Detecting correlated errors in state-preparation-and-measurement tomography

Christopher Jackson and S. J. van Enk

Detecting correlated errors in SPAM tomography

Christopher Jackson and S.J. van Enk

Oregon Center for Optics and Department of Physics
University of Oregon, Eugene, OR 97403
(Dated: July 30, 2015)

Whereas in standard quantum state tomography one estimates an unknown state by performing various measurements with known devices; and whereas in detector tomography one estimates the POVM elements of a measurement device by subjecting to it various known states, we consider here the case of SPAM (state preparation and measurement) tomography where neither the states nor the measurement device are assumed known. For d-dimensional systems measured by d-outcome detectors, we find there are at most $d^2(d^2-1)$ "gauge" parameters that can never be determined by any such experiment, irrespective of the number of unknown states and unknown devices. For the case $d=2$ we find new gauge-invariant quantities that can be accessed directly experimentally and that can be used to detect and describe SPAM errors. In particular, we identify conditions whose violations detect the presence of correlations between SPAM errors. From the perspective of SPAM tomography, standard quantum state tomography and detector tomography are protocols that fix the gauge parameters through the assumption that some set of fiducial measurements is known or that some set of fiducial states is known, respectively.

I. INTRODUCTION

Quantum tomography has become an important tool for characterizing quantum devices [1, 2]. For example, in order to estimate the state ρ of quantum systems produced by a given quantum source, we let the source create many copies of ρ, and subject those copies to different measurements. If we describe the whole set of measurements by POVM elements Π_i (such that $\sum_i \Pi_i = \mathbb{1}$), the experiment will produce estimates of probabilities

$$p_i = \text{Tr}(\rho\Pi_i), \qquad (1)$$

with $\sum_i p_i = 1$. If we have at least as many observed independent relative frequencies f_i (the number of times outcome i was observed divided by the total number of measurements) as there are parameters in ρ, then we may estimate ρ via a variety of different methods (from linear inversion to hedged maximum likelihood estimation and Bayesian methods, see, e.g., [3–5]). Legitimately ending up with an estimate of a *single* density matrix ρ does require some assumptions about the experimental setup [6–8], namely, invariance under permutations of the copies. In particular, even if there is drift (for example, due to a slowly changing magnetic field or a slowly changing phase of the laser field used to prepare our quantum systems), such that copy ρ_n differs slightly from the previous copy ρ_{n-1}, we still obtain an estimate of a single density matrix, namely, the average:

$$\overline{\rho} = \frac{1}{N_t}\sum_{n=1}^{N_t}\rho_n, \qquad (2)$$

in an experiment with N_t copies of states in total, provided the different measurements are performed in random order.

In more recent times several experiments on detector tomography have been performed [9–12]. Here the assumption is that known states are fed into an unknown measurement device, whose observed outcomes now tell us about the POVM elements that describe the detector. Detector tomography makes sense: A photodetector, for example, is not such a simple device theoretically. There are infinitely many modes of the electromagnetic field, and each mode is described by an infinite-dimensional Hilbert space. A simple candidate model for a photodetector can be formulated in terms of a quantum efficiency η and a dark count rate r, but these two parameters could drift over time and be field-mode dependent. In the case of detector tomography, we estimate probabilities of the form

$$p_\mu = \text{Tr}(\rho_\mu\Pi), \qquad (3)$$

for different known states ρ_μ (where μ runs from 1 to M), in order to estimate the POVM element Π, where the latter is really an average (under the proviso that we test the measurement device with different states created in random order, to ensure symmetry under permutations of the measurements) over the course of the experiment of the different POVM elements Π_n describing the drifting detector. In complete analogy to (2) we may write then

$$\overline{\Pi} = \frac{1}{N_t}\sum_{n=1}^{N_t}\Pi_n, \qquad (4)$$

if we performed this particular measurement N_t times in total.

The present paper focuses on two questions: First, what exactly can we find out about measurements and states when we do not assume we know either one or the other? That is, what can we infer about the POVM elements Π_i describing different measurement settings and about different states ρ_μ from estimates of probabilities

$$p_{\mu,i} = \text{Tr}(\rho_\mu\Pi_i)? \qquad (5)$$

Second, what changes if the fluctuations (or drift) in state preparation and measurement are correlated? That

is, what if the observed average relative frequencies are not determined by averaged states and averaged measurements? How can we infer there are correlations such that

$$\overline{\mathrm{Tr}(\rho_\mu \Pi_i)} \neq \mathrm{Tr}(\overline{\rho}_\mu \overline{\Pi}_i)? \qquad (6)$$

Here the left-hand side is a quantity for which we obtain direct estimates from observed relative frequencies, but the two quantities $\overline{\rho}_\mu$ and $\overline{\Pi}_i$ on the right-hand side are not directly accessible individually and have to be inferred.

Such questions have become of interest in the context of debugging quantum devices that are meant to serve as fault tolerant quantum computers. The requirements on fault tolerance are quite stringent and more and more precise tools for analyzing tomography experiments have been developed very recently [13–15]. Correlations between errors are particularly bad for fault tolerance, hence our focus on detecting correlated SPAM errors.

We are going to analyze SPAM tomography for the case of a single qubit and two-outcome POVMs. Even for this simple case our results suggest novel experiments. Most results generalize easily to multiple qubits and/or higher-dimensional systems and measurements, and in particular to the combination of d dimensional systems and d-outcome POVMs.

II. SPAM TOMOGRAPHY

A. Notation

To set the notation, let us first consider the simplest version of our problem: a single qubit state, ρ, and a single two-outcome qubit detector, with the two outcomes described by POVM elements $\{E, \neg E\}$ (read: E and NOT-E). From the conditions $\rho \geq 0, E \geq 0, \neg E \geq 0$, $\mathrm{Tr}\rho = 1$, and $E + \neg E = \mathbb{1}$, a general parameterization can be written in terms of the Pauli matrices $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)^T$ as

$$\begin{aligned} \rho &= \frac{1}{2}\big(\mathbb{1} + \vec{p} \cdot \vec{\sigma}\big), \\ E &= \frac{1}{2}\big((1+u)\mathbb{1} + \vec{w} \cdot \vec{\sigma}\big), \\ \neg E &= \frac{1}{2}\big((1-u)\mathbb{1} - \vec{w} \cdot \vec{\sigma}\big), \end{aligned} \qquad (7)$$

where positivity is ensured by the inequalities

$$\begin{aligned} |\vec{p}| &\leq 1, \\ |\vec{w}| + |u| &\leq 1. \end{aligned} \qquad (8)$$

Altogether there are seven parameters. Three of these correspond to a choice of coordinate system (when we think of the qubit as a spin-1/2 system, we have to specify what we mean by the components of spin in the x, y, z directions), or, equivalently, to a choice of basis for the

2D Hilbert space describing our qubit. The other four parameters can be thought of as state purity ($|\vec{p}|$), the detector's discrimination power ($|\vec{w}|$), detector bias (u), and the alignment (or fidelity) between detector and state ($\hat{p} \cdot \hat{w}$). We may conveniently represent the two-outcome POVM by a single observable:

$$\Sigma \equiv E - \neg E = u\mathbb{1} + \vec{w} \cdot \vec{\sigma}. \qquad (9)$$

Knowledge of its expectation value S in the state ρ gives us a single constraint on these seven parameters:

$$S \equiv \mathrm{Tr}\rho\Sigma = \vec{p} \cdot \vec{w} + u. \qquad (10)$$

B. Counting parameters

Now consider the more interesting case where we perform two-outcome measurements with a variety of states and observables. We press the μ-th button of a quantum source, where μ runs from 1 to M, representing the preparation of state ρ_μ, while the dial of the measurement device is turned to the i-th setting (where i runs from 1 to N), symbolizing the observable Σ_i. The outcome of measurement i is either that an indicator light blinks (E_i) or does not blink ($\neg E_i$).

Considering the number of undeterminable parameters, γ, a first guess might be that there are

$$\gamma = (\# \text{ unknowns}) - (\# \text{ constraints}) = 3M + 4N - NM \qquad (11)$$

of them, where we assume we always subject all M states to all N measurements. This kind of counting is correct so long as either $N < 3$ or $M < 4$ but otherwise it fails to take the following important observation into account: For $M = 4$, there are sufficiently many states to perform complete detector tomography on any number of detectors. Similarly, $N = 3$ would be a sufficient number of 2-outcome detectors to perform complete state tomography. Hence for any $N > 3$ and $M > 4$, there are just as many undeterminable parameters as when $N = 3$ and $M = 4$, because the addition of any states (or detectors) could reveal no more information about the detectors (or states) than the "original" 4 states (or 3 detectors) already yield. (This argument generalizes to d-outcome measurements performed on d-dimensional systems: each d-outcome measurement yields $d-1$ independent frequencies. $d+1$ known d-outcome POVMs are, therefore, sufficient to determine the d^2-1 parameters of any unknown state. Conversely, d^2 known states (each needing $d^2 - 1$ parameters for its specification) are needed to determine the d^2 parameter of any unknown POVM element. Thus, the maximum number of gauge parameters is $d^2(d^2-1)$.) Therefore the number of undeterminable parameters is

$$\gamma = \begin{cases} 3M + 4N - NM & : N < 3 \text{ or } M < 4 \\ 12 & : N \geq 3 \text{ and } M \geq 4 \end{cases} \qquad (12)$$

(From now on we assume we do have at least 4 states and at least 3 detectors, so that the number of undeterminable parameters is 12.) In order to find what these

undeterminable parameters are we first introduce some convenient notation. First of all, the set of all constraints can be written succinctly as a matrix equation,

$$S = PW \qquad (13)$$

where

$$P \equiv \begin{bmatrix} \cdots & \vec{p}_\mu & \cdots \\ \cdots & 1 & \cdots \end{bmatrix}^T \qquad (14)$$

and

$$W \equiv \begin{bmatrix} \cdots & \vec{w}_i & \cdots \\ \cdots & u_i & \cdots \end{bmatrix} \qquad (15)$$

are the $N \times 4$ and $4 \times M$ matrices of state and detector parameters, respectively, where we included a constant 1 (which stands for $\mathrm{Tr}(\rho_\mu)$) in the set of state parameters just to make the treatment of states and detectors more symmetric. The form of equation (13) makes obvious that the system of constraints has the symmetry:

$$P \longrightarrow PG^{-1}, \quad W \longrightarrow GW, \qquad (16)$$

where G is of the general form

$$G = \begin{bmatrix} H & \vec{0} \\ \vec{a}^T & 1 \end{bmatrix} \qquad (17)$$

where H is any real 3×3 invertible matrix, $H \in \mathrm{GL}(3)$, and $\vec{a}$ is any real 3-vector, $\vec{a} \in \mathbb{R}^3$. The collection of all matrices $G \in \mathbb{R}^3 \rtimes GL(3)$ forms the affine group.

C. Blame Gauges: A New Perspective on State/Detector Tomography

Some of the 12 parameters determining G should be familiar. Most important are perhaps the three parameters needed to describe matrices of the form

$$G = \begin{bmatrix} R & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix}. \qquad (18)$$

where R is in the adjoint representation of SU(2) (which here is just SO(3).) This space of symmetries is always present (also for higher-dimensional Hilbert spaces of size d where we would have the adjoint representation of SU(d)), and corresponds to the arbitrary choice of basis for the underlying Hilbert space. This sort of symmetry tends to be trivial, but becomes nontrivial (and interesting) in the context of quantum communication (the information to fix a reference frame was coined "unspeakable" by Asher Peres [16], for the reason that two distant parties cannot communicate the choice of coordinate system over the phone, unless they already share a reference frame).

Another simple undeterminable parameter is revealed by considering G of the form

$$G = \begin{bmatrix} g\mathbb{1}_3 & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix}, \qquad (19)$$

where $g \in \mathbb{R}$. This kind of scaling transformation trades overall state purity for the detector's discrimination power. That is, imperfections may be blamed, within some limits, more on the preparation of the states than on the quality of our detectors or *vice versa* [18]. Let us refer to parameters like this as *blame gauge* degrees of freedom.

We can characterize all 12 blame gauge parameters as follows. Every H (as appearing in (17)) has a polar decomposition $H = BR$ where R is real orthogonal and B is positive symmetric. This yields the following interpretation: We can always diagonalize B: the three eigenvalues are like the scale parameter g of (19). The basis in which B is diagonal (determined by three additional parameters) determine which components of $\vec{p}$ and $\vec{w}$ can be rescaled in concert.

The three parameters of R still correspond to the choice of Hilbert space basis (or of a coordinate system) as in (18) [19].

Finally, the last three parameters feature in transformations of the form

$$G = \begin{bmatrix} \mathbb{1}_3 & \mathbf{0} \\ \vec{a}^T & 1 \end{bmatrix}, \qquad (20)$$

where $\vec{a}$ is a real 3-vector. It is perhaps clearer what this transformation does when we write out the action on the system parameters explicitly:

$$\begin{cases} \vec{p}_\mu \longrightarrow \vec{p}_\mu - \vec{a} \\ \vec{w}_i \longrightarrow \vec{w}_i \\ u_i \longrightarrow u_i + \vec{a} \cdot \vec{w}_i. \end{cases} \qquad (21)$$

We see that this exchanges state-detector alignment for detector bias, and thus is another type of blame gauge.

Having now completely classified the 12 gauge degrees of freedom in qubit SPAM tomography, let us briefly revisit the more conventional types of (qubit) tomography. Having at our disposal $M = 4$ fiducial states assumed to be fully known allows for complete detector tomography precisely because we can write in that case $W = P^{-1}S$. Similarly, $N = 3$ known fiducial detectors allows us to perform complete quantum state tomography because we can write in that case $P = SW^{-1}$, where W^{-1} is to be interpreted as a right (pseudo-)inverse.

From the perspective of SPAM tomography we see, therefore, that the assumptions underlying state and detector tomography simply boil down to fixing the *blame gauge*. In state tomography on qubits one fixes four parameters each of three fiducial measurements, and in detector tomography one fixes four parameters each of three fiducial states.

III. DETECTING CORRELATED ERRORS

A. Uncorrelated States and Detectors

A measurement yields averages of the form

$$S_{\mu i} = \overline{\langle \Sigma_i \rangle}_\mu = \overline{\vec{p}_\mu \cdot \vec{w}_i + u_i}, \qquad (22)$$

which contains two types of averages: one quantum expectation value (in the state ρ_μ, denoted as usual by $\langle . \rangle$), and one average over the different runs of the experiment (indicated by the over line). The existence of a model (P, W) that satisfies Eq. (13) is precisely the statement that a system consists of *effectively uncorrelated* states and detectors. The point is now that, certainly for $M > 4$ and $N > 3$, not every possible set of data can be effectively uncorrelated simply because the number of data parameters is larger than the number of distinct uncorrelated model parameters whenever

$$MN > 3M + 4N - 12. \tag{23}$$

We will now answer the question under what conditions there is an uncorrelated model and when there is not.

B. Constraints for Uncorrelated Quantum Data: Partial Determinants

Suppose we have quantum data S that is effectively uncorrelated, i.e., described by a model (P, W). For the sake of presentation, let us consider first the case $N = M = 8$. Think of the data as partitioned into four 4×4 corners of an 8×8 matrix,

$$S \equiv \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}, \tag{24}$$

such that

$$Q_{ab} = P_a W_b \tag{25}$$

where

$$P = \begin{bmatrix} P_1 \\ P_2 \end{bmatrix}; \quad W = \begin{bmatrix} W_1 & W_2 \end{bmatrix}, \tag{26}$$

— that is,

$$P_1 = \begin{bmatrix} \vec{p}_1 & \cdots & \vec{p}_4 \\ 1 & \cdots & 1 \end{bmatrix}^T; \quad P_2 = \begin{bmatrix} \vec{p}_5 & \cdots & \vec{p}_8 \\ 1 & \cdots & 1 \end{bmatrix}^T \tag{27}$$

are the top and bottom halves of P and

$$W_1 = \begin{bmatrix} \vec{w}_1 & \cdots & \vec{w}_4 \\ u_1 & \cdots & u_4 \end{bmatrix}; \quad W_2 = \begin{bmatrix} \vec{w}_5 & \cdots & \vec{w}_8 \\ u_5 & \cdots & u_8 \end{bmatrix} \tag{28}$$

are the left and right halves of W.

These submatrices of the data, because of Eq. (25), have the property

$$Q_{11}^{-1} Q_{12} Q_{22}^{-1} Q_{21} = \mathbb{1}_4, \tag{29}$$

under the assumption that the two inverse matrices on the left-hand side exist. (For example, when state 1 and state 2 are the same, or measurements 1 and 3 are the same, Q_{11}^{-1} does not exist. Also, if one of the four vectors $\vec{p}_i$ for $i = 1, 2, 3, 4$ is a mixture of the other three, then

Q_{11}^{-1} does not exist either.) We will refer to the left hand side of (29) as a *partial determinant* of S and write

$$\Delta(S) \equiv Q_{11}^{-1} Q_{12} Q_{22}^{-1} Q_{21}. \tag{30}$$

One should note that the ordering of the rows of P and columns of W was arbitrary and any such reordering would result in a different partial determinant. Note that a partial determinant can be defined for values of N and M as low as 5 [20].

The important conclusion is that the condition that $\Delta S = \mathbb{1}_4$ can be used to detect correlations: a violation of this equality (while taking into account error bars) shows that state preparation and measurement must have been correlated.

C. Nonsingular Partial Determinants

To simplify notation, instead of using Q_{ab}s, let us write the partitions of the data, S, simply as

$$S \equiv \begin{bmatrix} A & B \\ C & D \end{bmatrix} \tag{31}$$

in which case we have

$$\Delta(S) = A^{-1} B D^{-1} C. \tag{32}$$

An awkward feature of this quantity is that it may display singular behavior because of the presence of inverse matrices. A simple way to remedy this is to define a *nonsingular partial determinant* [21]

$$\nabla(S) \equiv \det(D) A (\Delta - \mathbb{1}) = B \bar{D}^T C - (\det D) A \tag{33}$$

where $\bar{D}$ is the cofactor matrix of D so that, for example,

$$A^{-1} = \frac{1}{\det A} \bar{A}^T. \tag{34}$$

We chose to multiply by A simply to reduce computational complexity. This then gives rise to another (nonsingular) condition on having uncorrelated SPAM errors. It has the disadvantage that satisfaction of the condition $\nabla(S) = 0$ may occur not because one's SPAM errors are uncorrelated, but because of having multiplied both sides of (29) with zero.

D. Estimates and Uncertainty for the Partial Determinant of Qubit Data

We consider here briefly error bars in the partial determinant, which are needed to decide whether or not one's data provides significant evidence for correlated SPAM errors.

The data is estimated to maximize the likelihood of MN binomial distributions:

$$\langle S \rangle_{\mu i} \equiv \langle S_{\mu i} \rangle = 2 \frac{k_{\mu i}}{n_{\mu i}} - 1 \tag{35}$$

where $n_{\mu i}$ is the number of trials for the μ-th state and i-th detector settings and $k_{\mu i}$ is the number of positive outcomes. As such, the deviation of values from the actual mean is

$$(\delta S_{\mu i})^2 \equiv \langle S_{\mu i}{}^2 \rangle - \langle S_{\mu i} \rangle^2 = \frac{1 - S_{\mu i}{}^2}{n_{\mu i}}. \qquad (36)$$

We can then propagate these uncertainties to the non-singular partial determinant,

$$\nabla(S) = B \bar{D}^T C - (\det D) A. \qquad (37)$$

Assuming independent Gaussian random variables, the relative coefficients for the qubit case are

$$\frac{\partial \nabla_{ab}}{\partial A_{cd}} = -\delta_{ac} \delta_{bd} \det D \qquad (38)$$

$$\frac{\partial \nabla_{ab}}{\partial B_{cd}} = \delta_{ac} \varepsilon_{ijk} \varepsilon_{dlm} C_{ib} D_{jl} D_{km} \qquad (39)$$

$$\frac{\partial \nabla_{ab}}{\partial C_{cd}} = \delta_{bd} \varepsilon_{cij} \varepsilon_{klm} B_{ak} D_{il} D_{jm} \qquad (40)$$

$$\frac{\partial \nabla_{ab}}{\partial D_{cd}} = \varepsilon_{ijc} \varepsilon_{kld} \left(B_{ak} D_{jl} C_{ib} - \frac{1}{2} D_{ik} D_{jl} A_{ab} \right) \qquad (41)$$

where δ and ε are the Kronecker and Levi-Civita symbols, reps., and repeated indices are to be summed over. If we choose $N = M = 8$, then the subdata can be chosen to be statistically independent and the error bar in the partial determinant is then simply

$$(\delta \nabla_{ab})^2 = \sum_{c,d} \left(\frac{\partial \nabla_{ab}}{\partial A_{cd}} \delta A_{cd} \right)^2 + \left(\frac{\partial \nabla_{ab}}{\partial B_{cd}} \delta B_{cd} \right)^2$$
$$+ \left(\frac{\partial \nabla_{ab}}{\partial C_{cd}} \delta C_{cd} \right)^2 + \left(\frac{\partial \nabla_{ab}}{\partial D_{cd}} \delta D_{cd} \right)^2. \qquad (42)$$

Otherwise, if a smaller number of states and detectors is considered, one must take into account that the entries are not all statistically independent when calculating error bars.

E. Examples of Correlated Models

A significant feature of these partial determinants is that they can be estimated directly from measured data. Stressing the point explicitly: "$\nabla(S) = 0$" may serve as an operational definition for S to have no state-detector correlations. Certainly, if one finds that $\nabla(S) \neq 0$ with high confidence, then one has made a definitive statement that there are state-detector correlations.

It is perhaps helpful if we illustrate the concept of correlated errors with a couple of simple models. For the sake of simplicity, let us restrict our attention to measurement devices whose biases u_i may be chosen to all vanish.

1. Gaussian model

A first way to model correlation is to treat the Bloch vectors for all states and detectors as randomly fluctuating:

$$\begin{aligned} \vec{p}_\mu &= \vec{p}_{\mu o} + \vec{f}_\mu \\ \vec{w}_i &= \vec{w}_{io} + \vec{g}_i \end{aligned} \qquad (43)$$

where the $\vec{p}_{\mu o}$ and $\vec{w}_{io}$ are constants and $\vec{f}_\mu$ and $\vec{g}_i$ are Gaussian noise with

$$\langle \vec{f}_\mu \rangle = \langle \vec{g}_i \rangle = 0. \qquad (44)$$

Our data is then of the form:

$$S = P_o W_o + X \qquad (45)$$

where

$$X_{\mu i} \equiv \langle \vec{f}_\mu \cdot \vec{g}_i \rangle \qquad (46)$$

is the scalar noise correlation.

As a simple special case, consider $X = \chi \mathbb{1}_8$ where $\chi \ll 1$. We then have

$$S \equiv \begin{bmatrix} A & B \\ C & D \end{bmatrix} + \chi \mathbb{1}_8 = \begin{bmatrix} A + \chi \mathbb{1}_4 & B \\ C & D + \chi \mathbb{1}_4 \end{bmatrix} \qquad (47)$$

where $A^{-1} B D^{-1} C = \Delta(P_o W_o) = \mathbb{1}_4$ so

$$\Delta^{-1} = \mathbb{1}_4 + \chi(C^{-1} B^{-1} A + C^{-1} D B^{-1})$$

and

$$\nabla = -\chi(\det D)\left(A(BC)^{-1} A - \mathbb{1}_4 \right). \qquad (48)$$

As expected, the entries of ∇ are linear in the size of the correlations, χ.

2. Causal correlations

A second model considers data of the form

$$S_{\mu i} = \vec{p}_\mu \cdot \vec{w}_{i|\mu} \qquad (49)$$

where the states, $\vec{p}_\mu$, are constants but each detector has a value conditioned on the state they are detecting. Correlations of this type are causal in the sense that the state preparation event is considered to be prior to the detection event and therefore the former cannot depend on the latter, but the latter can depend on the former. More specifically, one could have, for example,

$$\vec{w}_{i|\mu} = \vec{w}_{io} + \chi \, \vec{p}_\mu \qquad (50)$$

where $\vec{w}_{io}$ is fixed and $\chi \ll 1$ is a sort of linear susceptibility. In this case, the data is of the form

$$S = P W_o + Y \qquad (51)$$

where

$$P \equiv \begin{bmatrix} \cdots & \vec{p}_\mu & \cdots \\ \cdots & 1 & \cdots \end{bmatrix}^T,$$

$$W_o \equiv \begin{bmatrix} \cdots & \vec{w}_{io} & \cdots \\ \cdots & 0 & \cdots \end{bmatrix}^T, \qquad (52)$$

$$Y_{\mu i} \equiv \chi |\vec{p}_\mu|^2.$$

We can, obviously, construct additional more complex correlated models containing larger numbers of adjustable parameters. Fitting actual data to the different models and applying model selection would allow us to find the best description of the correlations present in our experiment (as explained further in Ref. [17] in a similar context).

IV.   CONCLUSIONS

We considered a simple version of "gate-set tomography" [13], in which only state preparations and measurements (SPAM) are performed, and no gates in between. As a special case we considered in detail qubits in M *unknown* states measured by N *unknown* two-outcome detectors. These simplifications allowed us to fully analyze errors in SPAM and especially correlations between such errors. We found a simple test for detecting such correlated errors in terms of a "partial determinant" (see Eq. (29)) or in terms of a nonsingular version, see Eq. (37), both of which are determined directly by the observed frequencies of the outcomes of the $M \times N$ different measurements. Such conditions can be written down as soon as $M \geq 5$ and $N \geq 5$.

Finally, by considering the full set of gauge parameters (i.e., parameters not determinable by any (set of) SPAM experiment(s)), we found a novel interpretation of standard quantum state tomography and of the more recently developed detector topography: both can be seen as fixing the gauge parameters in their own particular way: the former by assuming some set of fiducial measurements is fully known, the latter by assuming some set of fiducial states is fully known.

[1] A. I. Lvovsky and M. G. Raymer, Rev. Mod. Phys. **81**, 299 (2009).

[2] M. Paris and J. Rehacek, *Quantum state estimation*, vol. 649 (Springer Science & Business Media, 2004).

[3] R. Blume-Kohout, Phys. Rev. Lett. **105**, 200504 (2010).

[4] R. Blume-Kohout, New. J. Phys. **12**, 043034 (2010).

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

[6] C. M. Caves, C. A. Fuchs, and R. Schack, J. Math. Phys. **43**, 4537 (2002).

[7] R. Renner, Nature Physics **3**, 645 (2007).

[8] S. J. van Enk and R. Blume-Kohout, New J. Phys. **15**, 025024 (2013).

[9] J. Lundeen, A. Feito, H. Coldenstrodt-Ronge, K. Pregnell, C. Silberhorn, T. Ralph, J. Eisert, M. Plenio, and I. Walmsley, Nature Physics **5**, 27 (2009).

[10] A. Feito, J. Lundeen, H. Coldenstrodt-Ronge, J. Eisert, M. Plenio, and I. Walmsley, New Journal of Physics **11**, 093038 (2009).

[11] J. J. Renema, G. Frucci, M. J. de Dood, R. Gill, A. Fiore, and M. P. van Exter, Phys. Rev. A **86**, 062113 (2012).

[12] J. Renema, R. Gaudio, Q. Wang, Z. Zhou, A. Gaggero, F. Mattioli, R. Leoni, D. Sahin, M. De Dood, A. Fiore, et al., Phys. Rev. Lett. **112**, 117604 (2014).

[13] R. Blume-Kohout, J. K. Gamble, E. Nielsen, J. Mizrahi, J. D. Sterk, and P. Maunz, arXiv preprint arXiv:1310.4492 (2013).

[14] S. T. Merkel, J. M. Gambetta, J. A. Smolin, S. Poletto, A. D. Córcoles, B. R. Johnson, C. A. Ryan, and M. Steffen, Phys. Rev. A **87**, 062119 (2013).

[15] C. Stark, Phys. Rev. A **89**, 052109 (2014).

[16] A. Peres and P. F. Scudo, arXiv preprint quant-ph/0201017 (2002).

[17] L. Schwarz and S. J. van Enk, Phys. Rev. A **88**, 032318 (2013).

[18] This parameter is sometimes referred to as the SPAM gauge.

[19] The interpretation of B and R is more subtle for higher Hilbert space dimensions, $d > 2$, because we loose the "accidental isomorphism" $SO(3) \cong SU(2)/Z_2$.

[20] It is worth mentioning that the statement $\Delta(S) = 1$ for all combinations of measurements and states is generically equivalent to the statement that rank$(S) \leq 4$. Geometrically, partial determinants are interesting because they formalize the subspace of effectively uncorrelated data as a variety — simply, solutions to the system of equations $\Delta = 1$ (over various Δs from different permutations of the data.)

[21] This quantity is in fact proportional to what's called the Schur complement, $A - BD^{-1}C$, of A.