1$ .

Although this was intended as a (very fast) numerical algorithm, we will manipulate it (by a series of approximations) to derive a closed-form expression for the average  $\langle \lambda \rangle$ .

Computing  $\langle \lambda_{\text{kite}} \rangle$  is a bit harder, because the boundary *does* constrain its elements. Here, we turn to the truncation algorithm given above for finding  $\hat{\rho}_{\text{MLE}}$ , which is most naturally performed in the eigenbasis of  $\hat{\rho}$ . Exact diagonalization of  $\hat{\rho}$  is not feasible analytically, but only the *small* eigenvalues of  $\hat{\rho}$  are critical in truncation. As long as all the nonzero eigenvalues of  $\rho_0$  are much larger than  $\epsilon$ , the eigenbasis of  $\hat{\rho}$  is accurately approximated by: (1) the eigenvectors of  $\rho_0$  on its support; and (2) the eigenvectors of  $\delta_{\text{ker}} = \Pi_{\text{ker}} \delta \Pi_{\text{ker}}$ , where  $\Pi_{\text{ker}}$  is the projector onto the kernel of  $\rho_0$ .

Changing to this basis diagonalizes the “kite” portion of  $\delta$ , and leaves all elements of the “L” unchanged (at  $\mathcal{O}(\epsilon)$ ). The diagonal elements of  $\hat{\rho}$  now fall into two categories:

1.  $r$  elements corresponding to the eigenvalues of  $\rho_0$ , which are given by  $p_j = \rho_{jj} + \delta_{jj}$  where  $\rho_{jj}$  is the  $j^{\text{th}}$  eigenvalue of  $\rho_0$ , and  $\delta_{jj} \sim \mathcal{N}(0, \epsilon^2)$ .

2.  $N \equiv d - r$  elements that are eigenvalues of  $\delta_{\text{ker}}$ , which we denote by  $\kappa = \{\kappa_j : j = 1 \dots N\}$ ,

and  $\lambda_{\text{kite}}$  is

$$\epsilon^2 \lambda_{\text{kite}} = \sum_{j=1}^r [\rho_{jj} - (p_j - q)^+]^2 + \sum_{j=1}^N [(\kappa_j - q)^+]^2, \quad (20)$$

where  $(x)^+ = \max(x, 0)$ .  $q$  is implicitly defined such that  $f(q) \equiv \text{Tr}[\text{Trunc}(\hat{\rho} - q\mathbb{I})]$  satisfies  $f(q) = 1$ . In terms of the eigenvalues of  $\hat{\rho}$ , this means  $q$  is the solution to

$$\sum_{j=1}^r (p_j - q)^+ + \sum_{j=1}^N (\kappa_j - q)^+ = 1 \quad (21)$$

To solve Equation (??), and derive an approximation for (??), we need to understand the behavior of the eigenvalues of  $\delta_{\text{ker}}$ . It is to this problem we now turn.

### 3. Approximating the Eigenvalues of a GUE( $N$ ) Matrix

We first observe that while the  $\kappa_j$  are random variables, they are not normally distributed. Instead, because  $\delta_{\text{ker}}$  is proportional to a GUE( $N$ ) matrix, for  $N \gg 1$ , the distribution of any eigenvalue  $\kappa_j$  converges to a Wigner semicircle distribution [?] given by  $\text{Pr}(\kappa) = \frac{2}{\pi R^2} \sqrt{R^2 - \kappa^2}$  for  $|\kappa| \leq R$ , with  $R = 2\epsilon\sqrt{N}$ . The eigenvalues are not independent; they tend to avoid collisions (“level avoidance” [?]), and typically form a surprisingly regular array over the support of the Wigner semicircle. Since our goal is to compute  $\langle \lambda_{\text{kite}} \rangle$ , we can capitalize on this behavior by replacing each random sample of  $\kappa$  with a *typical sample*  $\bar{\kappa}$  given by its order statistics. These are the average values of the *sorted*  $\kappa$ , so  $\bar{\kappa}_j$  is the average value of the  $j^{\text{th}}$  largest value of  $\kappa$ . Large random samples are usually well approximated (for many purposes) by their order statistics even when the elements of the sample are independent, and level avoidance makes the approximation even better.

Suppose that  $\kappa$  are the eigenvalues of a GUE( $N$ ) matrix, sorted from highest to lowest. Figure ?? illustrates such a sample for  $N = 100$ . It also shows the *average* values of 100 such samples (all sorted). These are the *order statistics*  $\bar{\kappa}$  of the distribution (more precisely, what is shown is a good *estimate* of the order statistics; the actual order statistics would be given by the average over infinitely many samples). The point of the figure is to show that, while the order statistics *are* slightly more smoothly and predictably distributed than a single (sorted) sample, the two are remarkably similar. A single sample  $\kappa$  will fluctuate around the order statistics, but these fluctuations are relatively small, partly because the sample is large, and partly because the GUE eigenvalues experience level repulsion. Thus, the “typical” behavior of a sample – by which we mean the mean value of a statistic of the sample – is well captured by the order statistics (which have no fluctuations at all).

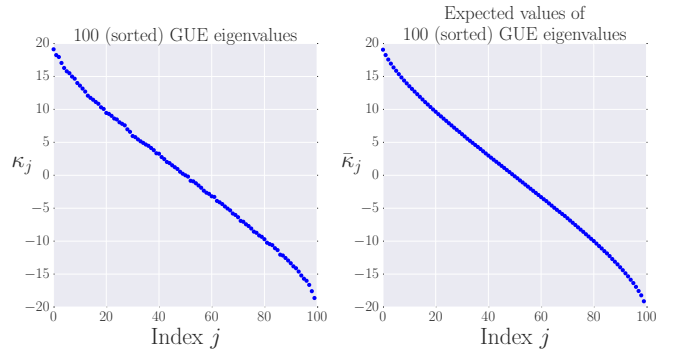


FIG. 4. Typical samples of GUE( $N$ ) eigenvalues are accurately approximated by order statistics of the distribution (average values of a sorted sample). **Left:** The sorted eigenvalues (i.e., order statistics  $\kappa_j$ ) of one randomly chosen GUE(100) matrix. **Right:** Approximate expected values of the order statistics,  $\bar{\kappa}_j$ , of the GUE(100) distribution, computed as the average of the sorted eigenvalues of 100 randomly chosen GUE(100) matrices.

We now turn to the problem of modeling  $\kappa$  quantitatively. We note up front that we are only going to be interested in certain properties of  $\kappa$ : specifically, partial sums of all  $\kappa_j$  greater or less than the threshold  $q$ , or partial sums of functions of the  $\kappa_j$  (e.g.  $(\kappa_j - q)^2$ ). We require only that an ansatz be accurate for such quantities. We do not use this fact explicitly, but it motivates our approach – and we do not claim that our ansatz is accurate for *all* conceivable functions.

In general, if a sample  $\kappa$  of size  $N$  is drawn so that each  $\kappa$  has the same probability density function  $\text{Pr}(\kappa)$ , then a good approximation for the  $j^{\text{th}}$  order statistic is given by the inverse *cumulative* distribution function (CDF):

$$\bar{\kappa}_j \approx \text{CDF}^{-1} \left( \frac{j - 1/2}{N} \right). \quad (22)$$

This is closely related to the observation that the histogram of a sample tends to look similar to the underlying probability density function. More precisely, it is equivalent to the observation that the empirical distribution function (the CDF of the histogram) tends to be (even more) similar to the underlying CDF. (For i.i.d. samples, this is the content of the Glivenko-Cantelli theorem [?]). Figure ?? compares the order statistics of GUE(100) and GUE(10) eigenvalues (computed as numerical averages over 100 random samples) to the inverse CDF for the Wigner semicircle distribution. Even though the Wigner semicircle model of GUE eigenvalues is only exact as  $N \rightarrow \infty$ , it provides a nearly-perfect model for  $\bar{\kappa}$  even at  $N = 10$  (and remains surprisingly good all the way down to  $N = 2$ ).

We make one further approximation, by assuming that  $N \gg 1$ , so the distribution of the  $\bar{\kappa}_j$  is effectively continuous and identical to  $\text{Pr}(\kappa)$ . For the quantities that we compute, this is equivalent to replacing the empirical distribution function (which is a step function) by the

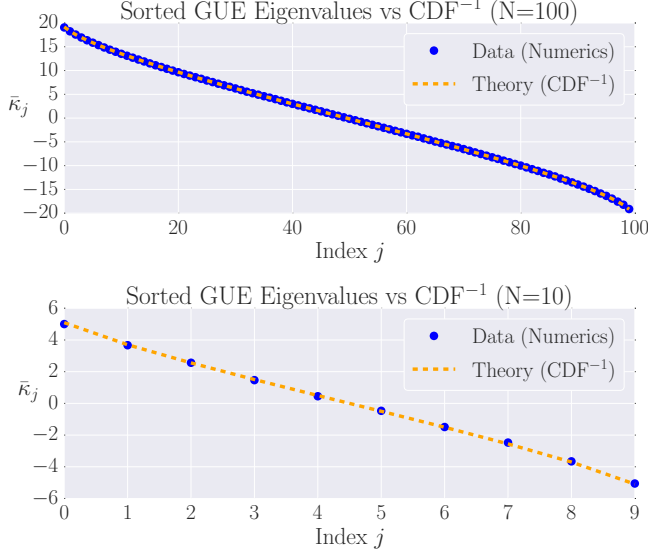


FIG. 5. Order statistics of the GUE( $N$ ) eigenvalue distribution are very well approximated by the inverse CDF of the Wigner semicircle distribution. In both figures, we compare the order statistics of a GUE( $N$ ) distribution to the inverse CDF of the Wigner semicircle distribution. **Top:**  $N = 100$ . **Bottom:**  $N = 10$ . Agreement in both cases is essentially perfect.

CDF of the Wigner semicircle distribution. So, whereas for any given sample the partial sum of all  $\kappa_j > q$  jumps discontinuously when  $q = \kappa_j$  for any  $j$ , in this approximation it changes smoothly. This accurately models the *average* behavior of partial sums.

#### 4. Deriving an Approximation for $q$

The approximations of the previous section allow us to use the following ansatz for the eigenvalues of  $\hat{\rho}$ ; namely,  $\{p_j\} \cup \{\bar{\kappa}_j\}$ , where the  $p_j$  are  $\mathcal{N}(\rho_{jj}, \epsilon^2)$  random variables, and the  $\bar{\kappa}_j$  are the (fixed, smoothed) order statistics of a Wigner semicircle distribution. In turn, the defining equation for  $q$  (Equation (??)) is well approximated as

$$\sum_{j=1}^r (p_j - q)^+ + \sum_{j=1}^N (\bar{\kappa}_j - q)^+ \approx 1 \quad (23)$$

To solve this equation, we observe that the  $\bar{\kappa}_j$  are symmetrically distributed around  $\kappa = 0$ , so half of them are negative. Therefore, with high probability,  $\text{Tr}[\text{Trunc}(\hat{\rho})] > 1$ , and so we will need to subtract  $q$  from  $\hat{\rho}$  before truncating. (This is in distinction to the case where we have to *add*  $q$ .)

We make another assumption; namely, that the eigenvalues of  $\rho_0$  are large compared to the perturbations  $\delta_{jj}$  and  $q$ . This implies  $(p_j - q)^+ = p_j - q$ . Under this

assumption,  $q$  is the solution to

$$\begin{aligned} 1 &\approx \sum_{j=1}^r (p_j - q)^+ + \sum_{j=1}^N (\bar{\kappa}_j - q)^+ \\ &\approx 1 - rq + \Delta + N \int_{\kappa=q}^{2\epsilon\sqrt{N}} (\kappa - q) \text{Pr}(\kappa) d\kappa \\ \implies 0 &= -rq + \Delta + \frac{\epsilon}{12\pi} \left[ (q^2 + 8N) \sqrt{-q^2 + 4N} - 12qN \left( \frac{\pi}{2} - \sin^{-1} \left( \frac{q}{2\sqrt{N}} \right) \right) \right], \end{aligned} \quad (24)$$

where  $\Delta = \sum_{j=1}^r \delta_{jj}$  is a  $\mathcal{N}(0, r\epsilon^2)$  random variable. We choose to replace a discrete sum (line 1) with an integral (line 2). This approximation is valid when  $N \gg 1$ , as we can accurately approximate a discrete collection of closely spaced real numbers by a smooth density or distribution over the real numbers that has approximately the same CDF. It is also remarkably accurate in practice.

In yet another approximation, we replace  $\Delta$  with its average value, which is zero. We could obtain an even more accurate expression by treating the fluctuations in  $\Delta$  more carefully, but this crude approximation turns out to be quite accurate already.

To solve Equation (??), it is necessary to further simplify the complicated expression resulting from the integral (line 3). To do so, we assume  $\rho_0$  is relatively low-rank, so  $r \ll N$ . In this case, the sum of the positive  $\bar{\kappa}_j$  is large compared with  $r$ , almost all of them need to be subtracted away, and therefore  $q$  is close to  $2\epsilon\sqrt{N}$ . [?] We therefore replace the complicated expression with its leading order Taylor expansion around  $q = 2\epsilon\sqrt{N}$ , substitute into Equation (??), and obtain the equation

$$\frac{rq}{\epsilon} = \frac{4}{15\pi} N^{1/4} \left( 2\sqrt{N} - \frac{q}{\epsilon} \right)^{5/2}. \quad (25)$$

This equation is a quintic polynomial, so it has no closed-form solution. However, its roots have a well-defined asymptotic ( $N \rightarrow \infty$ ) expansion that becomes accurate quite rapidly (e.g., for  $N > 4$ ):

$$z \equiv q/\epsilon \approx 2\sqrt{N} - \frac{(240r\pi)^{2/5}}{4} N^{1/10} + \frac{(240r\pi)^{4/5}}{80} N^{-3/10}. \quad (26)$$

Do we want a figure of how well this does?

#### 5. Computing $\langle \lambda_{\text{kite}} \rangle$

Now that we know how much to subtract off in the truncation process, we can approximate  $\langle \lambda_{\text{kite}} \rangle$ , originally



given in Equation (??):

$$\begin{aligned}
\langle \lambda_{\text{kite}} \rangle &\approx \frac{1}{\epsilon^2} \left\langle \sum_{j=1}^r [\rho_{jj} - (p_j - q)^+]^2 + \sum_{j=1}^N [(\bar{\kappa}_j - q)^+]^2 \right\rangle \\
&\approx \frac{1}{\epsilon^2} \left\langle \sum_{j=1}^r [-\delta_{jj} + q]^2 + \sum_{j=1}^N [(\bar{\kappa}_j - q)^+]^2 \right\rangle \\
&\approx r + rz^2 + \frac{N}{\epsilon^2} \int_{\kappa=q}^{2\epsilon\sqrt{N}} \text{Pr}(\kappa)(\kappa - q)^2 d\kappa \\
&= r + rz^2 + \frac{N(N+z^2)}{\pi} \left( \frac{\pi}{2} - \sin^{-1} \left( \frac{z}{2\sqrt{N}} \right) \right) \\
&\quad - \frac{z(z^2 + 26N)}{24\pi} \sqrt{4N - z^2} \quad (27)
\end{aligned}$$

#### IV. RESULT FOR $\langle \lambda \rangle$ , COMPARISON TO NUMERICAL EXPERIMENTS

The total expected value,  $\langle \lambda \rangle = \langle \lambda_L \rangle + \langle \lambda_{\text{kite}} \rangle$ , is thus

$$\begin{aligned}
\langle \lambda(\rho_0, \mathcal{M}_d) \rangle &\approx 2rd - r^2 + rz^2 \\
&\quad + \frac{N(N+z^2)}{\pi} \left( \frac{\pi}{2} - \sin^{-1} \left( \frac{z}{2\sqrt{N}} \right) \right) \\
&\quad - \frac{z(z^2 + 26N)}{24\pi} \sqrt{4N - z^2} \quad (28)
\end{aligned}$$

where  $z$  is given in Equation (??),  $N = d - r$ , and  $r = \text{Rank}(\rho_0)$ .

Equation (??) is our main result. To test its validity, we compare it to numerical simulations for  $d = 2, \dots, 30$  and  $r = 1, \dots, 10$ , in Figure ???. The prediction of the Wilks Theorem is wildly incorrect for  $r \ll d$ . In contrast, Equation (??) is almost perfectly accurate when  $r \ll d$ , but it does begin to break down (albeit fairly gracefully) as  $r$  becomes comparable to  $d$ . We conclude that our analysis (and Equation (??)) correctly models tomography *if* the Fisher information is isotropic ( $\mathcal{I} \propto \mathbb{I}$ ).

#### V. COMPARISON TO HETERODYNE TOMOGRAPHY

In practice, the Fisher information is rarely isotropic. So we tested our idealized result by applying it to a realistic, challenging, and experimentally relevant problem: quantum heterodyne (equivalent to double homodyne) state tomography [?] of a single optical mode. (See Figure ?? for a plot of the *condition number* – the ratio of the largest to smallest eigenvalue – of the estimated Fisher information. It is clear that, for such a tomographic setup,  $\mathcal{I} \not\propto \mathbb{I}$ .) States of this continuous-variable system are described by density operators on the infinite-dimensional Hilbert space  $L^2(\mathbb{R})$ . Fitting these infinitely many parameters to finitely much data demands simpler models.

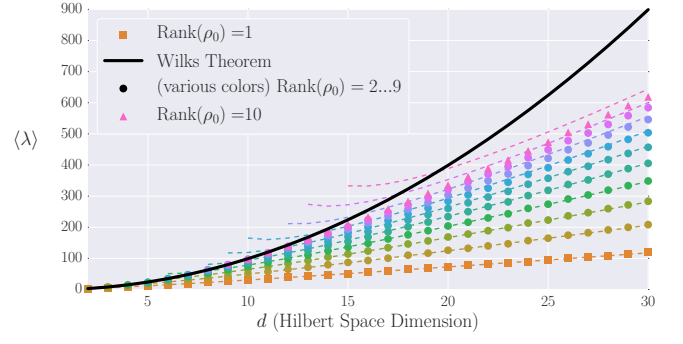


FIG. 6. Numerical results for  $\langle \lambda \rangle$  compared to the prediction of the Wilks Theorem (solid line) and our replacement theory as given in Equation (??), (dashed lines). Our formula depends on the rank  $r$  of  $\rho_0$  (unlike the Wilks prediction), and is nearly perfect for  $r \ll d$ . It becomes less accurate as  $r$  approaches  $d/2$ , and is invalid when  $r \approx d$ .

We consider a family of nested models motivated by a low-energy (few-photon) ansatz, and choose the Hilbert space  $\mathcal{H}_d$  to be that spanned by the photon number states  $\{|0\rangle, \dots, |d-1\rangle\}$ . Heterodyne tomography reconstructs  $\rho_0$  using data from repeated measurements of the coherent-state POVM,  $\{|\alpha\rangle\langle\alpha|/\pi, \alpha = x+ip \in \mathbb{C}\}$ , which corresponds to sampling directly from the state's Husimi  $Q$ -function [?].

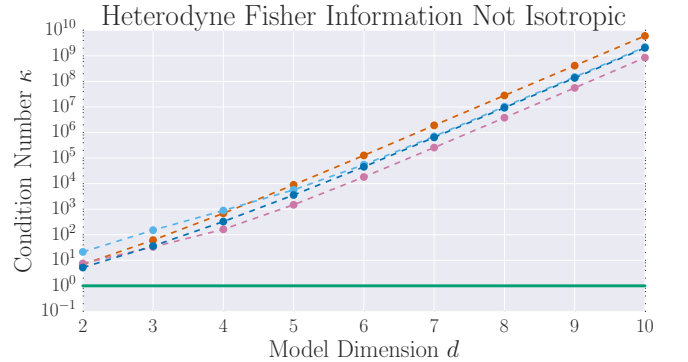


FIG. 7. The condition number  $\kappa$  – the ratio of the largest to smallest eigenvalue – of the estimated heterodyne Fisher information grows with model dimension, indicating an increase in anisotropy. (Estimates are the average over 100 Hessians of the loglikelihood function.) The dashed lines indicate different states  $\rho_0$ , and the solid line is  $\kappa = 1$  (i.e.,  $\mathcal{I} \propto \mathbb{I}$ ).

We examined the behavior of  $\lambda$  for 13 distinct states  $\rho_0$ , both pure and mixed, supported on  $\mathcal{H}_2, \mathcal{H}_3, \mathcal{H}_4$ , and  $\mathcal{H}_5$ . We used rejection sampling to simulate 100 heterodyne datasets with up to  $N_{\text{samples}} = 10^5$ , and found MLEs over each of the 9 models  $\mathcal{M}_2, \dots, \mathcal{M}_{10}$  using numerical optimization [?]. For each  $\rho_0$  and each  $d$ , we averaged  $\lambda(\rho_0, \mathcal{M}_d)$  over all 100 datasets to obtain an empirical average loglikelihood ratio  $\bar{\lambda}$  for each  $(\rho_0, d)$  pair.

Results of this test are shown in Figure ??, where we plot the predictions for  $\langle \lambda \rangle$  given by the Wilks Theorem

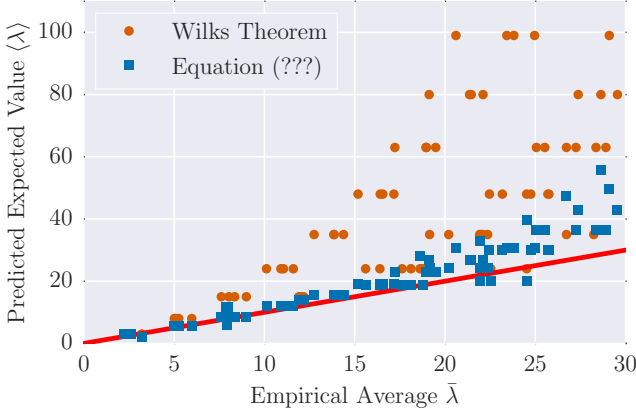


FIG. 8. The Wilks Theorem (orange dots) dramatically overestimates  $\langle \lambda(\rho_0, \mathcal{M}_d) \rangle$  in optical heterodyne tomography. Our formula, Equation ?? (blue squares), is far more accurate. Residual discrepancies occur in large part because  $N_{\text{samples}}$  is not yet “asymptotically large”. The solid red line corresponds to perfect correlation between theory ( $\langle \lambda \rangle$ ) and practice ( $\bar{\lambda}$ ).

and Equation (??), against the empirical average  $\bar{\lambda}$ , for a variety of  $\rho_0$  and  $d$ . Our formula correlates very well with the empirical average, while the Wilks Theorem (unsurprisingly) overestimates  $\lambda$  dramatically for low-rank states. Whereas a model selection procedure based on Wilks Theorem would tend to falsely reject larger Hilbert spaces (by setting the threshold for acceptance too high), our formula provides a reliable null theory.

Interestingly, as  $d$  grows, Equation (??) also begins to overpredict. As Figure ?? indicates, a more accurate description is that the numerical experiments are *under-achieving*, because  $\bar{\lambda}$  is still growing with  $N_{\text{samples}}$ . Both the Wilks Theorem and our analysis are derived in an asymptotic limit  $N \rightarrow \infty$ ; for finite but large  $N$ , both may be invalid. Figure ?? shows that, even at  $N \sim 10^5$ , the behavior of  $\bar{\lambda}$  has failed to become asymptotic. This is surprising, and suggests heterodyne tomography is a particularly exceptional and challenging case to model statistically.

However, our model *does* get some of the qualitative features correct. In Figure ??, we look at  $\langle \lambda_{jk} \rangle$ , where we assume an isotropic Fisher information, and when we simulate heterodyne tomography. While the numbers given for  $\langle \lambda_{jk} \rangle$  do not agree exactly, they still break down into two groups, the “L” and the “kite”. (See Figure ?? for an analysis of the exact differences in the values.)

## VI. CONCLUSIONS AND DISCUSSION

The Wilks Theorem is not generally reliable in quantum state tomography, but our Equation (??) provides a much more broadly applicable replacement that can be used in model selection methods. This includes protocols like the AIC and BIC [?] that do not explicitly use the Wilks Theorem, but rely on the same

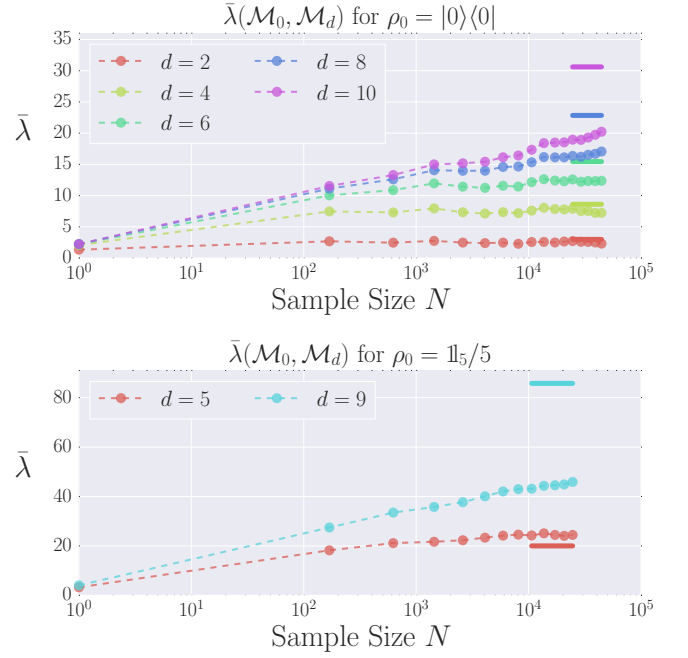


FIG. 9. The empirical average  $\bar{\lambda}$  may have achieved its asymptotic value, or is still growing, depending on the true state  $\rho_0$  and the model dimension  $d$ . Solid lines indicate the value of our formula for the asymptotic expected value, given in Equation (??).

Isotropic Model (10000 Trials)									Heterodyne Tomography (100 Trials)								
0	5.3	0.99	0.97	0.98	1	1	0.99	0.99	0	1.3	1	0.92	0.94	0.85	0.95	0.85	0.67
1	0.99	0.17	0.06	0.06	0.06	0.06	0.06	0.06	1	1	0.11	0.12	0.05	0.02	0.01	0.01	0.01
2	0.97	0.06	0.16	0.06	0.06	0.06	0.06	0.06	2	0.92	0.12	0.06	0.06	0.04	0.03	0.02	0.01
3	0.98	0.06	0.06	0.16	0.06	0.06	0.06	0.06	3	0.94	0.05	0.06	0.05	0.04	0.02	0.02	0.01
4	1	0.06	0.06	0.06	0.16	0.06	0.06	0.06	4	0.85	0.02	0.04	0.04	0.04	0.02	0.02	0.02
5	1	0.06	0.06	0.06	0.06	0.16	0.06	0.06	5	0.95	0.01	0.03	0.02	0.02	0.02	0.01	0.02
6	0.99	0.06	0.06	0.06	0.06	0.06	0.16	0.06	6	0.85	0.01	0.02	0.02	0.02	0.01	0.02	0.02
7	0.99	0.06	0.06	0.06	0.06	0.06	0.06	0.16	7	0.67	0.01	0.01	0.01	0.02	0.02	0.02	0.02
0	1	2	3	4	5	6	7		0	1	2	3	4	5	6	7	
0	2.7	0.99	0.97	0.98	1	1	0.99	0.99	0	0.79	1.1	0.94	0.77	0.51	0.38	0.28	0.35
1	0.99	2.6	1	0.99	1	1	1	0.99	1	1.1	1.8	1.1	0.79	0.89	0.79	0.7	0.57
2	0.97	1	0.33	0.12	0.12	0.12	0.12	0.12	2	0.94	1.1	0.14	0.08	0.04	0.03	0.01	0.01
3	0.98	0.99	0.12	0.34	0.12	0.12	0.12	0.12	3	0.77	0.79	0.08	0.11	0.04	0.03	0.02	0.02
4	1	1	0.12	0.12	0.33	0.12	0.12	0.12	4	0.51	0.89	0.04	0.04	0.08	0.04	0.03	0.02
5	1	1	0.12	0.12	0.12	0.34	0.12	0.12	5	0.38	0.79	0.03	0.03	0.04	0.08	0.04	0.03
6	0.99	1	0.12	0.12	0.12	0.12	0.33	0.12	6	0.28	0.7	0.01	0.02	0.03	0.04	0.05	0.03
7	0.99	0.99	0.12	0.12	0.12	0.12	0.12	0.34	7	0.35	0.57	0.01	0.02	0.02	0.03	0.03	0.05
0	1	2	3	4	5	6	7		0	1	2	3	4	5	6	7	

FIG. 10. The values of  $\langle \lambda_{jk} \rangle$  assuming an isotropic Fisher information (left), and for heterodyne tomography (right). **Top:**  $\rho_0 = |0\rangle\langle 0|$ . **Bottom:**  $\rho_0 = \mathbb{I}_2/2$ . **Discussion:** Qualitatively, the behavior is the same, though there are quantitative differences, particularly within the kite.

assumptions (asymptotic normality, etc). Null theories of loglikelihood ratios have many other applications, including hypothesis testing [?] and confidence regions [?], and our result is directly applicable to them. Refs. [?] both point out explicitly that their methods are unreliable near boundaries and therefore cannot be

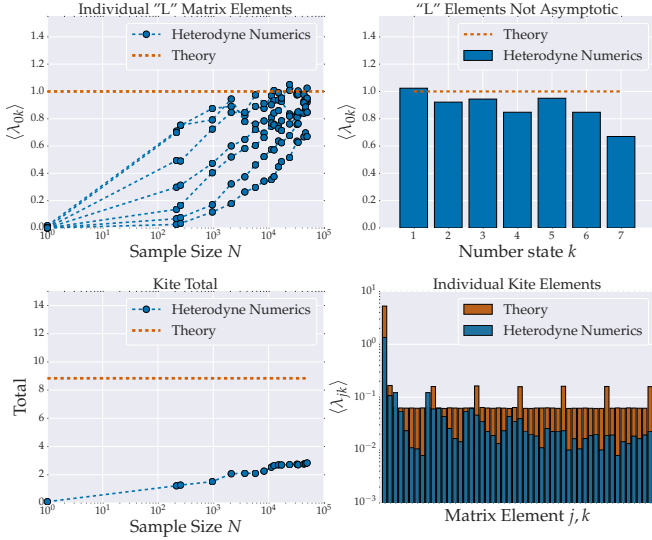


FIG. 11. Examining how our predicted values for  $\langle \lambda_{jk} \rangle$  disagree with simulated heterodyne experiments. We take  $\rho_0 = |0\rangle\langle 0|$  and  $d = 8$ . **Top Left:** The values of  $\langle \lambda_{0k} \rangle$  in the “L” as a function of sample size  $N$ . **Top Right:** Even at the largest  $N$  studied,  $\langle \lambda_{0k} \rangle$  is nontrivially less than 1, especially for the higher number states. **Bottom Left:** The total from the “kite” versus  $N$ . It is clear the total is still growing. **Bottom Right:** The individual “kite” elements  $\langle \lambda_{jk} \rangle$  at the largest  $N$  studied; most are small compared to values they would have in the isotropic case.

applied to rank-deficient states; our result fixes this out-

standing problem. However, our numerical experiments with heterodyne tomography show unexpected behavior, indicating that quantum tomography can still surprise, and may violate *all* asymptotic statistics results. In such cases, bootstrapping [?] may be the only reliable way to construct null theories for  $\lambda$ . Finally, the *methods* presented here have application beyond the analysis of loglikelihoods. They shed light on the behavior of  $\hat{\rho}_{\text{MLE}}$  for rank-deficient states, and can be used to predict other derived properties such as the average rank of the estimate, which is independently interesting for (e.g.) quantum compressed sensing [? ? ? ?].

## VII. ACKNOWLEDGEMENTS:

The authors are grateful for those who provide support for the following software packages: iPython/Jupyter [?], matplotlib [?], mpi4py [?], NumPy [?], pandas [?], Python 2.7 [?], seaborn [?], and SciPy [?]. TLS thanks Jonathan A Gross for helpful discussions on software design, coding, and statistics, John King Gamble for useful insights on parallelized computation and feedback on earlier drafts of this paper, and Daniel Seuss for proofreading edits.

Sandia National Laboratories is a multi-mission laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-AC04-94AL85000.

- 
- [?] M. G. A. Paris and J. Rehacek, eds., *Quantum State Estimation* (Springer, Berlin-Heidelberg, 2004).
  - [?] N. Reid and D. R. Cox, *International Statistical Review* **83**, 293 (2015).
  - [?] L. Wasserman, *All of Statistics: A Concise Course in Statistical Inference* (Springer New York, 2004).
  - [?] J. B. Altepeter, E. R. Jeffrey, and P. G. Kwiat, in *Advances in Atomic, Molecular and Optical Physics*, Vol. 52 (Elsevier, 2005) pp. 105–159.
  - [?] J. I. Bertrand and P. Bertrand, *Foundations of Physics* **17**, 397 (1987).
  - [?] A. I. Lvovsky and M. G. Raymer, *Reviews of Modern Physics* **81**, 299 (2009).
  - [?] G. Breitenbach, S. Schiller, and J. Mlynek, *Nature* **387**, 471 (1997).
  - [?] U. Leonhardt and H. Paul, *Progress in Quantum Electronics* **19**, 89 (1995).
  - [?] F. Motzoi, J. M. Gambetta, P. Rebentrost, and F. K. Wilhelm, *Physical Review Letters* **103**, 110501 (2009).
  - [?] R. Fazio, G. Palma, and J. Siewert, *Physical Review Letters* **83**, 5385 (1999).
  - [?] K. P. Burnham, *Sociological Methods & Research* **33**, 261 (2004).
  - [?] E. J. Candes and T. Tao, *IEEE Transactions on Information Theory* **52**, 5406 (2006).
  - [?] S. T. Flammia, D. Gross, Y.-K. Liu, and J. Eisert, *New Journal of Physics* **14** (2012), 10.1088/1367-2630/14/9/095022.
  - [?] D. Suess, L. Rudnicki, and D. Gross, arXiv:1608.00374 (2016).
  - [?] A. Carpentier, J. Eisert, D. Gross, and R. Nickl, arXiv:1504.03234v2 (2015).
  - [?] L. Schwarz and S. J. van Enk, *Physical Review A* **88**, 032318 (2013).
  - [?] M. Guta, T. Kypraios, and I. Dryden, *New Journal of Physics* **14** (2012), 10.1088/1367-2630/14/10/105002.
  - [?] S. J. van Enk and R. Blume-Kohout, *New Journal of Physics* **15**, 025024 (2013).
  - [?] N. K. Langford, *New Journal of Physics* **15** (2013), 10.1088/1367-2630/15/3/035003.
  - [?] J. O. S. Yin and S. J. van Enk, *Physical Review A* **83**, 062110 (2011).
  - [?] T. Moroder, M. Kleinmann, P. Schindler, T. Monz, O. Guhne, and R. Blatt, *Physical Review Letters* **110**, 180401 (2013).
  - [?] L. Knips, C. Schwemmer, N. Klein, J. Reuter, G. Tóth, and H. Weinfurter, arXiv:1512.06866 (2015).
  - [?] The index  $j$  may be continuous or discrete.
  - [?] Z. Hradil, *Physical Review A* **55**, R1561 (1997).
  - [?] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G.

- White, *Physical Review A* **64**, 052312 (2001).
- R. Blume-Kohout, J. O. S. Yin, and S. J. van Enk, *Physical Review Letters* **105**, 170501 (2010).
  - J. Neyman and E. S. Pearson, *Philosophical Transactions of the Royal Society of London* **231**, 289 (1933).
  - H. Akaike, *IEEE Transactions on Automatic Control* **19**, 716 (1974).
  - S. S. Wilks, *The Annals of Mathematical Statistics* **9**, 60 (1938).
  - L. Le Cam, *Annals of Mathematical Statistics* **41**, 802 (1970).
  - L. Le Cam, in *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, edited by J. Neyman (1956) pp. 129–156.
  - S. Boyd, *Convex Optimization*, 7th ed. (Cambridge University Press, New York, New York, USA, 2009).
  - J. A. Smolin, J. M. Gambetta, and G. Smith, *Physical Review Letters* **108**, 070502 (2012).
  - That is, we let  $\text{Tr}(\delta)$  fluctuate as well.
  - Y. V. Fyodorov, *Arxiv:math-ph/0412017* (2004).
  - E. P. Wigner, *Annals of Mathematics* **67**, 325 (1958).
  - T. Tao and V. Vu, *Random Matrices: Theory and Applications* **2** (2013), 10.1142/S201032631350007X.
  - A. W. van der Vaart, *Asymptotic Statistics* (Cambridge University Press, 2000).
  - This justifies the assumption that  $\rho_{jj} + \delta_{jj} - q > 0$ .
  - A. I. Lvovsky, H. Hansen, T. Aichele, O. Benson, J. Mlynek, and S. Schiller, *Physical Review Letters* **87**, 050402 (2001).
  - K. Husimi, in *Proceedings of the Physico-Mathematical Society of Japan*, Vol. 22 (1940) pp. 264 – 314.
  - The model  $\mathcal{M}_1$  is trivial, as  $\mathcal{M}_1 = \{|0\rangle\langle 0|\}$ . This model will almost always be wrong, in general.
  - G. Schwarz, *The Annals of Statistics* **6**, 461 (1978).
  - R. E. Kass and A. E. Raftery, *Journal of the American Statistical Association* **90**, 773 (1995).
  - S. Glancy, E. Knill, and M. Girard, *New Journal of Physics* **14** (2012), 10.1088/1367-2630/14/9/095017.
  - B. Efron, *The Annals of Statistics* **7**, 1 (1979).
  - J. Higgins, *An Introduction to Modern Nonparametric Statistics* (Brooks/Cole, 2004).
  - A. Steffens, C. Riofrio, W. McCutcheon, I. Roth, B. A. Bell, A. McMillan, M. S. Tame, J. G. Rarity, and J. Eisert, *arXiv:1611.01189* (2016).
  - A. Kalev and C. H. Baldwin, *arXiv:1511.01433v1* (2015).
  - A. Kalev, R. L. Kosut, and I. H. Deutsch, *npj Quantum Information* **1**, 15018 (2015).
  - F. Pérez and B. E. Granger, *Computing in Science and Engineering* **9**, 21 (2007).
  - J. D. Hunter, *Computing in Science and Engineering* **9**, 90 (2007).
  - L. Dalcin, R. Paz, and M. Storti, *Journal of Parallel and Distributed Computing* **65**, 1108 (2005).
  - S. Van Der Walt, S. C. Colbert, and G. Varoquaux, *Computing in Science and Engineering* **13**, 22 (2011).
  - W. McKinney, in *Proceedings of the 9th Python in Science Conference*, edited by S. van der Walt and J. Millman (2010) pp. 51–56.
  - G. van Rossum, “Python Language Reference,” (1995).
  - M. Waskom, “seaborn,” (2016).
  - T. E. Oliphant, *Computing in Science and Engineering* **9**, 10 (2007).