[15] I. Kurniawan, G. Dirr, and U. Helmke, "Controllability aspects of quantum dynamics: A unified approach for closed and open systems," *IEEE Trans. Autom. Control*, vol. 57, no. 8, Aug. 2012.
[16] H. Yuan, "Characterization of majorization monotone quantum dynamics," *IEEE Trans. Autom. Control*, vol. 55, no. 4, pp. 955–959, Apr. 2010.
[17] C. O'Meara, G. Dirr, and T. Schulte-Herbrüggen, "Illustrating the Geometry of Coherently Controlled Unital Quantum Channels," Tech. Rep., 2011 [Online]. Available: http://arXiv.org/pdf/1103.2703
[18] K. H. Hofmann and W. A. F. Ruppert, *Lie Groups and Subsemigroups With Surjective Exponential Function*, ser. Memoirs Amer. Math. Soc. Providence, RI: American Mathematical Society, 1997, vol. 130.
[19] D. D'Alessandro, *Introduction to Quantum Control and Dynamics*. Boca Raton, FL: Chapman & Hall/CRC, 2008.
[20] G. Dirr and U. Helmke, "Lie theory for quantum control," *GAMM-Mitteilungen*, vol. 31, pp. 59–93, 2008.
[21] D. Elliott, *Bilinear Control Systems: Matrices in Action*. London, U.K.: Springer, 2009.
[22] V. Jurdjevic and H. Sussmann, "Control systems on Lie groups," *J. Diff. Equations*, vol. 12, pp. 313–329, 1972.
[23] C. Altafini, "Coherent control of open quantum dynamical systems," *Phys. Rev. A*, vol. 70, pp. 062321–062321, 2004.
[24] I. Kurniawan, "Controllability Aspects of the Lindblad-Kossakowski Master Equation—A Lie-Theoretical Approach," Ph.D. dissertation, Universität Würzburg, Würzburg, Germany, 2009.
[25] H. J. Sussmann, *Differential Geometric Control Theory*, ser. Progress in Mathematics, Real Analyticity, and Geometric Control. Boston, MA: Birkhäuser, 1983, pp. 1–116.
[26] R. Alicki and K. Lendi, *Quantum Dynamical Semigroups and Applications*, ser. Lecture Notes in Physics. Berlin, Germany: Springer, 1987, vol. 286.
[27] T. Ando, "Majorization, doubly stochastic matrices, and comparison of eigenvalues," *Lin. Multilin. Alg.*, vol. 118, pp. 163–248, 1989.
[28] A. Uhlmann, "Sätze über Dichtematrizen," *Wiss. Z. Karl-Marx-Univ. Leipzig, Math. Nat. R.*, vol. 20, pp. 633–637, 1971.
[29] A. Marshall, I. Olkin, and B. Arnold, *Inequalities: Theory of Majorization and its Applications*, 2nd ed. New York: Springer, 2011.
[30] K. L. Teo, C. J. Goh, and K. H. Wong, *A Unified Computational Approach to Optimal Control Problems*. New York: Longman, 1991.
[31] T. Schulte-Herbrüggen, A. Spörl, N. Khaneja, and S. J. Glaser, "Optimal control for generating quantum gates in open dissipative systems," *J. Phys. B*, vol. 44, pp. 154013–154013, 2011.
[32] M. Grace, C. Brif, H. Rabitz, I. Walmsley, R. Kosut, and D. Lidar, "Optimal control of quantum gates and suppression of decoherence in a system of interacting two-level particles," *J. Phys. B.*, vol. 40, pp. S103–S125, 2007.

Convex Optimization-Based Parameter Estimation and Experiment Design for Pauli Channels

Gábor Balló, *Student Member, IEEE*,
Katalin M. Hangos, *Member, IEEE*, and Dénes Petz

***Abstract*—In this technical note, we study the problem of parameter estimation for quantum channels, which are (completely positive) maps acting on quantum states. We focus on an important subclass of channels called Pauli channels which are characterized by a certain set of vectors (directions) and a number of scalar parameters. For the case where the directions are known, a special parametrization turns the parameter estimation problem into a convex optimization problem.**

For the case of unknown directions we give a simple algorithm to estimate the directions for qubit Pauli channels. These results assume that the identification experiment configuration is given, namely, a set of quantum states are given on which the channel acts, and a set of positive measurement (operator)s are fixed from which information is gathered.

In the second part of the technical note we consider the problem of determining the experiment design, namely, determining states and measurements for optimal parameter identification of the channels. We formalize this problem as a maximization problem for Fisher information and, assuming known channel directions, prove that this problem is convex. We also prove that the optimal states to be used in experiments are pure and the optimal measurements are extremal. For qubit Pauli channels we prove that both the optimal pure input states and projective measurements should be directed towards the channel directions. We illustrate the results of the technical note with numerical examples and simulations.

***Index Terms*—Optimization, quantum information and control, system identification.**

I. Introduction

Quantum systems are special stochastic nonlinear systems, where the stochasticity and nonlinearity are caused by the back-action of the measurements on the measured system [1], [2]. Therefore, even in the simplest static case, when the parameters of a non-dynamic quantum system are to be estimated, one needs special estimation methods [3]; this case is called state tomography in theoretical quantum physics.

Quantum channels are widely used information transfer devices in quantum information theory [1], that map an input quantum state into an output one usually in a static way. The task of the estimation of quantum channels – commonly known as *quantum process tomography* (QPT) in theoretical quantum physics – got a significant attention over about the last ten years. The problem was investigated by several authors [1], [4], [5]. The work [6] gives a comprehensive survey on the different strategies used for channel estimation.

Manuscript received January 26, 2011; revised July 06, 2011; accepted February 15, 2012. Date of publication May 03, 2012; date of current version July 19, 2012. This work was supported in part by the Hungarian Research Fund through grant K83440, and by the project TAMOP-4.2.2/B-10/1-2010-0025. Recommended by Associate Editor C. Altafini.

G. Balló is with the Department of Electrical Engineering and Information Systems, University of Pannonia, Veszprém H-8201, Hungary (e-mail: ballo.g@virt.uni-pannon.hu).

K. M. Hangos is with the Process Control Research Group, Computer and Automation Research Institute, Hungarian Academy of Sciences, Budapest H-1111, Hungary (e-mail: hangos@scl.sztaki.hu).

D. Petz is with the Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, Budapest H-1053, Hungary (e-mail: petz@renyi.hu; petz@math.bme.hu).

Color versions of one or more of the figures in this technical note are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2012.2195835

From a methodological point of view, there are two principally different approaches to the problem of quantum tomography, the *statistical approach* and the convex optimization based approach [4]. The former gives information on the statistics of the estimate and on its covariance matrix, but it has the drawback, that it leads to hard computations in higher dimensions. In spite of this, the majority of existing methods belong to this category. In contrast to this, an *optimization based method* does not give as much information, but it leads to relatively easy computations. This approach has been pursued in the work [7] where the problem of channel estimation is considered, assuming a completely general channel without any assumption on its inner structure. The author uses random input states, and random measurements on the output, and formulates a maximum likelihood problem. A similar method is used in the work [8], which formulates the task of process tomography as a least squares problem, which is convex. However, as it is stated in [9], it is a reasonable assumption to consider only a certain family of channels given with a model, based on a priori knowledge about the structure of the channel.

It is a commonly known fact that system identification is intimately related to *experiment design*, whose general aim is determining experimental conditions that result in good or even optimal identification results [10]. Thus the method of identification, that is, model parameter and structure estimation, determines the methods applied for experiment design, too. In addition, the nature and properties of the system to be identified have also a major influence on identification and experiment design. The experiment design for quantum channel parameter estimation includes the design of the quantum input to the channel, and the measurements to be applied on the resulting quantum output state. These are called the *experiment configuration*, together with the number of measurements to be performed in the different experiment configurations.

The results on experiment design for quantum state and channel estimation appear sparsely in the quantum state and process tomography literature. Most often, the authors investigate the optimality of their experiment configurations. The problem of optimal *experiment design for quantum state estimation* was first investigated by Kosut *et al.* [8] who developed methods using convex optimization for the determination of the number of measurements to be performed in the different experiment configurations. Since then, a few more papers can be found about optimal experiment design for quantum state estimation (see, e.g., [11] for a recent paper), but the problem is far from being solved for all cases.

The problem of finding an *optimal estimation of* one parameter *quantum channels* is discussed in [12], for different qubit (a two-level quantum system) input cases using statistical methods. An efficient estimation scheme is proposed in [13], where the quantum Fisher information and information geometrical considerations lead to an optimal measurement configuration for an important subclass of channels, the so called generalized Pauli channels. These are considered in this technical note as well (see Section II-B for their definition). A recent paper of [14] gives a good overview of the state-of-the-art in the field of optimal channel estimations.

Motivated by the above experiment design problems for quantum process tomography and by our recent work of optimization based quantum channel estimation [15], the aim of this study is to extend our work on parameter estimation using convex optimization. Moreover, we want to propose an experiment design method applying convex maximization for Pauli channels, that can be extended to generalized Pauli channels, too.

II. Basic Notions

Some basic notions for finite dimensional quantum systems [1], [2] are given below.

A. Quantum Measurements and Fisher Information

1) State Representation of Finite Dimensional Quantum Systems: The state of a finite dimensional quantum system is described by a so called density operator or *density matrix* ρ that acts on the underlying finite dimensional complex Hilbert space $\mathcal{H}$. Density matrices are self-adjoint positive semidefinite matrices with unit trace, i.e., $\rho \geq 0, \rho^* = \rho$, $\mathrm{Tr}(\rho) = 1$, where ρ^* denotes the adjoint of ρ.

Two-level quantum systems are called *quantum bits*, their density matrices are 2×2 complex matrices

$$\rho = \frac{1}{2}\begin{bmatrix} 1+\theta_3 & \theta_1 - \mathrm{i}\theta_2 \\ \theta_1 + \mathrm{i}\theta_2 & 1-\theta_3 \end{bmatrix} = \frac{1}{2}\left(I + \sum_{i=1}^{3}\theta_i\sigma_i\right), \quad \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \sigma_2 = \begin{bmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix}, \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{1}$$

are the so-called Pauli matrices defined in [2], and I is the unit matrix. The vector $\theta = (\theta_1, \theta_2, \theta_3)^{\mathrm{T}}$ is in the 3-dimensional unit ball of $\mathbb{R}^3$. This state representation is called *Bloch vector*. Let $|\phi_{i,1}\rangle$ and $|\phi_{i,2}\rangle$ be the normalized eigenvectors (column vectors) of σ_i ($i = 1, 2, 3$). Then

$$\rho = \frac{1}{2}\left(1 - \sum_{i=1}^{3}\theta_i\right) I + \sum_{i=1}^{3}\theta_i |\phi_{i,1}\rangle\langle\phi_{i,1}| \tag{2}$$

where $\langle\phi_{i,1}|$ denotes the conjugate transpose of $|\phi_{i,1}\rangle$. Note that we can uniquely represent any density matrix with its Bloch vector $\boldsymbol{\theta}$ given the normalized eigenvectors. A Bloch vector corresponds to a *pure state* if it has unit length, i.e., $\|\boldsymbol{\theta}\|_2 = 1$.

2) Quantum Measurements: Quantum measurements (observables) can be described mathematically as sets $\mathbf{M} = \{M_1, \ldots, M_m\}$, where the self-adjoint positive operators M_i act on the Hilbert space and $M_1 + \cdots + M_m = I$. Such $\mathbf{M} = \{M_1, \ldots, M_m\}$ is called to be a *positive operator-valued measure* (POVM). If the positive operators M_i are all projections, then we get a so called projective (or von Neumann) measurement. If a POVM $\mathbf{M}$ is performed as a measurement on the state ρ_θ parametrized by the Bloch vector, then the possible outcomes are $1, 2, \ldots, m$ and the probability of the outcome i is $\mathrm{Tr}(\rho_\theta M_i)$. POVMs also form a convex set. The extremal points of this set are called *extremal POVMs* [16].

3) Fisher Information: The Fisher information reflects the amount of information that a measured random variable can carry about the parameter $\theta = (\theta_1, \ldots, \theta_k)^{\mathrm{T}}$. In other words, it measures the accuracy of the unbiased estimator $\hat{\theta}$ of θ. Fisher information is a classical concept in statistics [2] and in system identification [10].

The Fisher information matrix for the quantum case [17] is

$$F(\theta)_{i,j} = \sum_{\ell} \frac{1}{\mathrm{Tr}(\rho_\theta M_\ell)} \frac{\partial}{\partial\theta_i}\mathrm{Tr}(\rho_\theta M_\ell) \frac{\partial}{\partial\theta_j}\mathrm{Tr}(\rho_\theta M_\ell) \tag{3}$$

and the Cramér-Rao matrix inequality describes its relation with the covariance matrix: $\mathrm{Var}(\hat{\theta}) \geq F(\theta)^{-1}$. This bound shows that the higher the Fisher information, the better estimation we can have. Note that the formula for F depends on the actual set of measurement observables $\mathbf{M}$ with which the experiments had been performed, i.e., $F(\theta, \mathbf{M})$.

B. Quantum Channels

Quantum channels model the information transfer between quantum systems, i.e., they transform the source quantum system into a target one. A quantum channel $\mathcal{E} : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$ is defined to be a completely positive and trace preserving (CPTP) map, where $\mathcal{B}(\mathcal{H}_i)$ is the operator algebra on the Hilbert space $\mathcal{H}_i$. This means that quantum channels map density matrices to density matrices.

A common way of representing quantum channels is the use of the *Choi matrix*. Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces, and $\mathcal{E} : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$ be a linear mapping that represents the quantum channel. To define the Choi matrix of $\mathcal{E}$ we take an orthonormal basis $f_1, \ldots, f_n$ in $\mathcal{H}_1$. Then $|f_i\rangle\langle f_j| \in \mathcal{B}(\mathcal{H}_1)$ and $\mathcal{E}$ acts on this operator. Then the Choi matrix [2] is

$$X_{\mathcal{E}} = \sum_{i,j} |f_i\rangle\langle f_j| \otimes \mathcal{E}\Big(|f_i\rangle\langle f_j|\Big) \in \mathcal{B}(\mathcal{H}_1) \otimes \mathcal{B}(\mathcal{H}_2) \qquad (4)$$

where $\otimes$ denotes the tensor product.

1) Pauli Channels: A notable wide class of quantum channels are the *Pauli channels*. In the qubit case when the input density matrix ρ has the form (1)

$$\mathcal{E}(\rho) = \frac{1}{2}\left(I + \sum_{i=1}^{3} \lambda_i \theta_i \sigma_i\right) \qquad (5)$$

defines the qubit Pauli channel.

An equivalent definition of a qubit Pauli channel acting on a density matrix (2) is

$$\mathcal{E}(\rho) = \frac{1}{2}\left(1 - \sum_{i=1}^{3} \lambda_i \theta_i\right) I + \sum_{i=1}^{3} \lambda_i \theta_i |\phi_{i,1}\rangle\langle\phi_{i,1}|. \qquad (6)$$

To describe this channel we need the three real constants λ_1, λ_2, λ_3 and the vectors $|\phi_{1,1}\rangle$, $|\phi_{2,1}\rangle$, $|\phi_{3,1}\rangle$. Below the vectors $|\phi_{1,1}\rangle$, $|\phi_{2,1}\rangle$, $|\phi_{3,1}\rangle$ will be called *channel directions*. The effect of the channel can then be described as depolarizing in each direction $|\phi_{i,1}\rangle$ with the corresponding parameter λ_i.

Definition (6) can be generalized for the multiple level case [18]. In this case one refers to these channels as *generalized Pauli channels*.

C. Quantum Channel Parameter Estimation as an Optimization Problem

The parameter estimation of quantum channels, or quantum process tomography is a widely investigated problem in mathematical physics. The pioneering works of Kosut and co-workers [8] formulated it as a convex optimization problem.

Consider an unknown quantum channel $\mathcal{E} : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$, which is to be estimated. We use a so called *tomography configuration* for this purpose that contains a known input quantum state with density matrix ρ and a set of positive operators (a POVM) $\mathbf{M}$. Note that we can use multiple different tomography configurations, i.e., different input states and observables in order to achieve better estimation on $\mathcal{E}$. In this work, the (input, POVM) pair corresponding to the γth configuration is denoted by $(\rho_\gamma, \mathbf{M}_\gamma)$.

1) The Least Squares Estimation: In order to have a convex optimization problem, the Choi matrix of the channel $\mathcal{E}$ is used as optimization variable. By the use of (4) we get for the probability of a measurement outcome $p_{\alpha,\gamma} = \mathrm{Tr}(C_{\alpha,\gamma} X_{\mathcal{E}})$, where $X_{\mathcal{E}}$ is the Choi matrix, and the *configuration matrix* $C_{\alpha,\gamma} = \rho_\gamma^{\mathrm{T}} \otimes M_{\alpha,\gamma}^*$ depends on the channel input ρ and on the measured POVM elements in configuration γ [15]. The probability $p_{\alpha,\gamma}$ can be estimated by using the relative frequency $\hat{p}_{\alpha,\gamma}$ that can be calculated from the measurement results. The variance of this unbiased estimate after n_γ independent measurements is known to be

$$\mathrm{Var}\Big(\hat{p}_{\alpha,\gamma}\Big) = \frac{1}{n_\gamma} p_{\alpha,\gamma}\Big(1 - p_{\alpha,\gamma}\Big) \qquad (7)$$

because $\hat{p}_{\alpha,\gamma}$ has a binomial distribution. This implies that for large n_γ, $\hat{p}_{\alpha,\gamma} \rightarrow p_{\alpha,\gamma}$ and $\mathrm{Var}\Big(\hat{p}_{\alpha,\gamma}\Big)$ tends to 0 as $n_\gamma \rightarrow \infty$, so $\hat{p}_{\alpha,\gamma}$ is a reasonable unbiased estimate of the real value $p_{\alpha,\gamma}$. This leads to formulating the parameter estimation as the following least squares problem:

$$\arg\min_{X_{\mathcal{E}}} \sum_{\alpha,\gamma} \left[\hat{p}_{\alpha,\gamma} - \mathrm{Tr}\Big(C_{\alpha,\gamma} X_{\mathcal{E}}\Big)\right]^2, \quad \text{so that} \quad X_{\mathcal{E}} \geq 0, \quad \mathrm{Tr}_2(X_{\mathcal{E}}) = I. \qquad (8)$$

This problem is a convex optimization problem in the Choi matrix $X_{\mathcal{E}}$ (see e.g., [8]), thus it can be solved relatively easily using existing numerical algorithms [19], [20].

2) Estimation of Pauli Channel Model Families With Known Channel Directions: The above derived least squares method gives an estimate of the elements of the Choi matrix. These depend in a complex way on the specific parameters of a given channel type, thus this method can suffer significantly from overparametrization. Some authors proposed approaches based on prior information on the channel, thus obtaining a well conditioned parameter estimation problem. These are mainly derived from physical interactions involved in the dynamics [14], and in [21] the authors also consider the problem of finding the optimal series of experiments to estimate the channel parameters.

As another possible solution, we can study the internal structure of the Choi matrix, and use this information to select more appropriate, model specific parameters for optimization. The natural choice would be to select just the unknown channel parameters. However it can be easily seen, that this choice would ruin convexity, as the Choi matrix can be an arbitrarily nonconvex function of these in the most general case. Instead of this, we approximate the Choi matrix by an affine structure. Let $h_1(\lambda), \ldots, h_m(\lambda)$ denote $\mathbb{R} \mapsto \mathbb{R}$ functions of the channel parameters, and let $H_0, H_1, \ldots, H_m$ denote constant Hermitian matrices. Then we can expand the Choi matrix as an affine function [15]

$$X_{\mathcal{E}} = \sum_k H_k h_k(\lambda) + H_0 \qquad (9)$$

and use the functions $h_k(\lambda)$ as optimization variables.

The above affine parameter estimation problem becomes really simple in the case of qubit *Pauli channels*. Assume that the channel directions are known or have been determined (see Section III later). Using (9), the Choi matrix of the qubit Pauli channel can be decomposed using simple constant Hermitian matrices [15], and the new optimization variables will be $h_1 = \lambda_1$, $h_2 = \lambda_2$, $h_3 = \lambda_3$. Thus, in this representation, *the parameter estimation of any two dimensional Pauli channel is a convex problem, as the optimization variables are exactly the channel parameters to be estimated.*

The above result remains valid for the higher dimensional generalized Pauli channel case, too [15].

III. Estimating the Channel Directions

During the parameter estimation of Pauli channels it is generally assumed that the Pauli channel directions are known. This, however, is not true in general so this section describes a method to estimate these directions *for the qubit case*. This also gives a first estimate on the parameters [15].

If we do not know the exact three channel directions $|\phi_{1,1}\rangle$, $|\phi_{2,1}\rangle$, $|\phi_{3,1}\rangle$ in which the Pauli channel is depolarizing, then quantum state estimation steps can be used to determine them.

Let us fix three vectors $|\varphi_{1,1}\rangle$, $|\varphi_{2,1}\rangle$, $|\varphi_{3,1}\rangle$ satisfying $|\langle\varphi_{i,1}|\varphi_{j,1}\rangle| = 1/2$, $i \neq j$. Then the operators $|\phi_{i,1}\rangle\langle\phi_{i,1}|$, $i = 1, 2, 3$ formed by the channel directions can also be expressed in the form of (2) on the basis determined by $|\varphi_{1,1}\rangle$, $|\varphi_{2,1}\rangle$, $|\varphi_{3,1}\rangle$ with Bloch vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. These vectors form a basis in $\mathbb{R}^3$. Let us further assume that the input qubit to the Pauli channel is represented by the Bloch vector $\mathbf{b}$ *in the* $\{\mathbf{v}_i\}$ *basis representing the channel directions*. In the rest of this section, the words "vector" and "state" are used as synonyms, both referring to Bloch vectors.

Using these notations, *the task is to estimate the three depolarizing directions of a qubit Pauli channel* $\mathcal{E}$. Let the set of found channel directions be $\mathbf{D}$. Let $\mathbf{D} = \{\}$ and $n = 0$, this is the initialization step. The following algorithm describes the direction estimation procedure.

Algorithm 1 Direction estimation

1: **repeat**

2: Prepare a pure state $\tilde{\mathbf{b}}^{(n)} \in \mathbf{D}^{\perp}$.

3: **repeat**

4: Put $\tilde{\mathbf{b}}^{(n)}$ into the composite channel $\mathcal{E}^k$ formed by cascading k instances of the channel $\mathcal{E}$, then get the output $\mathbf{b}^{(n+1)}$.

5: Perform quantum state tomography on $\mathbf{b}^{(n+1)}$.

6: Project $\mathbf{b}^{(n+1)}$ to $\mathbf{D}^{\perp}$ to get $\mathbf{b}^{(n+1)}_{\mathrm{proj}}$.

7: Normalize $\mathbf{b}^{(n+1)}_{\mathrm{proj}}$ to get the pure state $\tilde{\mathbf{b}}^{(n+1)}$.

8: Increase n by 1.

9: **until** The distance $\left\|\tilde{\mathbf{b}}^{(n)} - \tilde{\mathbf{b}}^{(n+1)}\right\|$ is smaller than some prescribed value.

10: Put $\tilde{\mathbf{b}}^{(n+1)}$ into $\mathbf{D}$, set n to 0.

11: **until** Dimension of $\mathbf{D}^{\perp}$ is 0.

We give the *mathematical arguments* that support the steps of the above algorithm for the most common case when all of the λ_i channel parameters have different values, and $|\lambda_i| < 1$. We refer to [22] for a discussion of the general case. We recall that the effect of the channel for the input Bloch vector $\mathbf{b} = \sum_{i=1}^{3} b_i \mathbf{v}_i$ ($\|\mathbf{b}\| \le 1$) can be written as $\mathcal{E}(\mathbf{b}) = \sum_{i=1}^{3} \lambda_i b_i \mathbf{v}_i$.

A. Step 4

Assume we use a pure state $\tilde{\mathbf{b}}^{(n)}$ with $\left\|\tilde{\mathbf{b}}^{(n)}\right\| = 1$ as an input to the channel. Then the output $\mathbf{b}^{(n_1)}$ can be expanded in the $\{\mathbf{v}_i\}$ basis, thus $\mathbf{b}^{(n_1)} = \sum_{i=1}^{3} \lambda_i \tilde{b}_i^{(n)} \mathbf{v}_i$. Let the channel parameter with the largest absolute value be λ_m. As $|\lambda_i| < 1$ from the positivity and trace preserving constraints, the absolute value of component $\tilde{b}_m^{(n)}$ will decrease the least among nonzero components of $\tilde{\mathbf{b}}^{(n)}$. If we continued this procedure, and put the channel output $\mathbf{b}^{(n_1)}$ back into the channel as input to get the output $\mathbf{b}^{(n_2)}$, then the sequence $\left\{\mathbf{b}^{(n_\ell)} / \left\|\mathbf{b}^{(n_\ell)}\right\|\right\}_{\ell=0}^{\infty}$ would be a Cauchy sequence, thus would converge to the direction $\mathbf{v}_m$ that corresponds to the parameter with the largest absolute value. However, the length $\left\|\mathbf{b}^{(n_\ell)}\right\|$ of the sequence will converge to zero.

B. Steps 5 and 7

To avoid this to happen, we normalize the output Bloch vector $\mathbf{b}^{(n+1)} = \mathbf{b}^{(n_k)}$ after each step. First we perform quantum state tomography in Step 5 to get an estimate $\hat{\mathbf{b}}^{(n+1)}$, and normalize it in Step 7 to obtain the pure state $\tilde{\mathbf{b}}^{(n+1)} = \hat{\mathbf{b}}^{(n+1)} / \left\|\hat{\mathbf{b}}^{(n+1)}\right\|$. This will be put again in the channel. This way, the Cauchy sequence of vectors will indeed converge to $\mathbf{v}_m$.

C. Steps 2, 6, 10 and 11

After the first channel direction $\mathbf{v}_m$ was found using this procedure, we can continue the search in the plane $\mathbf{D}^{\perp}$ orthogonal to $\mathbf{v}_m$ (Steps 10 and 2). However, due to the inaccuracies in state tomography, the direction we will find will not be exactly $\mathbf{v}_m$, rather some vector $\tilde{\mathbf{b}}^* \approx \mathbf{v}_m$. Thus, it is more robust if we apply a projection to the output vector in Step 6, onto the subspace $\mathbf{D}^{\perp}$. When the second direction is found, then the third can be easily obtained, as it will be the one orthogonal to both the first and the second direction. Thus the direction estimation procedure is finished in only two iteration steps (Step 11).

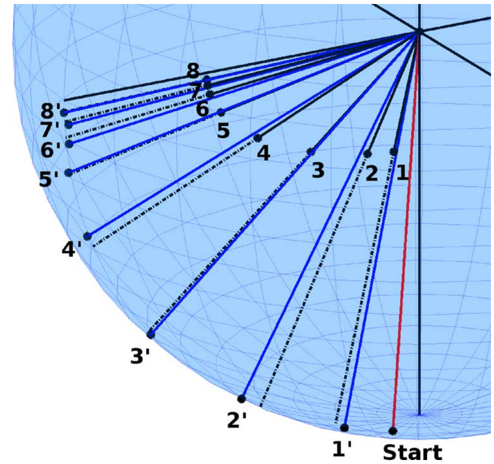

Fig. 1. Channel direction estimation example for the qubit channel with parameters $\lambda_1 = 0.6$, $\lambda_2 = 0.3$ and $\lambda_3 = 0.1$ estimating the direction $m = 1$. The unknown channel directions are shown by the black axes in the Bloch sphere. The numbered vectors indicate the channel output (n) and its perturbed and normalized form (n') in the nth step. The starting input vector was chosen randomly at the beginning of the search for each direction.

D. A Simple Numerical Example

In order to illustrate the operation and properties of the above proposed channel direction estimation algorithm, a simple illustrative numerical example is presented here for a qubit channel with different parameters $\lambda_1 = 0.6$, $\lambda_2 = 0.3$ and $\lambda_3 = 0.1$.

The three unknown channel directions were chosen to be the eigenvectors of the Pauli matrices. The uncertainty in the estimated channel output state arising from quantum state estimation was simulated using random perturbations in the output state. The perturbation for the ith direction is a random term added to the Bloch vector component θ_i, and it is of the form $\xi\sqrt{(1-\theta_i)/N}$, where ξ is a random number taken form the standard normal distribution, N is the number of measurements in the state tomography step, and $(1-\theta_i)/N$ is the variance of the estimator $\hat{\theta}_i$.

The result of the numerical test can be seen in Fig. 1. It can be seen from the figure, that the sequence of input states converges to the channel direction of the highest absolute parameter value in the subspace of searching in a few iteration steps.

IV. Experiment Design in the Known Channel Direction Case

The field of experiment design for quantum channel parameter estimation has not matured yet. Even the main problems have not been formalized completely. Only a few papers exist that aim at determining the elements of the tomography configuration, i.e., the input quantum state and the measurement observables, (see, e.g., [12], [13]). These papers, however, fix one of the elements – the input quantum state, for example – and determine the other (say the observable) according to some optimality criteria. To the best of our knowledge the only paper that treats experiment design using convex optimization solves a restricted problem, i.e., the determination of the number of measurements to be performed in the different experiment configurations [8].

A. The General Case

Suppose we have a quantum channel $\mathcal{E}_\lambda$ with some fixed channel parameter vector λ. The task is to *find the input state ρ and a measurement observable $\mathbf{M}$ for which the Fisher information $F(\lambda)$ of the channel parameters estimated from the channel output $\mathcal{E}_\lambda(\rho)$ using the observable $\mathbf{M}$ is maximal*. By maximization of $F(\lambda)$ we mean the maximization of an appropriately selected scalar function of $F(\lambda)$. The following statement summarizes our result.

Statement 1: The optimal input state will be pure, and the optimal measurement (POVM) will be an extremal POVM.

In order to prove this, we use the definition of the Fisher information matrix (3) expressed with the configuration matrix $C_\alpha = (\rho^{\mathrm{T}} \otimes M_\alpha)$ and the Choi matrix X_λ of the channel to obtain

$$[F(\lambda)]_{i,j} = \sum_\alpha \frac{1}{\mathrm{Tr}(C_\alpha X_\lambda)} \frac{\partial}{\partial \lambda_i} \mathrm{Tr}\left(C_\alpha X_\lambda\right) \frac{\partial}{\partial \lambda_j} \mathrm{Tr}\left(C_\alpha X_\lambda\right) .$$

A scalar valued objective function is needed for the maximization, thus we take the trace of the Fisher information matrix

$$\tilde{F}(\lambda) = \sum_{i,\alpha} \frac{1}{\mathrm{Tr}(C_\alpha X_\lambda)} \left(\frac{\partial}{\partial \lambda_i} \mathrm{Tr}\left(C_\alpha X_\lambda\right) \right)^2 \tag{10}$$

using property $\mathrm{Tr}(A) \le \mathrm{Tr}(B)$ whenever $A \le B$ for the Hermitian matrices A and B. It can be shown that the function $\tilde{F}$ is convex in the configuration matrix C_α on the set of valid C_α matrices, thus convex both in the input ρ and in the set of measurement observable $\mathbf{M}$ if we fix the other to be a constant. From this it follows that $\tilde{F}$ should attain its maximum at an extremal point of the feasible region containing the possible experiment configurations.

B. Optimal Configuration for Qubit Pauli Channels

In the special case of qubit Pauli channels, a stronger statement can be made when the three depolarizing directions of the Pauli channel are known. Because of the rotational symmetry of the Bloch ball, the obtained results can be applied to any other Pauli channel, with different directions. The experiment design problem is solved for *projective measurements*.

Statement 2: The three Pauli channel directions can be used as optimal directions for both measurements and input states.

For the proof we recall, that projective measurements can be represented with two-element extremal observables $\{|\psi\rangle\langle\psi|, I - |\psi\rangle\langle\psi|\}$ [16]. Let these projections be represented with the Bloch vectors $\mathbf{m}$ and $-\mathbf{m}$ with $\|\mathbf{m}\|_2 = 1$. Let also the pure input state be in Bloch parametrization (1), with the Bloch vector denoted as $\mathbf{b}$. Then the channel output with channel parameter vector $\lambda = [\lambda_1, \lambda_2, \lambda_3]^{\mathrm{T}}$ will be (5), and if we write the trace of the Fisher information matrix of the channel parameters, we get

$$\tilde{F}(\lambda) = \frac{m_1^2 b_1^2 + m_2^2 b_2^2 + m_3^2 b_3^2}{1 - (m_1 b_1 \lambda_1 + m_2 b_2 \lambda_2 + m_3 b_3 \lambda_3)^2} . \tag{11}$$

Recall that the unit length requirement on the vectors $\mathbf{b}$ and $\mathbf{m}$ follows from the convexity of $\tilde{F}$, which we want to maximize. Note also that the above formula is a special case of (10).

Let us now define the vector $\mathbf{c} = [m_1 b_1, m_2 b_2, m_3 b_3]^{\mathrm{T}}$, which is the configuration vector of the channel estimation problem including not only the input state and measurement information, but also the assumptions on the channel structure. The objective (11) will then be

$$\tilde{F}(\lambda) = \frac{\mathbf{c}^{\mathrm{T}}\mathbf{c}}{1 - (\mathbf{c}^{\mathrm{T}}\lambda)^2} = \frac{\mathbf{c}^{\mathrm{T}}\mathbf{c}}{1 - \mathbf{c}^{\mathrm{T}}\lambda\lambda^{\mathrm{T}}\mathbf{c}}.$$

By Hölders inequality it is easy to see that the set of all possible $\mathbf{c}$ vectors forms an octahedron inside the Bloch sphere, whose vertices are the unit vectors pointing to the three directions of the channel. Thus the set of all $\mathbf{c}$ vectors is convex, moreover we have equality if and only if $|b_i|^2 = |m_i|^2$, i.e., when the vectors $\mathbf{b}$ and $\mathbf{m}$ are parallel.

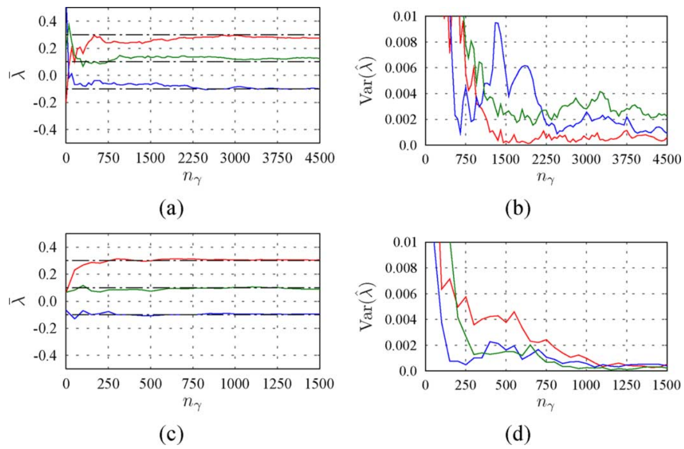

Fig. 2. Estimation with nonoptimal ((a) and (b)) and optimal ((c) and (d)) configuration for the channel parameters $\lambda_1 = 0.3$, $\lambda_2 = -0.1$, $\lambda_3 = 0.1$.

As the objective is convex in both $\mathbf{b}$ and $\mathbf{m}$ and thus in $\mathbf{c}$ we know that it attains its maximum at a vertex of the octahedral feasible set. Thus the optimal $\mathbf{c}$ has not only unit 1-norm, but unit 2-norm, too. This can only happen if only one component of $\mathbf{c}$ is nonzero, which means that both the input and the measurement have to be in the same channel direction. This implies that the objective is maximized if the direction of $\mathbf{c}$ is that direction, for which $|\lambda_i|$ is maximal. Let this be for example λ_1, then the optimal objective will be $\tilde{F}(\lambda) = 1/(1 - \lambda_1^2)$.

Now, we see that performing experiments in this direction does not give any information on the other directions, so we have to search for additional experiment configurations. Let the direction of the optimal configuration found first be the direction x, i.e., for $i = 1$. If we now constrain the objective (11) to the plane orthogonal to x, then we get the constraints $m_1 = 0$ and $b_1 = 0$. Using the same derivation as in the general three dimensional case, we get that the next optimal configuration will be the y (with $i = 2$) or z (with $i = 3$) direction, and so on.

C. Numerical Example

Simulation experiments were used to analyse the effect of experiment design on the performance of the numerical optimization based estimation of *qubit Pauli channels*. Results were generated in MATLAB environment, using simulated random measurement data. The optimization problem (8) was solved using YALMIP modelling language [23] and the SDPT3 solver [24].

1) Tomography Configurations: Two configurations have been used for comparison purposes. A *nonoptimal configuration* applied the so called minimal POVM ($\gamma = 1$) described by [25] with pure state $(1/\sqrt{3})[1, 1, 1]^{\mathrm{T}}$ as input state. The total number of measurements was $n_{\mathrm{tot}} = n_\gamma = 4500$.

In the *optimal experiment configuration* both the input and measurement were optimal with respect to the channel directions. Here $\gamma = 3$, and the number of measurements in each direction was $n_\gamma = 1500$.

2) Simulation Results: Each experiment setup was repeated five times and their average was taken. The channel parameters in each test were $\lambda_1 = 0.3$, $\lambda_2 = -0.1$, $\lambda_3 = 0.1$.

As a result, the empirical mean $\bar{\lambda}$ and the empirical variances of the estimated parameters $\mathrm{Var}(\hat{\lambda})$ were computed and plotted in Fig. 2 for both the nonoptimal and the optimal cases. The results indicate, that the efficiency of the optimal experiment configuration highly outperforms the nonoptimal one. We can also see that in the optimal setting, we can reach a very accurate estimation with only about $n_\gamma = 1000$ number of measurements in each configuration.

The detailed simulation results of both the qubit and the generalized Pauli channel cases can be found in [22], together with a simple robustness analysis in the qubit case.

V. Conclusion

Convex optimization-based parameter estimation and convex maximization-based experiment design methods were proposed in this technical note for Pauli channels. The development is based on our earlier method [15] for the parameter estimation of Pauli channels with known channel directions. This requires the solving of a purely convex optimization problem.

We have first proposed an efficient iterative method of estimating the channel directions for the qubit Pauli channel case. The extension of this method to the general higher dimensional case is a possible direction of our further work.

An experiment design procedure based on maximizing the Fisher information of the output of a quantum channel was also presented. It was shown that the Fisher information is a convex function both in the input and in the measurement parameters. This way we proved that the optimal input state is pure and the measurement POVM is extremal. For qubit Pauli channels this formulation leads to an optimal setting that includes pure input states and projective measurements directed towards the channel directions.

References

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge, U.K.: Cambridge Univ. Press, 2000.
[2] D. Petz, *Quantum Information Theory and Quantum Statistics*. Berlin, Germany: Springer-Verlag, 2008, Theoretical and Mathematical Physics.
[3] M. Paris and J. Rehácek, *Quantum State Estimation*. Berlin, Germany: Springer, 2004, vol. 649.
[4] G. M. D'Ariano, M. G. A. Paris, and M. F. Sacchi, "Quantum tomography," *Adv. Imag. Electron Phys.*, vol. 128, p. 205, 2003.
[5] M. G. A. Paris and J. Rehácek, *Quantum State Estimation*. Berlin, Germany: Springer Verlag, 2004, vol. 649.
[6] M. Mohseni, A. T. Rezakhani, and D. A. Lidar, "Quantum process tomography: Resource analysis of different strategies," *Phys. Rev. A*, vol. 77, p. 032322, 2008.
[7] M. F. Sacchi, "Maximum-likelihood reconstruction of completely positive maps," *Phys. Rev. A*, vol. 63, no. 5, p. 054104, Apr. 2001.
[8] R. Kosut, I. A. Walmsley, and H. Rabitz, "Optimal experiment design for quantum state and process tomography and hamiltonian parameter estimation," *ArXiv:quant-ph*, vol. 0411093, pp. 1–51, 2004.
[9] M. Sasaki, M. Ban, and S. M. Barnett, "Optimal parameter estimation of a depolarizing channel," *Phys. Rev. A*, vol. 66, no. 2, p. 022308, Aug. 2002.
[10] L. Ljung, *System Identification: Theory for the User*. Upper Saddle River, NJ: Prentice-Hall, 1999.
[11] J. Nunn, B. J. Smith, G. Puentes, I. A. Walmsley, and J. S. Lundeen, "Optimal experiment design for quantum state tomography: Fair, precise, and minimal tomography," *Phys. Rev. A*, vol. 81, p. 042109, 2010.
[12] M. Sarovar and G. Milburn, "Optimal estimation of one-parameter quantum channels," *J. Phys. A: Math. General*, vol. 39, p. 8487, 2006.
[13] A. Fujiwara and H. Imai, "Quantum parameter estimation of a generalized pauli channel," *J. Phys. A: Math. General*, vol. 36, pp. 8093–8103, 2003.
[14] M. Branderhorst, J. Nunn, I. Walmsley, and R. Kosut, "Simplified quantum process tomography," *New J. Phys.*, vol. 11, p. 115010, 2009.
[15] G. Balló and K. M. Hangos, "Parameter estimation of quantum processes using convex optimization," in *Proc. 19th Int. Symp. Math. Theory Netw. Syst. (MTNS'10)*, Apr. 2010, pp. 2043–2050.
[16] G. D'Ariano, P. Lo Presti, and P. Perinotti, "Classical randomness in quantum measurements," *J. Phys. A Math. General*, vol. 38, pp. 5979–5991, Jul. 2005.
[17] R. D. Gill and S. Massar, "State Estimation for Large Ensembles," Tech. Rep., 2002 [Online]. Available: arXiv:quant-ph/9902063v2
[18] D. Petz and H. Ohno, "Generalizations of Pauli channels," *Acta Math. Hungar*, vol. 124, pp. 165–177, 2009.
[19] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Rev.*, vol. 38, no. 1, pp. 49–95, 1996.
[20] K. Audenaert and B. De Moor, "Optimizing completely positive maps using semidefinite programming," *Phys. Rev. A*, vol. 65, no. 3, p. 030302, Feb. 2002.
[21] K. C. Young, M. Sarovar, R. Kosut, and K. B. Whaley, "Optimal quantum multiparameter estimation and application to dipole- and exchange-coupled qubits," *Phys. Rev. A*, vol. 79, no. 6, p. 062301, Jun. 2009.
[22] G. Balló and K. M. Hangos, "Experiment design and parameter estimation of pauli channels using convex optimization," *ArXiv Quantum Phys. e-prints*, pp. 1–29, Jul. 2011.
[23] J. Löfberg, "Yalmip : A toolbox for modeling and optimization in MATLAB," in *Proc. CACSD Conf.*, Taipei, Taiwan, 2004, pp. 284–289.
[24] K. C. Toh, M. Todd, and R. Tütüncü, "SDPT3 – A MATLAB software package for semidefinite programming," *Optim. Methods Softw.*, vol. 11, pp. 545–581, 1998.
[25] J. Rehácek, B.-G. Englert, and D. Kaszlikowski, "Minimal qubit tomography," *Phys. Rev. A*, vol. 70, no. 5, p. 052321, 2004.