Selective and efficient quantum process tomography

Ariel Bendersky,[1] Fernando Pastawski,[2] and Juan Pablo Paz[1]

[1]*Departamento de Física, FCEyN, UBA, Pabellón 1, Ciudad Universitaria, 1428 Buenos Aires, Argentina*
[2]*Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Str. 1, D-85748 Garching, Germany*

In this paper we describe in detail and generalize a method for quantum process tomography that was presented by Bendersky *et al.* [Phys. Rev. Lett. **100**, 190403 (2008)]. The method enables the efficient estimation of any element of the χ matrix of a quantum process. Such elements are estimated as averages over experimental outcomes with a precision that is fixed by the number of repetitions of the experiment. Resources required implementing it scale polynomially with the number of qubits of the system. The estimation of all diagonal elements of the χ matrix can be efficiently done without any ancillary qubits. In turn, the estimation of all the off-diagonal elements requires an extra clean qubit. The key ideas of the method, which is based on efficient estimation by random sampling over a set of states forming a 2-design, are described in detail. Efficient methods for preparing and detecting such states are explicitly shown.

I. INTRODUCTION

For quantum information processing to become feasible, it is necessary to be able to efficiently characterize quantum processing elements. This characterization, taking the name of *quantum process tomography* (QPT), is a very helpful tool, for instance, to design appropriate quantum error correcting codes. In general, QPT is a challenging task due to the exponential amount of parameters involved as a function of the number of qubits (n). Another source of complexity lies in the fact that tomographic methods are typically indirect as the parameters characterizing a quantum process are not directly accessible to experiments but have to be inferred after a large number of such experiments followed by a highly demanding classical postprocessing. To be more precise, it is convenient to represent a quantum process in an abstract way, independently of the physical carriers of information and the actual physical time required for the process to occur. Under rather general conditions the evolution of the system is given by a completely positive linear superoperator mapping initial states into final states: $\mathcal{E}(\rho_{in}) = \rho_{out}$ (more precisely, for this to be valid the initial state must be such that all correlations between the system of interest and its environment are truly classical—i.e., with vanishing discord [1]). This map represents the evolution of quantum states between two snapshots in time. Trace preserving maps having the same output dimension as input dimension are of particular interest and will be the focus of our work. One possible representation of a process is through its χ matrix, which is defined with respect to a certain basis of the space of operators. Choosing the basis $\{E_m\}$ consisting of D^2 operators ($D = 2^n$ is the dimensionality of the Hilbert space of a system of n qubits), the χ-matrix representation for $\mathcal{E}$ is such that

$$\mathcal{E}(\rho) = \sum_{mm'} \chi_{mm'} E_m \rho E_{m'}^{\dagger}. \qquad (1.1)$$

Any completely positive linear map can be written in this way. If the map preserves the trace its χ matrix is such that the condition $\sum_{mm'} \chi_{mm'} E_{m'}^{\dagger} E_m = I$ holds. The map $\mathcal{E}$ is completely characterized by the positive Hermitian matrix $\chi_{mm'}$

(satisfying the above trace preserving condition). Thus, it is clear that a complete characterization of a map $\mathcal{E}$ requires determining $D^4 - D^2$ real parameters, a number that scale exponentially with the number of subsystems (n). Moreover, as we will describe below, existing methods determine any element of the χ matrix after a number of experiments that are exponentially large in n (and also require classical postprocessing that is exponentially dependent on n).

Here, we will describe a method whose main advantage is that it enables us to extract important tomographic information by investing resources that scale polynomially with n. Thus, we will show that we can directly estimate (any) matrix element of the χ matrix by investing resources that scale polynomially with n and depend on the precision required in the estimation. The main idea of this method was presented first in [2]. Here we will describe it in detail and show that to estimate diagonal elements of the χ matrix, the preparation and detection of states from a state 2-design is the main ingredient. For the estimation of off-diagonal coefficients, besides preparing and detecting such states, an ancillary qubit will be needed. The method is "selective" since one can use it to estimate any coefficient (or any set of coefficients). For each coefficient there is an efficient estimation strategy that we describe below. For this reason we denote our strategy as *selective efficient quantum process tomography* (SE-QPT).

It is convenient to briefly describe here the features of other existing tomographic schemes. The first tomographic method proposed [3] is known as standard quantum process tomography. It involves preparing a set of input states ρ_k and then performing full quantum state tomography on the resulting output states obtained after evolution. By doing this, we directly measure coefficients $\lambda_{jk} = \text{tr}[\rho_k \mathcal{E}(\rho_j)]$. However, if one wants to find the matrix elements $\chi_{mm'}$ it is necessary to invert an exponentially large system of equations relating λ with χ [4]. For this reason, the method is indirect (since it requires inversion to obtain matrix elements χ_{mn}). It is also inefficient since, in the most general case, in order to estimate any of the coefficients χ_{mn} one needs to perform an exponentially large number of experiments and classical

postprocessing. Another method known as direct characterization of quantum dynamics was recently proposed [5,6] and it requires an extra ancillary system of n clean qubits. The ancillary qubits must go through a clean quantum channel. Provided such a resource is available, the method enables the direct estimation of all diagonal χ_{mm} by associating them with survival probabilities of entangled (Bell) states of the system and the ancilla. The estimation of off-diagonal elements $\chi_{mm'}$ is also possible in this context but it turns out that it requires the inversion of a system of equations which, in the most general case, involves an exponential number of terms. More recently, the method of symmetrized characterization of noisy quantum processes [7] (SCNQPs) was introduced. It is based on the idea of transforming the original channel $\mathcal{E}$ into a symmetrized channel $\mathcal{E}'$ via twirling operations. After symmetrization, only diagonal χ'_{mm} coefficients remain, being the averages over the original coefficients of the same Hamming weight. The twirling is achieved using only $O(n)$ single qubit gates with constant depth. The values of the averaged coefficients are linearly related to output probabilities through an upper diagonal square matrix of size $n+1$. The method is ideally tailored for evaluating the applicability of relevant quantum error correcting codes [8] as it allows the evaluation of diagonal χ_{mm} coefficients averaged over operators of the same Hamming weight (i.e., χ_{00}, average over one qubit errors, etc.). However, it is not possible to estimate any of the off-diagonal $\chi_{mm'}$ coefficients, which are wiped out by the symmetrization protocol, nor distinguish among specific Pauli errors of the same Hamming weight.

Thus, existing methods do not allow the efficient estimation of an arbitrary coefficient $\chi_{mm'}$. This will be, in fact, one of the main characteristics of the method we will discuss below. The other great advantage is that only a single ancillary qubit will be required for off-diagonal coefficient estimation and none for diagonal no matter the number of qubits of the system. Our method has a similar flavor to SCNQP adding the possibility to determine any of the coefficients $\chi_{mm'}$ (including off-diagonal ones) with polynomial resources. The method is based on two observations. The first is the fact that any matrix element $\chi_{mm'}$ can be related to an average survival probability of input states under the action of the channel (or a related quantity as described below). The average involved here is over the entire Hilbert space using the so-called Haar measure. The second observation is that such averages can be efficiently estimated by sampling over a finite set of states (2-design, as described below).

This paper is organized as follows. In Sec. II, we review the method we will use to compute averages over the entire Hilbert space: we define and discuss the concept of 2-design. In Sec. III we present the core of the SEQPT method: we show how any element of the χ matrix can be efficiently estimated. We separately describe the estimation of diagonal and nondiagonal matrix elements presenting a detailed budget for the resources required for the estimation. Section IV shows how this very same algorithm can be extended. Thus, we present a method where all the information required for the simultaneous estimation of all diagonal coefficients of the χ matrix is obtained from the same experiment. Also, we give a complete error analysis for the method. In Appendix A we give a brief review on mutually unbiased bases (MUBs)

and show that they are proper 2-designs, and in Appendix B we show how to actually prepare any state on a complete set of MUBs (i.e., on a 2-design) by giving an explicit construction of change in basis circuits among the different bases of the MUBs.

II. COMPUTING AVERAGES IN HILBERT SPACE USING 2-DESIGNS

A crucial part of the method we will describe below consists in estimating averages over the entire Hilbert space of products of expectation values of two operators. The computation of quantities of this type was analyzed before and, as shown in [9], for any pair of operators O_1 and O_2 we have

$$\int \langle\psi|O_1|\psi\rangle\langle\psi|O_2|\psi\rangle d\psi = \frac{\mathbf{tr}(O_1)\mathbf{tr}(O_2) + \mathbf{tr}(O_1 O_2)}{D(D+1)}.$$

(2.1)

The integral above is over the entire Hilbert space using the so-called Haar measure (which is the only unitarily invariant one).

Experimentally measuring the above quantities, which involve averages over the entire Hilbert space, seems completely unrealistic. However, the beautiful recent work on the theory of 2-designs [10–13] provides the means for doing so. Delsarte [14] showed how integrating polynomials on the sphere could be reduced to averaging the integrand on a finite set of points coined spherical designs (the important fact is that one can use the same set of points to evaluate the average of any polynomial of a fixed degree). The same idea can be extended to integrals over the entire Hilbert space. A state 2-design X is a set of states satisfying

$$\int \langle\psi|O_1|\psi\rangle\langle\psi|O_2|\psi\rangle d\psi = \frac{1}{|X|}\sum_{\psi\in X}\langle\psi|O_1|\psi\rangle\langle\psi|O_2|\psi\rangle$$

(2.2)

for all operators $O_{1,2}$. Thus, averaging over the entire Hilbert space is equivalent to averaging over the finite set X (of cardinality $|X|$). A state 2-design with a finite (but exponentially large) number of states exists. Although it may seem that averaging over an entire 2-design is an exponentially hard task, it is possible to estimate this average by randomly sampling over initial states $|\psi\rangle$ chosen from the set X. The size of the sample of initial states required will scale polynomially on the number of qubits, allowing to efficiently estimate the required average. This idea will be crucial for the method we will present below.

Luckily, it is rather easy to produce a state 2-design for n qubits. One possibility is to find $D+1$ MUBs that automatically form a state 2-design [13]. Each basis will be labeled with an index $J=0,\dots,D$ and the states within each basis will be labeled with the index $m=1,\dots,D$. In order for the orthonormal bases to be unbiased, the $D(D+1)$ states of the MUBs must satisfy $|\langle\psi^J_m|\psi^K_n\rangle|^2 = \frac{1}{D}$ for all $J\neq K$. Since generalized Pauli operators may be partitioned into $D+1$ maximally sets of D commuting operators so that each pair of sets only holds the identity I as common element [15], there are

FIG. 1. Circuit for measuring χ_{mm} for a given channel $\mathcal{E}$.

$D+1$ MUBs, each one diagonalizing each of these commuting subsets of Pauli operators [16]. In this way, the set of states in the MUBs can be efficiently described and also can be efficiently generated with $O(n^2)$ one and two-qubit gates [17]. It is simple to adapt the procedure used to efficiently generate any state in any MUB to compute survival probabilities of such states and also to compute the transition rates from the (J,m) to (J,m') states.

It is interesting to mention that there are other sets of states that form 2-designs but are not MUBs. For example, Dankert *et al.* [10,11] proposed to use of approximate unitary 2-design (which are designs on the space of unitary operators) showing that they can be efficiently approximated. Approximate unitary 2-designs with accuracy $\epsilon + 1/D^2$ can be obtained by employing only $O(n \log \frac{1}{\epsilon})$ gates. A unitary 2-design acting on any fixed state induce state 2-design fitting into the previous scheme. Dually the action of the random unitaries may be interpreted as symmetrizing the channel $\mathcal{E}$ through twirling. Following this line, we may also use weaker symmetrization protocols as in SCNQP [7] for estimating fidelities of modified channels, as is shown in Figs. 1 and 2.

The importance of 2-designs for the task of quantum process parameter estimation was first pointed out by Dankert *et al.* [11]. This first work proved that 2-designs provide the means for efficiently measuring the fidelity of a quantum process $\bar{F}(\mathcal{E})$ defined as

$$\bar{F}(\mathcal{E}) = \int_{U(D)} dU \, \mathbf{tr}[U|0\rangle\langle 0|U^\dagger \mathcal{E}(U|0\rangle\langle 0|U^\dagger)], \quad (2.3)$$

where the integral is over the Haar measure. Since the integrand is a polynomial of degree 2 in U and degree 2 in $U^\dagger$ unitary 2-design allow evaluating the expression exactly as an average over a finite set of operators U. If we think of the operator U in the integrand as acting over $|0\rangle$, then the integral over U may be cast in terms of an integral over $|\psi\rangle$,

$$\bar{F}(\mathcal{E}) = \int_{\psi(D)} d\psi \langle \psi | \mathcal{E}(|\psi\rangle\langle\psi|) | \psi\rangle. \quad (2.4)$$

This equation makes clear that the state $|0\rangle$ does not play a special role in defining the fidelity of the process $\mathcal{E}$. At the same time, since the integrand is a polynomial of degree 2 in $|\psi\rangle$ and degree 2 in $\langle\psi|$, it opens the possibility of using a quantum state 2-design for evaluating the average fidelity for $\mathcal{E}$.
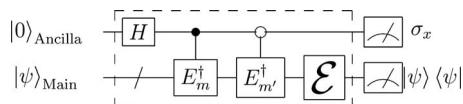
FIG. 2. Circuit for measuring $\mathbf{Re}(\chi_{mn})$ for a given channel $\mathcal{E}$.

III. SELECTIVELY MEASURING CHANNEL COEFFICIENTS

In this section we will present the main idea that enables SEQPT. Below, we will separately discuss the evaluation of diagonal and nondiagonal elements of the χ matrix. However, the evaluation of both types of coefficients is based on the use of a mathematical identity that relates such coefficients with an average fidelity of a modified channel. Thus, using Eq. (2.1) above, together with the χ matrix representation of the channel $\mathcal{E}$, it is simple to show the validity of the following equation:

$$\bar{F}_{ab}(\mathcal{E}) = \int_{\psi(D)} \langle\psi|\mathcal{E}(E_a^\dagger|\psi\rangle\langle\psi|E_b)|\psi\rangle d\psi = \frac{D\chi_{ab} + \delta_{ab}}{D+1}. \quad (3.1)$$

This equation is valid provided we use an operator basis $\{E_m\}$ which is orthogonal $[\mathbf{tr}(E_m E_n^\dagger) = D\delta_{mn}]$ and such that $\mathbf{tr}(E_m) = D\delta_{m0}$ (we also assumed that the channel $\mathcal{E}$ is trace preserving). An example of such kind of basis is the one formed by the generalized Pauli operators (obtained by n-fold tensor products of the identity and/or the three Pauli operators on each qubit). The proof of the above equation begins by replacing $\mathcal{E}$ by its definition given in Eq. (1.1). This and the orthogonality of the operator basis yield

$$\bar{F}_{ab}(\mathcal{E}) = \frac{D\chi_{ab}}{D+1} + \frac{1}{D(D+1)} \mathbf{tr}[\mathcal{E}(E_a^\dagger E_b)] \quad (3.2)$$

and using that the channel is trace preserving and the orthogonality of the operator basis yields the desired result from Eq. (3.1).

A. Evaluating diagonal coefficients

In particular, the above equation shows that diagonal coefficients χ_{mm} are directly related to averaged fidelities of the modified channel $\mathcal{E}_m$ defined as $\mathcal{E}_m(\rho) = E_m^\dagger \mathcal{E}(\rho) E_m$. Thus, it is straightforward to show that

$$\bar{F}(\mathcal{E}_m) = \frac{D\chi_{mm} + 1}{D+1}. \quad (3.3)$$

The channel $\mathcal{E}_m$ is simply obtained as the application of the original channel $\mathcal{E}$ followed by the operator $E_m^\dagger$ (which we assume to be unitary in what follows). Hence, if a method for measuring fidelity is available, diagonal elements of the χ matrix are also accessible. One such method is shown in Fig. 1.

According to our previous discussion, the fidelity averaged over the entire Hilbert space can be evaluated also by averaging over a state 2-design. Thus, the diagonal coefficients are evaluated as the average fidelity over all the states of the 2-design. Below, we will show that estimating the average fidelity with fixed precision requires a number of measurements that scale polynomially with the number of qubits. To do this, one should simply be able to randomly sample over the states of the 2-design. In such case the standard deviation in the estimation decreases as $1/\sqrt{M}$, where M is the number of experimental runs. Also, in Appendix A we explicitly show an efficient method to prepare the set of

states of the 2-design. This completes the method to evaluate diagonal coefficients.

B. Evaluating off-diagonal coefficients

The evaluation of off-diagonal elements requires a slightly different strategy. This is the case because, unlike diagonal coefficients, $\chi_{mm'}$ is not related with the average fidelity of a completely positive channel. In fact, for $m \neq m'$ the above expression reduces to

$$\int \langle\psi|\mathcal{E}(E_m^\dagger|\psi\rangle\langle\psi|E_{m'})|\psi\rangle d\psi = \frac{D\chi_{mm'}}{D+1}. \quad (3.4)$$

As $\rho \to \mathcal{E}(E_m^\dagger \rho E_{m'})$ is in general not a physical map (it does not even preserve hermiticity and trace of the state) the measurement of off-diagonal coefficients is not as straightforward. However, as we will now show, it can be achieved by using a single qubit as an ancilla.

The measurement scheme is very similar to the one used in the determinstic quantum computing with only one pure qubit (DQC1) model of quantum computation [18] and is described by the circuit exhibited in Fig. 2. As before, the state $|\psi\rangle$ used as input should be randomly chosen from the 2-design. Let us now analyze circuit 2 to show that it indeed measures $\chi_{mm'}$.

The input state is $\rho_0 = |0\rangle\langle 0| \otimes |\psi\rangle\langle\psi|$. After going through the circuit, but prior to measurement, the state is given by

$$\rho_f = \frac{1}{2}[|0\rangle\langle 0| \otimes \mathcal{E}(E_{m'}^\dagger|\psi\rangle\langle\psi|E_{m'}) + |0\rangle\langle 1| \otimes \mathcal{E}(E_{m'}^\dagger|\psi\rangle\langle\psi|E_m)$$
$$+ |1\rangle\langle 0| \otimes \mathcal{E}(E_m^\dagger|\psi\rangle\langle\psi|E_{m'}) + |1\rangle\langle 1| \otimes \mathcal{E}(E_m^\dagger|\psi\rangle\langle\psi|E_m)].$$
$$(3.5)$$

It can be easily shown that the expectation value of σ_x or σ_y on the ancilla qubit conditioned to the survival of the state $|\psi\rangle$ on the main system is related to the off-diagonal $\chi_{mm'}$ coefficients as

$$\int \mathbf{tr}[\rho_f(\sigma_x \otimes |\psi\rangle\langle\psi|)]d\psi = \frac{D \operatorname{Re}(\chi_{mm'}) + \delta_{mm'}}{D+1}, \quad (3.6)$$

$$\int \mathbf{tr}[\rho_f(\sigma_y \otimes |\psi\rangle\langle\psi|)]d\psi = \frac{D \operatorname{Im}(\chi_{mm'})}{D+1}. \quad (3.7)$$

This shows how to measure off-diagonal χ coefficients.

It is worth noticing that the method independently estimates all coefficients—both diagonal and off-diagonal ones. Because of the unavoidable estimation errors, by completely estimating the χ matrix in this way we would certainly obtain a noncompletely positive and trace preserving map, as required. The problem of finding a completely positive map that is, in some sense, the closest one to the map reconstructed with the experimental data is important but out of the scope of our present work. It is important to notice that this problem arises when doing full tomography of the channel. However, one of the strengths of our protocol is to enable an efficient partial characterization of the channel by measuring, for example, certain elements of the χ matrix
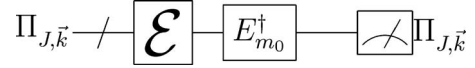

FIG. 3. Two-design circuit for measuring χ_{mm} for a given channel $\mathcal{E}$.

(estimating average fidelities, etc.). For this case, the above mentioned problem is certainly does not exist.

IV. ERROR ANALYSIS AND GENERALIZATIONS

A. Error analysis for selective and efficient quantum process tomography

The method discussed above requires the use of states forming a 2-design as inputs and the detection of such states to estimate the survival probabilities. The use of a 2-design is crucial to evaluate the required average. In particular, it is useful to use a special type of 2-design formed by the $D(D+1)$ states belonging to $D+1$ mutually unbiased bases of the Hilbert space. We will denote projector onto the $\vec{k}$th state of the J'th base as $\Pi_{J,\vec{k}}$ (see Appendix A). Using such notation, the circuit in Fig. 3 describes the protocol for the estimation of the $\chi_{m,m}$, which is obtained by randomly sampling over J and $\vec{k}$. An effective way of estimating this average, known as Monte Carlo sampling, is to randomly choose both J and $\vec{k}$ for each experiment and take an average over a set of M such experiments. The values obtained in this way are unbiased estimators of $\chi_{m_0 m_0}$. Furthermore, since the results of each individual experiment are either 0 or 1, we can bound the variance of the average over M experiments by $\frac{1}{4M}$. If we want to ensure with a probability p that the error is lower than ϵ, then a Chernoff bound implies that a number of experiments satisfying $M \geq \ln[2(1-p)^{-1}]/(2\epsilon^2)$ are needed. Each of these M experiments will have a complexity of $O(n^2)$ arising from the number of elementary quantum gates required for the change of basis circuit (see Appendix B). The estimation of off-diagonal elements of the χ matrix requires four times more experiments because of the measurement on the ancillary qubit.

B. Generalization I: Simultaneous estimation of diagonal coefficient using transition probabilities

It is useful to notice that the method can be easily extended if one is able to prepare the states corresponding to the MUBs associated with the operator basis E_m used to describe the channel. Let us focus on the case where the operators E_m are generalized Pauli operators. Suppose that the projectors $\Pi_{J,\vec{k}}$ are the states stabilized by commuting subgroups of Pauli operators (see Appendix A). Then, we can simply find out how the operator E_{m_0} acts on the state $\Pi_{J,\vec{k}}$ upon conjugation. Hence, the expectation value of the observable

$$\mathbf{tr}[E_{m_0}^\dagger \mathcal{E}(\Pi_{J,\vec{k}})E_{m_0}\Pi_{J,\vec{k}}]$$

corresponding to instances of the experiment is equal to

$$\mathbf{tr}[\mathcal{E}(\Pi_{J,\vec{k}})\Pi_{J,\vec{k}''}] \qquad (4.1)$$

for some $\vec{k}''$ which depends on $\vec{k}$, J, and E_{m_0}. Then, it is clear that by detecting not only the survival of the state but also recording all the possible transitions we can obtain all the required information to estimate any diagonal coefficient and not a single one. Thus, the strategy is simple: one maintains the random preparation step of a state labeled by the indices $(J,\vec{k})$ and one stores the information about the recorded state labeled with $(J,\vec{k}')$. Given the operator E_m, the event should be counted in the estimation of χ_{mm} only if the input state $(J,\vec{k})$ is mapped onto the output state $(J,\vec{k}')$ by the action of E_m. Suppose that we perform a set of repetitions of this experiment and denote each event by a triplet $(J,\vec{k},\vec{k}')$, where J indicates the basis randomly chosen for the experiment, $\vec{k}$ indicates the randomly chosen input state, and $\vec{k}'$ indicates the measured output state. Then, the experiment should be counted positively in the fidelity of E_m if and only if $\vec{k}'-\vec{k}$ is precisely the commutation vector of the operator E_m with respect to basis $\mathcal{B}_J$. Every experiment counts positively for exactly D values of m. The commutation vector of an operator E with respect to basis $\mathcal{B}_J$ is the binary vector $\vec{v}$ such that

$$J_i E = (-1)^{v_i} E J_i, \qquad (4.2)$$

where $J_1, J_2, \ldots, J_n$ are the canonical generators for the stabilizer group of $\mathcal{B}_J$. We may restate this as saying that v_i is the simplectic inner product between the vector describing E and the one describing J_i.

Since the commutation vector is so important for the estimation of the χ_{mm} coefficients, we will show how it can be efficiently calculated from the canonical descriptions of E_m and $\mathcal{B}_J$. For E_m, we assume a canonical description through a binary vector with two parts, $\vec{m_X}$ and $\vec{m_Z}$, such that $E_m = X^{\vec{m_X}} Z^{\vec{m_Z}}$. For $\mathcal{B}_J$, there are two possibilities. Either $\mathcal{B}_J$ is the computational basis $\mathcal{B}_Z$ or $\mathcal{B}_J$ is a basis stabilized by the group described by the binary vector $\vec{J}$ as in Eq. (A3) from Appendix A. In either case, a canonical representation of the n generators of the stabilizer may be obtained with only $O(n^2)$ operations. The calculation of the commutation vector then additionally requires the calculation of n simplectic inner products thus maintaining the algorithmic complexity.

Thus, the estimation of a specific χ_{mm} from such a set of M experiments requires $O(Mn)$ storage for the description of the M individual experiments. The amount of classical processing required is $O(Mn^2)$ mainly for the verification of M commutation vectors. Since each of the M experiment already requires $O(n^2)$ elementary quantum gates and $O(n^3)$ classical processing to determine these quantum gates, this will be the dominant complexity term for the estimation of χ_{mm}.

C. Generalization II: Detecting and measuring large χ_{mm} coefficients

We have shown that a single set of experiments is capable of providing information to estimate any of the diagonal χ_{mm} coefficients. We will now extend this result to determine the operators E_m related to the largest χ_{mm} coefficients. A

straightforward search by estimating all χ_{mm} coefficient is not a reasonable approach to doing this as the number of such coefficients is exponential in n. Surprisingly, we will see that finding and estimating the set of such χ_{mm} are actually possible under the condition that there be only *a few* χ_{mm} with high values. This is precisely the case that can be effectively remedied by quantum error correction, where the set of correctable error syndromes is typically small.

The idea behind this estimation is that the larger coefficients will be compatible with many of the triplets $(J,\vec{k},\vec{k}')$. Finding the larger coefficients will be done by taking experimental results (i.e., triplets) by pairs and finding the E_m compatible with both experiments. Those E_m found will be the ones associated with the larger χ_{mm}.

Suppose that we estimate all $\bar{F}(\mathcal{E}_m) = (D\chi_{mm}+1)/(D+1)$ coefficients from a set of M experiments. The smallest nonzero value we could obtain in such an estimation is $\frac{1}{M}$, while the second smallest value is $\frac{2}{M}$, and so on. The efficiency of our criteria for finding the E_m with large χ_{mm} coefficients is based on the fact that we can efficiently detect all $F(\mathcal{E}_m)$ for which the estimated value is greater or equal to $\frac{2}{M}$. Furthermore, we will be able to quantify how unlikely it is for a $F(\mathcal{E}_m)$ larger than ϵ to have an estimated value smaller than $\frac{2}{M}$.

Suppose we wish to determine the operator E_m consistent with two experiments $(J,\vec{k}_1,\vec{k}_2)$, $(J',\vec{k}_1',\vec{k}_2')$. If $J=J'$ there will be either 0 or D such operators E_m. Otherwise, there is exactly one such operator which we will show how to determine. We will start by noting that any operator E_m can, up to a phase, be written as

$$E_m \cong \prod_{i=0}^{n-1} J_i^{q_i} \prod_{i=0}^{n-1} J_i'^{q_i'}, \qquad (4.3)$$

where the J_i and J_i', with $i \in \{1 \ldots n\}$, are the canonical generators of the stabilizer J and J', respectively. Once the vectors $\vec{q}$ and $\vec{q}'$ are obtained, it is straightforward to obtain the canonical representation for E_m with $O(n^2)$ classical operations. Note that representation [Eq. (4.3)] of E_m is just a generalization of the canonical representation, where $\mathcal{B}_J = \mathcal{B}_X = \mathcal{B}_0$ and $\mathcal{B}_{J'} = \mathcal{B}_Z$.

We must obtain $\mathbf{q}$ and $\mathbf{q}'$ such the E_m given by representation [Eq. (4.3)] satisfies

$$J_i E_m = (-1)^{k_2 - k_1} E_m J_i,$$

$$J_i' E_m = (-1)^{k_2' - k_1'} E_m J_i'. \qquad (4.4)$$

To do this, we determine the nonsingular binary matrix C such that

$$J_i J_j' = (-1)^{C_{i,j}} J_j' J_i. \qquad (4.5)$$

Condition (4.4) may now be translated as

$$\mathbf{k}_2 - \mathbf{k}_1 = C\mathbf{q}',$$

$$\mathbf{k}_2' - \mathbf{k}_1' = C^T \mathbf{q}. \qquad (4.6)$$

Thus, by inverting C and C^T, we may obtain the necessary values for $\mathbf{q}$ and $\mathbf{q}'$. This procedure may be repeated on each

of the $\binom{M}{2}$ pairs of experiment data to find the E_m for which the estimate for $\bar{F}(\mathcal{E}_m) \geq 2/M$.

D. Error analysis for simultaneous determination of coefficients

One may consider now how many experiments are needed to obtain the χ_{mm} coefficients with a certain precision. If one wishes to measure all χ_{mm} coefficients greater than ϵ, all within individual uncertainty δ, one may give the number of experiments M sufficient for achieving this with a probability P as [19]

$$M \geq \frac{2\left(D + \frac{1}{\epsilon}\right)(D+1)}{D^2 \delta^2 (1-P)} = \frac{2\left(1 + \frac{1}{D\epsilon}\right)\left(1 + \frac{1}{D}\right)}{\delta^2 (1-P)}. \quad (4.7)$$

This is obtained by considering that the estimation of every $\bar{F}(\mathcal{E}_m)$ is obtained with a variance of $2\bar{F}(\mathcal{E}_m)[1 - \bar{F}(\mathcal{E}_m)]/M$ and by using a bound on the error function.

If one further considers $\epsilon \gg \frac{1}{D}$ this expression may be simplified to

$$M \gtrsim \frac{2}{\delta^2 (1-P)}. \quad (4.8)$$

This means that we may perform full diagonal tomography with only polynomial resources in both the number of qubits in the system and the desired precision δ. A loophole in this argument is that most of the coefficients χ_{mm} may take values close to $\frac{1}{D}$ when the basis $\{E_m\}$ is chosen without any knowledge about the channel since a bias toward only a few of the χ_{mm} is not expected. The proposed method will yield good results when the channel under consideration is not random, particularly for characterizing highly local noise.

V. SUMMARY

In this paper we have shown how to selectively and efficiently measure any coefficient of the χ-matrix representation of a channel. To estimate such coefficients we could adapt any method capable of efficiently estimate the average fidelity of a channel. In particular, we described in detail how to do this by estimating the fidelity of states randomly sampled over the states of a set of mutually unbiased basis, which have the property of forming a state 2-design. We presented an explicit construction of the circuits implementing the change in bases for the set of MUB formed by the bases stabilized by tensor product Pauli operators. The fact that these MUBs form a 2-design allows us to sample over a finite set of states to obtain the mean values required for the coefficient estimation. We are also able to profit from the interesting properties of stabilizer states appearing in this construction. These properties enable the estimation of any diagonal χ_{mm} from the same set of experimental measurements. This construction, allowing the efficient and selective estimation of coefficients, is not the only strength of this method. It also enables us to efficiently determine every diagonal coefficient larger than a certain value. This is the

application profiting from both the 2-design averaging property and stabilizer properties of this set of states.

APPENDIX A: MUTUALLY UNBIASED BASES

MUBs are a construct from combinatorial mathematics that has become part of the theory toolbox for quantum information. We say that two orthonormal bases $\mathcal{B}_J = \{|\psi_m^J\rangle : m \in 1 \ldots D\}$ and $\mathcal{B}_K = \{|\psi_l^K\rangle : l \in 1 \ldots D\}$ are mutually unbiased if and only if

$$\forall \ m, l, \quad |\langle \psi_m^J | \psi_l^K \rangle| = \frac{1}{D}. \quad (A1)$$

The usual reading of this equation states that measurement in basis $\mathcal{B}_J$ gives absolutely no information about measurement in basis $\mathcal{B}_K$ and vice versa. It has been shown that there can be at most $D+1$ bases which are mutually unbiased and constructions are only known for prime power dimensions. Klappenecker and Roetteler [13] related maximal sets of MUB to 2-designs. They proved that the set of states belonging to $D+1$ mutually unbiased bases is itself a 2-design. Earlier, Bandyopadhyay *et al.* [15] had proven a strong connection between sets of mutually unbiased bases and maximally commuting sets of orthogonal unitary operators. One of their results is that if one partitions a complete set of $D^2 - 1$ mutually orthogonal traceless operators into $D+1$ subsets of $D-1$ commuting operators each, then the $D+1$ bases diagonalizing each of these subsets are mutually unbiased. We will provide an explicit construction of such sets as it will later be necessary to refer to it and invoke some of its properties.

When dealing with tensor product Pauli operators, the construction going from operators to states and back may be cast in terms of the stabilizer formalism [20–22]. In this formalism we will say that a state $|\psi\rangle$ is stabilized by an operator O if it holds that $O|\psi\rangle = \pm|\psi\rangle$. States are described by the set of operators stabilizing them and the corresponding eigenvalues. If the $2^{2n} - 1$ generalized Pauli operators are partitioned into $2^n + 1$ maximally commuting subsets of $2^n - 1$ Pauli operators, each of these subsets will be the stabilizer of a basis and each of these bases will be unbiased to each other. Thus, the problem of finding the stabilizer groups for the MUB set is reduced to that of partitioning the generalized Pauli operators into $2^n + 1$ Abelian groups.

The easiest way to construct this partition is via the finite field construction first introduced by Wootters [23] and used by Paz *et al.* [24]. The first requirement is the construction of the companion matrix M for a primitive polynomial of the finite field $GF(2^n)$,

$$M = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & & & & & 1 \\ r_0 & r_1 & r_2 & \cdots & & r_{n-1} \end{pmatrix}, \quad (A2)$$

where the primitive polynomial for the finite field is $p(x) = r_0 + r_1 x + r_2 x^2 + \cdots + r_{n-1} x^{n-1} + x^n$.

This matrix has the property that $M^D = M$ and $M^k \neq M$ $\forall$ $k < D$, where every operation on the matrix is performed, mod 2. Consider the following sets of generalized Pauli operators:

$$G_{\vec{b}} = \{1, P_{\vec{b},j} = X^{\vec{1}M^j} Z^{\vec{b}\hat{M}^j} : j = 1, \ldots, D-1\}, \qquad \text{(A3)}$$

where $\vec{b} \in \{0,1\}^n$ is an n bit vector, $\hat{M}$ is the transpose of M, $\vec{1} = (1, 0, 0, \ldots)$ is the first canonical binary basis vector, and $X^{\vec{b}} = \otimes_i X^{b_i}$. Note that since $M^D = M$, we have that $j = 0$ is equivalent to $j = D-1$.

It is easy to check that $G_{\vec{b}}$ is an Abelian group and that the only common operator between groups $G_{\vec{b}}$ and $G_{\vec{b}'}$ is the identity. Thus, the set $G_{\vec{b}}$, along with the group formed by the tensor product of Z operators—that is, the stabilizer group for the computational basis—is the partition needed.

This completes the explicit construction of the MUB set and of the required 2-design. On Appendix B we will see an operational approach on how to prepare the states from this MUB set and how to measure on each of these bases.

APPENDIX B: CHANGE IN BASIS CIRCUIT

The circuits presented in Sec. III assume that states are sampled over the whole Hilbert space. However, it is sufficient to sample over a complete set of MUBs since these form a state 2-design. In this appendix we will show how to build efficient change in basis circuits for a given set of MUBs that, along with translations in the computational basis, will allow us to sample over every state from the 2-design. The change in basis circuits should implement the unitary rotation V_J^K such that

$$\forall \quad a \quad V_J^K |\psi_a^J\rangle = |\psi_a^K\rangle. \qquad \text{(B1)}$$

These circuits are a main component of the tomographic scheme introduced in this work since they are used for the preparation of arbitrary MUB states and to measure in non-computational MUB basis. Hence, the efficiency achieved in this step will be reflected in that of the whole method.

The construction is divided into two stages:

(1) The construction, for a J basis, of the circuits T_a^b such that $T_a^b |\psi_b^J\rangle = |\psi_a^J\rangle$. This construction is trivial if the chosen J corresponds to the computational basis.

(2) The construction of the V_J^K for the same fixed J since every other change in basis can be built by combining two of these circuits via $V_L^M = V_J^M V_L^J = V_M^{J\dagger} V_L^J$.

With an efficient solution to the second stage, it will be possible to efficiently go from any state of any of the basis from the MUB set to any other state of any other basis from the set, thus solving the problem of averaging over a 2-design.

1. Circuits for the change in basis

In this section we will present an efficient construction for circuits that convert any state from the computational basis into the corresponding state of the basis stabilized by $G_{\vec{b}}$.

The problem of finding a change of basis quantum circuit to go from the computational basis to any other is equivalent

to that of finding a circuit that, under conjugation, transforms any tensor product of local Z operators—that is, the stabilizers of the computational basis—into the corresponding stabilizers of the target basis (i.e., the operators belonging to $G_{\vec{b}}$). In fact, we are looking for a unitary operator $V_{\vec{b}}$ such that

$$\forall \quad i, \forall \quad k \quad P_{\vec{b},k} V_{\vec{b}} |i\rangle = \pm V_{\vec{b}} |i\rangle, \qquad \text{(B2)}$$

where $P_{\vec{b},k} \in G_{\vec{b}}$ and $|i\rangle$ is the state i from the computational basis. So $V_{\vec{b}}^\dagger P_{\vec{b},k} V_{\vec{b}}$ has to be, for every k, a stabilizer operator for the computational basis.

The first step for the construction is to find a set of generators for the stabilizer group of the target basis. This can be done easily with definition (A3) using $j = 0, \ldots, n-1$, and it is efficient since it only requires $O(n^2)$ classical operations. Thanks to the fact that M is sparse allowing multiplication of vectors by M to be performed with only $O(n)$ operations. Then, in the following steps, each of the operators is going to be conjugated into stabilizers of the computational basis.

The second step is to choose the first operator from the generator group. That is, given the generator group in the form of Eq. (A3), we take the operator $P_{\vec{b},0}$, where $\vec{b}$ labels the generator group under consideration.

As the third step for the construction, single qubit rotations should be performed on each qubit to transform the operator into a product of single qubit Z and 1 as follows: for each qubit,

(i) if the operator has a 1 on the qubit in question, nothing should be done,

(ii) if the operator has a Z, nothing should be done,

(iii) if the operator has an X, Hadamard conjugation should be performed, and

(iv) if the operator has a Y, phase (for a single qubit in the computational basis it acts as $T|b\rangle = i^b |b\rangle$) and Hadamard conjugation should be performed.

This step is summarized as follows: the operator chosen should be conjugated by

$$S = \overset{n}{\underset{i=1}{\bigotimes}} R[(\mathbf{1}M^j)_i, (\mathbf{b}\hat{M}^j)_i], \qquad \text{(B3)}$$

where

$$R(0,0) = 1,$$

$$R(0,1) = H,$$

$$R(1,0) = 1,$$

$$R(1,1) = HT^\dagger, \qquad \text{(B4)}$$

and the subindex indicates the qubit in which each operator is acting.

So this transformed the first operator of the stabilizer into a product of Z and 1 for each single qubit using $O(n)$ quantum gates. The fourth step is to transform this product into a Z on the first qubit and identities on every other qubit. This is easily done via successive controlled-NOT conjugations, each
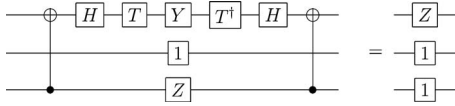
FIG. 4. Example of a first step of a change in basis circuit.

FIG. 5. Example of a change in basis circuit.

one with control on each of the qubits with Z except for the first one, and target on the first qubit.

So far we have conjugated the chosen operator via U_1 defined as

$$U_1 = \left\{ \prod_{i=2}^{n} [\text{controlled-NOT}(i,1)]^{[1-\delta_{(1M^j)_i,0}\delta_{(\vec{b}\hat{M}^j)_i,0}]} \right\}$$
$$\times \left\{ \bigotimes_{i=1}^{n} R[(1M^j)_i,(\vec{b}\hat{M}^j)_i] \right\}, \tag{B5}$$

transforming the first generator of the stabilizer group of the basis chosen into Z_1 using $O(n)$ quantum gates. However, all the other operators on the stabilizer group are changed due to the conjugation performed by U_1. The next step is to see how they are changed. This is easily done in the circuit formalism, constructing the circuit $U_1 P U_1^\dagger$ and permuting the controlled-NOT gates, and the rotations with the Pauli operators in P can be easily done and requires $O(n)$ classical operations for each of the $n-1$ remaining operators in $G_{\vec{b}}$.

It should be noticed, as it will be used later on, that the order of the generators is conserved. That is, the first generator of a stabilizer group corresponding to $j=0$ will be transformed into the first generator of another stabilizer group and so on.

We have found a way to transform the generators of $G_{\vec{b}}$ into the generators of a group $\tilde{G}_{\vec{b}}$ that has Z_1 as one of the stabilizers. So every other operator in $\tilde{G}_{\vec{b}}$ can have either a Z or a 1 on the first qubit. The remaining canonical generators (i.e., those with $j=1,\ldots,n-1$) will transform into operators having the identity on the first qubit. This way, the generator set $G_{\vec{b}}$ has been transformed into Z_1 and the generators of a stabilizer group of $n-1$ qubits in $O(n^2)$ classical operations and $O(n)$ quantum gates. Now we have to repeat the procedure from the second step and on for the remaining $n-1$ qubits to obtain the circuit for changing from the basis chosen into the computational basis. It is easy to see that the whole circuit will require $O(n^3)$ classical operations for its construction and $O(n^2)$ quantum gates.

2. Enlightening example

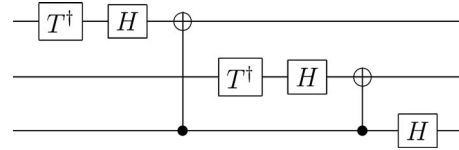We now illustrate the ideas of the method through an example. The zeroth step of the method is the choice of a primitive polynomial for the field $GF(2^n)$. In this example we will consider three qubits, i.e., $n=3$, and the one chosen for this example is $P(x)=1+x+x^3$. This polynomial gives the following companion matrix:

$$M = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}. \tag{B6}$$

And each basis of the MUB will be characterized by a $\vec{b}$, as shown on Eq. (A3). In this example we will consider the basis with $\vec{b}=(1,0,1)$. It is straightforward to find the corresponding stabilizer group generators,

$$P_{\vec{b},0} = Y \otimes 1 \otimes Z, \tag{B7}$$

$$P_{\vec{b},1} = 1 \otimes Y \otimes Z, \tag{B8}$$

$$P_{\vec{b},2} = Z \otimes Z \otimes Y. \tag{B9}$$

Once the generators for the stabilizer group are known we can follow the previously described steps. Take the first operator. Since it has a Y on the first qubit, conjugate it via $T^\dagger H$. Then apply a controlled-NOT with control in the third qubit and target in the first to transform this first operator into $Z \otimes 1 \otimes 1$, as shown in Fig. 4.

Next we have to see how the other generators of the stabilizer group transform under this circuit. It is not difficult to see that the transformed stabilizer group G' is

$$G' = \{Z \otimes 1 \otimes 1, 1 \otimes Y \otimes Z, 1 \otimes Z \otimes X\}. \tag{B10}$$

We have to repeat the above procedure for the last two generators. This will not modify the first generator since it only acts on the first qubit. This defines the recursive procedure that will, at last, generate the change in basis circuit, which will be the right side of the circuits. In Fig. 5 the full result is shown.

The iterative procedure described to generate the circuit is applied n times. Each iterative step incorporates $O(n)$ quantum gates, which means that the full circuit will use $O(n^2)$ single and two-qubit quantum gates. On the other hand, the classical overhead for obtaining a description of the circuit to apply requires $O(n^3)$ classical processing, giving rise to the dominant term in the efficiency of our method.

[1] A. Shabani and D. A. Lidar, Phys. Rev. Lett. **102**, 100402 (2009).

[2] A. Bendersky, F. Pastawski, and J. P. Paz, Phys. Rev. Lett. **100**, 190403 (2008).

[3] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

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

[5] M. Mohseni and D. A. Lidar, Phys. Rev. Lett. **97**, 170501 (2006).

[6] M. Mohseni and D. A. Lidar, Phys. Rev. A **75**, 062331 (2007).

[7] J. Emerson, M. Silva, O. Moussa, C. Ryan, M. Laforest, J. Baugh, D. G. Cory, and R. Laflamme, Science **317**, 1893 (2007).

[8] M. Silva, E. Magesan, D. W. Kribs, and J. Emerson, Phys. Rev. A **78**, 012347 (2008).

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

[10] C. Dankert, Master thesis, University of Waterloo, 2005.

[11] C. Dankert, R. Cleve, J. Emerson, and E. Livine, Phys. Rev. A **80**, 012304 (2009).

[12] A. Ambainis and J. Emerson, *Twenty-Second Annual Conference on Computational Complexity* (IEEE Computer Society, Los Alamitos, CA, 2007), Vol. 00, pp. 129–140.

[13] A. Klappenecker and M. Roetteler, Proceedings of the IEEE International Symposium on Information Theory, 2005 (IEEE, Adelaide, SA), pp. 1740–1744.

[14] P. Delsarte, J. M. Goethals, and J. J. Seidel, Geom. Dedic. **6**, 363 (1977).

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

[16] J. Lawrence, C. Brukner, and A. Zeilinger, Phys. Rev. A **65**, 032320 (2002).

[17] A. M. Bendersky, Master thesis, Universidad de Buenos Aires, 2006.

[18] C. Miquel, J. P. Paz, M. Saraceno, E. Knill, R. Laflamme, and C. Negrevergne, Nature (London) **418**, 59 (2002).

[19] F. Pastawski, Master thesis, Universidad Nacional de Córdoba, 2008.

[20] D. Gottesman, Ph.D. thesis, California Institute of Technology, 1997.

[21] D. Gottesman, e-print arXiv:quant-ph/9807006.

[22] D. Gottesman, in *Quantum Computation: A Grand Mathematical Challenge for the Twenty-First Century and the Millennium*, edited by S. J. Lomonaco, Jr. (American Mathematical Society, Providence, Rhode Island, 2002).

[23] W. K. Wootters, Ann. Phys. **176**, 1 (1987).

[24] J. P. Paz, A. J. Roncaglia, and M. Saraceno, Phys. Rev. A **72**, 012309 (2005).