Parameter estimation of quantum processes using convex optimization

Gábor Balló, Attila Magyar and Katalin M. Hangos

Abstract— **A convex optimization based method is proposed for quantum process tomography, in the case of known channel model structure, but unknown channel parameters. The main idea is to select an affine parametrization of the Choi matrix as a set of optimization variables, and formulate a semidefinite programming problem with a least squares objective function. Possible convex relations between the optimization variables are also taken into account to improve the estimation.**

Simulation case studies show, that the proposed method can significantly increase the accuracy of the parameter estimation, if the channel model structure is known. Beside the convex part, the determination of the channel parameters from the optimization variables is a nonconvex step in general. In the case of Pauli channels however, the method reduces to a purely convex optimization problem, allowing to obtain a globally optimal solution.

I. INTRODUCTION

In these days, the building of quantum computers that can be used to solve realistic, large scale problems – including the breaking of cryptographic codes and the simulation of complex quantum systems – has two main difficulties from the theoretical point of view. The first is decoherence, in other words the unavoidable coupling of quantum systems with their environment, which leads to the altering or even the complete destruction of the quantum system's state, thus causing errors in the calculation. The other main obstacle is the lack of ability to fully manipulate and extract full information from the physical system, i.e. the control and estimation of the quantum states and channels, called quantum processes. This latter problem has to be handled by the formalism of the modern methods of system- and control theory. It is certainly not a trivial problem, because in quantum mechanics, one has to face the difficulty of the threatment of measurements. Namely, that no measurement can be carried out on a quantum system without disturbing the state of the system itself.

The task of the estimation of quantum channels – commonly known as quantum process tomography (QPT) [1] – got a significant attention over about the last ten years. The problem was investigated by several authors [1], [2], [3]. The work [4] gives a comprehensive survey on the different strategies used for process tomography (or channel estimation). The problem can essentially be formulated in two type of methods: direct, and indirect. In the indirect method, we trace the problem back to quantum state tomography, i.e. the information about the unknown quantum channel is

G. Balló and A. Magyar are with the Department of Electrical Engineering and Information Systems, University of Pannonia, Veszprém, Hungary `{ballo.g,amagyar}@almos.uni-pannon.hu`

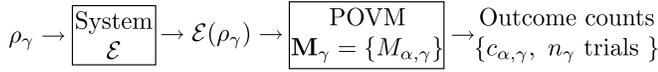K. M. Hangos is with Computer and Automation Research Institute, Budapest, Hungary `hangos@scl.sztaki.hu`

obtained by sending known probe quantum systems through the channel, and performing state tomography on the output states. In contrast, in the direct method, the experiments directly give information about the channel, without the need for a state tomography step.

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 [2]. The former gives information on the statistics of the estimate and on its covariance matrix, but it has the drawback, that it is hard to compute in higher dimensions. In spite of this, majority of the existing methods belong to this category.

In contrast to this, an *optimization based method* does not give as much information, but it is relatively easy to compute. This approach has been pursued in the work [5] where the problem of channel estimation (in the form of the Choi matrix) is considered, assuming a completely general channel, thus without any assumption on the inner structure of the Choi matrix. 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 [6], which formulates the task of process tomography as a least squares problem, which is convex. It also uses the Choi matrix as optimization variable, thus searches the optimal channel in the convex set of all CPTP maps using multiple input-measurement pairs.

However, as it is stated in [7], 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. The aim of this work is to develop a method which is capable of incorporating these constraints into the channel tomography problem, while still remaining – at least partially in the general case – solvable by convex optimization.

The rest of this paper is organised as follows. In Section II the formal description of quantum process tomography is given, and a least squares objective is derived for the solution. In Section III the variables and constraints of the convex optimization problem are considered, and the extraction of the channel parameters is discussed. In Section IV the procedure is demonstrated on several case studies. Finally, in Section V the conclusions are given, and the direction of further work is discussed.

II. QUANTUM PROCESS TOMOGRAPHY AS AN OPTIMIZATION PROBLEM

In this work, the indirect procedure of process tomography is followed, mostly relying on the work [6]. The formal mathematical description of process tomography, i.e. the

$$\rho_\gamma \rightarrow \boxed{\begin{array}{c} \text{System} \\ \mathcal{E} \end{array}} \rightarrow \mathcal{E}(\rho_\gamma) \rightarrow \boxed{\begin{array}{c} \text{POVM} \\ \mathbf{M}_\gamma = \{M_{\alpha,\gamma}\} \end{array}} \rightarrow \begin{array}{c} \text{Outcome counts} \\ \{c_{\alpha,\gamma}, \ n_\gamma \text{ trials }\} \end{array}$$

Fig. 1: The scheme of data collection for process tomography.

so called *tomography configuration* contains the following elements:

- A known input density operator ρ on the Hilbert space $\mathcal{H}$ of the system.
- The unknown quantum channel $\mathcal{E} : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H})$, which is to be estimated. The channel can be written for example in the Kraus representation, in which case the output of $\mathcal{E}$ is $\sigma = \mathcal{E}(\rho) = \sum_i E_i \rho E_i^\dagger$, where the E_i operators are the *operator elements* of the channel $\mathcal{E}$. These must satisfy the trace preserving constraint $\sum_i E_i^\dagger E_i = \hat{1}$, where $\hat{1}$ is the identity operator.[1]
- A set of POVMs. A POVM $\mathbf{M} = \{M_\alpha\}$ is a set of positive operators, with which we can perform quantum measurement on the channel output state σ. In order to be able to uniquely identify the channel output state, an important requirement in quantum process tomography is that the measurements must be *tomographically complete*. This means that the measured POVMs must provide all the information about the output state, and thus on the channel. Such a set of measurement operators is sometimes called a *quorum*.

Note that we can use multiple different tomography configurations, i.e. different input states and POVMs in order to achieve better estimation on $\mathcal{E}$. In this work, the input-POVM pair corresponding to the γth configuration is denoted by ρ_γ and $\mathbf{M}_\gamma$.

A. Data collection

The first stage of process tomography is the collection of the measurement data into a measurement record. The measurements are performed in each γ configuration n_γ times independently. This scheme can be seen in Fig. 1.

During data collection, the different α outcomes of the measurements in the configuration γ are counted in the variable $c_{\alpha,\gamma}$, and put in the measurement record $\mathbf{D}$. Then obviously $\sum_\alpha c_{\alpha,\gamma} = n_\gamma$. Thus, we have to perform a total number of $n_{\text{tot}} = \sum_\gamma n_\gamma$ independent measurements. The estimator $\hat{\mathcal{E}}$ of the channel $\mathcal{E}$ will be calculated from these measurement outcomes, based on some estimation procedure.

B. Least Squares estimation

The next step of the tomography problem is to choose a suitable estimation procedure. The least squares is a popular method, because it is easy to implement, and it results in an estimation procedure, which can be used in higher dimensions, and for multiple parameters. So in this work it is used for quantum process estimation. The exact form of the LS objective function for the process tomography problem is derived in the following.

If we write the $\sigma_\gamma = \mathcal{E}(\rho_\gamma)$ output of the channel in Kraus representation, then the probability density function of the measurement results can be modeled by

$$p_{\alpha,\gamma} = \text{Tr}(\sigma_\gamma M_{\alpha,\gamma}) = \text{Tr}\left(\sum_i E_i^\dagger M_{\alpha,\gamma} E_i \rho_\gamma\right). \quad (1)$$

In this expression, the optimization variables would be the Kraus operator elements E_i. However, as these do not form a convex set, the resulting LS optimization problem would be nonconvex. To overcome this difficulty, we can choose the Choi matrix as optimization variable.

A common definition of the Choi matrix is through the use of the Jamiołkowsky isomorphism [8]. Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces, and let A be a bounded linear operator in $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$. Then we can associate A with a vector in the Hilbert-space $\mathcal{H}_1 \otimes \mathcal{H}_2$:

$$|A\rangle\!\rangle := \sum_i A|i_1\rangle|i_2\rangle \ ,$$

where $\{|i_1\rangle\}$ and $\{|i_2\rangle\}$ are the respective bases of the spaces $\mathcal{H}_1$ and $\mathcal{H}_2$.[2] By this, noting that in our case $\mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H}$ is the Hilbert space of the input state, the Choi matrix of the channel $\mathcal{E}$ will be

$$X_\mathcal{E} := \sum_k |E_k\rangle\!\rangle\langle\!\langle E_k| \ ,$$

where the vector $|E_k\rangle\!\rangle$ in $\mathcal{H} \otimes \mathcal{H}$ is associated to the operator E_k. The condition of complete positivity can then be written as $X_\mathcal{E} \geq 0$, and the channel is trace preserving if $\text{Tr}_{\mathcal{H}_{(2)}}(X_\mathcal{E}) = \hat{1} \in \mathcal{L}(\mathcal{H}_{(1)})$.

The scalar product of these "double-ket" vectors is defined in a natural way:

$$\langle\!\langle A|B\rangle\!\rangle := \text{Tr}(A^\dagger B) \ , \quad (2)$$

what is by definition the Hilbert–Schmidt scalar product of the two operators. We can also identify a useful rule easily[3]:

$$(A \otimes B)|C\rangle\!\rangle = |BCA^\text{T}\rangle\!\rangle \ , \quad (3)$$

Now, using the above, we can continue the derivation of (1). By the use of relations (2) and (3) we get

$$\begin{aligned} p_{\alpha,\gamma} &= \sum_i \text{Tr}\left(E_i^\dagger M_{\alpha,\gamma} E_i \rho_\gamma\right) = \\ &= \sum_i \langle\!\langle M_{\alpha,\gamma}^\dagger E_i | E_i \rho_\gamma\rangle\!\rangle = \\ &= \sum_i \langle\!\langle E_i | (\hat{1} \otimes M_{\alpha,\gamma}^\dagger)(\rho_\gamma^\text{T} \otimes \hat{1}) | E_i\rangle\!\rangle = \\ &= \sum_i \text{Tr}\big[\underbrace{(\rho_\gamma^\text{T} \otimes M_{\alpha,\gamma}^\dagger)}_{R_{\alpha,\gamma}} |E_i\rangle\!\rangle\langle\!\langle E_i|\big] = \\ &= \text{Tr}(R_{\alpha,\gamma} X_\mathcal{E}) \ , \end{aligned}$$

where $X_\mathcal{E}$ is the Choi matrix of the channel $\mathcal{E}$, and the operator $R_{\alpha,\gamma}$ depends on the channel input ρ and on the measured POVM elements in configuration γ.

---

[1] Here $\dagger$ denotes the adjoint of the operator.

[2] The $|.\rangle\!\rangle$ notation indicates that these vectors represent operators.
[3] Here $^\text{T}$ denotes the transposition.

Proceeding to the derivation of the LS objective function, a statistics for estimating the probability $p_{\alpha,\gamma}$ can be given by the

$$\hat{p}_{\alpha,\gamma} = \frac{c_{\alpha,\gamma}}{n_\gamma} \qquad (4)$$

relative frequency that can be calculated from the measurement results. The variance of this unbiased estimate after n_γ independent measurements is known to be

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

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

$$\arg\min_{X_\mathcal{E}} V_{\mathrm{LS}}(X_\mathcal{E}) = \sum_{\alpha,\gamma} \mathrm{Tr}\big(P^{\mathrm{emp}}_{\alpha,\gamma} - R_{\alpha,\gamma} X_\mathcal{E}\big)^2 \ , \qquad (5)$$

$$\text{so that} \quad X_\mathcal{E} \geq 0, \quad \mathrm{Tr}_\mathcal{H}(X_\mathcal{E}) = \hat{1}$$

where the diagonal matrix $P^{\mathrm{emp}}_{\alpha,\gamma} = \hat{p}_{\alpha,\gamma} \frac{\hat{1}_{\mathcal{H}\otimes\mathcal{H}}}{\dim(\mathcal{H}\otimes\mathcal{H})}$ is defined for simplification purpose.

This problem is a convex optimization problem in the Choi matrix $X_\mathcal{E}$, thus it can be solved relatively easily using existing numerical algorithms [9], [10].

III. ESTIMATION OF CHANNEL MODEL FAMILIES

It is reasonable to assume that we know a model type of the channel, and only the unknown values of the parameters of this model have to be estimated. In such a problem, the above derived least squares objective cannot be used directly, as it assumes a completely general channel model, and estimates the elements of the Choi matrix. Thus, if the task is to estimate some specific model parameters, this method can suffer significantly from overparametrization.

As a 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. Effectively, this should reduce the set of optimal solutions of problem (5) to solutions, which are consistent with the desired model family.

A. The complex LMI constraint

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.

Thus, instead of this, the following method is used. Let $h_1(\theta), \dots, h_m(\theta)$ denote functions of the channel parameters, and let $H_0, H_1, \dots, H_m$ denote constant Hermitian matrices. Then we can expand the Choi matrix as an affine function:

$$X_\mathcal{E} = \sum_k H_k h_k(\theta) + H_0 \qquad (6)$$

It can be seen, that if we consider the functions $h_k(\theta)$ as optimization variables, we can replace the completely unknown matrix variable $X_\mathcal{E}$ in problem (5) with a matrix, which has a known structure up to an affine approximation.

Thus, the positivity constraint in (5) turns into a complex LMI constraint, and the trace preserving constraint can in turn be omitted, as if an exact channel model is known, it can always be taken into account in the model construction.

B. Convex constraints

It can occur, that some of the functions $h_k(\theta)$ depend on one or more other functions. If this relation f is convex, then we can use it to define additional constraints to make the estimation more accurate. We can determine a minimal set $\mathcal{V} = \{h_l(\theta)\}$, where $l \in I_\mathcal{V}$. The set $I_\mathcal{V}$ contains the indexes of those functions among $h_k(\theta)$ ($k = 1, \dots, m$), which do not depend on each other through a convex relation, thus are independent in this sense. These can in general be written as

$$h_k(\theta) = f^{\mathrm{cvx}}\big(h_1(\theta), \dots, h_{m'}(\theta)\big),$$
$$h_l(\theta) \in \mathcal{V}, \quad l = 1, \dots, m'$$

Thus, if we have convex relations between the functions $h_k(\theta)$ chosen as optimization variables, then the identification of these relations can help to reduce the number of these variables, which can be helpful in the parameter extraction step (see subsection III-C.).

However, these relations cannot be added simply as equations to the problem (5), because a strictly convex function with equality is not a convex constraint. In these cases we can relax the equality, and add the convex constraint to the problem as inequality.

If n_c such constraints are found, then we obtain the following optimization problem (omitting the trace preserving constraint):

$$\arg\min_{\mathbf{h}(\theta)} V_{\mathrm{LS}}(X_\mathcal{E}) \ ,$$
$$\text{so that} \quad X_\mathcal{E} \geq 0 \ , \qquad (7)$$
$$\text{and} \quad f^{\mathrm{cvx}}_k\big(h_1(\theta), \dots, h_{m'}(\theta)\big) \leq h_k(\theta), \quad k = 1, \dots, n_c$$

After solving this problem we get an optimal value $v_0^* = V_{\mathrm{LS}}(X_\mathcal{E}^*)$ for the objective, where the * denotes the optimality. Then we formulate for each k the following auxiliary problem:

$$\arg\min_{\mathbf{h}(\theta)} h_k(\theta) \ ,$$
$$\text{so that} \quad X_\mathcal{E} \geq 0 \ , \qquad (8)$$
$$\text{and} \quad f^{\mathrm{cvx}}_k\big(h_1(\theta), \dots, h_{m'}(\theta)\big) \leq h_k(\theta), \quad k = 1, \dots, n_c \ ,$$
$$\text{and} \quad v_k = V_{\mathrm{LS}}(X_\mathcal{E}) \leq v^*_{k-1}$$

By solving these for each k in succession, we make sure that the additional convex constraints in (7) get as close to equality as possible, thus making the parameter estimation procedure more accurate. However it would require further studies to see whether this method always guarantees a global optimum in the original objective, and whether it is reachable in finite steps. In practice, the effectiveness of this

heuristics is based only on the fact, that in case of the optimal solution (in the model of the real channel), all of the convex constraints are saturated.

C. Determining the model parameters

After we performed all the needed optimization steps and obtained an optimal Choi matrix $X_{\mathcal{E}}^*$ together with the optimal variables h_k^*, the next step of parameter estimation is to fit h_k^* to the predefined optimization variables $h_k(\theta)$. In essence, this problem can be written as

$$\arg \min_{\theta} \sum_{l \in I_{\mathcal{V}}} \left(h_l^* - h_l(\theta) \right)^2 , \tag{9}$$

where $I_{\mathcal{V}}$ denotes the index set of the independent optimization variables. This is a nonlinear least squares problem, which is not even convex in the general case, because of the $h_l(\theta)$ functions. If however these functions are convex, then the problem (9) will be convex too. Note that having fewer independent optimization variables results in fewer nonlinear equations.

IV. CASE STUDIES

The aim of the simulation experiments was to analyze the effect of additional constraints on the performance of the numerical optimization based estimation of quantum channels. Results were generated in MATLAB environment, using simulated random measurement data. The optimization problem (7) and (8) were solved using YALMIP modeling language [11] and the SDPT3 solver [12]. The extraction of the model parameters, i.e. the solving of (9) was in general done by the built in nonlinear least squares solver of MATLAB, solving it multiple times using random initial conditions to achieve a global minimum with high probability.

A. Tomography configurations

The experiments were set up as follows.
- The used input states were all pure states.
- To obtain a tomographically complete measurement, appropriate POVMs were selected on the Hilbert space of the system, and were used for measurement. In two dimensions, these were the *Pauli matrices*, and observables with similar properties in higher dimensions. Each of these can be decomposed into a POVM, which can be used in one configuration. Note that the tomographically complete measurement in this setting may not require an operator basis of observables to be used. The knowledge about the channel structure may allow us to identify parts of the state space where the effect of the channel is the same, thus requiring only less observables (see case study in subsection IV-D). The only important thing is to get information on all of the channel parameters.
- The total n_{tot} number of measurements was distributed among all the configurations equally, i.e. for each configuration γ, an equal number of experiment were used.

Each experiment setup was repeated five times and their average was taken. Each of the estimated process Choi matrices $X_{\mathcal{E}}$ and channel parameters were analyzed using the following four estimation performance measuring quantities:
- Fidelity of the estimated and ideal channel output: $F\big(\hat{\mathcal{E}}(\rho), \mathcal{E}(\rho)\big)$. The fidelity of two quantum states is defined by

$$F(\rho, \sigma) = \mathrm{Tr} \sqrt{\rho^{\frac{1}{2}} \sigma \rho^{\frac{1}{2}}} .$$

- The empirical mean $\bar{\theta}$ of the estimated parameters $\hat{\theta}$.
- The empirical covariance matrix of the estimated parameters $\hat{\theta}$. This could be used only when the empirical covariance could be made diagonal with appropriately selected measurement observables, because in other cases, the estimated parameters were correlated. One way to obtain a diagonal empirical covariance is, that the output state parametrization has to be compatible with the channel parametrization in the sense that each state parameter is affected only by one channel parameter. In this case we can use that set of operators for measurement in which the state parametrization is given, and through this compatibility, we will be able to estimate the channel parameters independently [13]. Thus, only the diagonal elements of the covariance matrix, i.e. the variances of the parameters were computed:

$$\mathrm{Var}(\hat{\theta}_i) = \frac{1}{4} \sum_{j=1}^{5} (\hat{\theta}_{i,j} - \bar{\theta}_i)^2 ,$$

using that each setup was repeated 5 times.
- Hilbert–Schmidt norm of the estimation error: $\|\hat{X} - X\|$.

B. The generalized amplitude damping channel

The Choi matrix of the generalized amplitude damping channel is

$$X_{\mathcal{E}} = \begin{bmatrix} 1 - \gamma + p\,\gamma & 0 & 0 & \sqrt{1-\gamma} \\ 0 & \gamma - p\,\gamma & 0 & 0 \\ 0 & 0 & p\,\gamma & 0 \\ \sqrt{1-\gamma} & 0 & 0 & -p\,\gamma + 1 \end{bmatrix} .$$

Here the parameter γ gives the strength of the noise effect in the interval $[0, 1]$, and p parameterizes the output state for $\gamma = 1$ also in the interval $[0, 1]$. This is $|0\rangle$ when $p = 1$, and $|1\rangle$ when $p = 0$ in the computational basis.

Using (6), the above Choi matrix can be decomposed in the following way:

$$H_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad H_1 = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$H_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \quad H_3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix},$$

and the optimization variables will be

$$h_1 = \gamma, \quad h_2 = p\gamma, \quad h_3 = \sqrt{1-\gamma} .$$

Notice that there is a convex relation between these variables, so we can further reduce their number: $h_3^2 - 1 = -h_1$.

If we now put these into the problem (7), we get the following optimization problem:

$$\arg\min_{\mathbf{h}(\theta)} v_0 = V_{\mathrm{LS}}(X_{\mathcal{E}}) \,,$$

$$\text{so that} \quad X_{\mathcal{E}} \geq 0, \quad \text{and} \quad h_3^2 - 1 \leq -h_1$$

The type (8) auxiliary problem will be:

$$\arg\min_{\mathbf{h}(\theta)} -h_1 \,,$$

$$\text{so that} \quad X_{\mathcal{E}} \geq 0, \ h_3^2 - 1 \leq -h_1 \text{ and } V_{\mathrm{LS}}(X_{\mathcal{E}}) \leq v_0^*$$

After solving these, the final step is the parameter extraction using (9). As the independent set $\mathcal{V} = \{h_1, h_2\}$ of optimization variables contains nonconvex functions, this problem will also be nonconvex. However, in this simple two dimensional case, we can see that the solution of (9) will be explicitly $\gamma = h_1$ and $p = \frac{h_2}{h_1}$.

The input state used in this example was the pure state with the Bloch vector $\frac{1}{\sqrt{3}}[1,1,1]^{\mathrm{T}}$. The applied observables were the Pauli matrices

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

These form a tomographically complete set of observables for the two dimensional case. Each was used as a two-element POVM in a different configuration, thus a total of three configurations were used. Unfortunately, this set does not allow the independent estimation of the parameters for this channel, thus the empirical covariance matrix cannot be used in this case as performance indicator.

The resulting estimations, and the used characteristic quantities are plotted in Fig. 2. The chosen (γ, p) pair was $(0.7, 0.3)$. More examples can be found in [14].

From the results, we can conclude that the estimated parameters are very accurate, even on the border of the parameter domain (see [14]).

C. T-representation of qubit Pauli channels

In the two dimensional case, we have the possibility to expand the quantum channels in the Pauli basis, thus giving the affine map of the channel, which takes Bloch vectors to Bloch vectors [15]:

$$\begin{bmatrix} 1 & 0 \\ t & T_3 \end{bmatrix}$$

where T_3 can be assumed diagonal with elements α, β, and γ without loss of generality.

If we restrict ourselves to Pauli channels, i.e. we set $t = 0$, then the Choi matrix of a quantum channel in this T-parametrization is

$$X_{\mathcal{E}} = \begin{bmatrix} \frac{1+\gamma}{2} & 0 & 0 & \frac{\alpha+\beta}{2} \\ 0 & \frac{1-\gamma}{2} & \frac{\alpha-\beta}{2} & 0 \\ 0 & \frac{\alpha-\beta}{2} & \frac{1-\gamma}{2} & 0 \\ \frac{\alpha+\beta}{2} & 0 & 0 & \frac{1+\gamma}{2} \end{bmatrix}.$$

Here, the allowed values for the three parameters α, β and γ for the channel to send density matrices to density matrices are

$$-1 \leq \alpha, \beta, \gamma \leq 1 \,,$$

and the positivity constraint is satisfied if and only if $|1 \pm \gamma| \geq |\alpha \pm \beta|$.

Using (6), the Choi matrix of this channel can be decomposed in the following way:

$$H_0 = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad H_1 = \frac{1}{2}\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix},$$

$$H_2 = \frac{1}{2}\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad H_3 = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

and the optimization variables will be

$$h_1 = \alpha, \quad h_2 = \beta, \quad h_3 = \gamma \,.$$

There are no convex relations in this case, so only the solution of (5) is required.

The fact, that the optimization variables h_k are exactly the channel parameters themselves, makes the parameter extraction step (9) also unnecessary for this channel. Thus, we can see that in this representation, the estimation of any two dimensional Pauli channel is a convex problem.

The input state and the POVMs were the same in this experiment as in subsection IV-B, so the same three configurations were used. Note however, that in this example, the set of observables are not only tomographically complete, but also make it possible to estimate each channel parameter independently, resulting in a diagonal empirical covariance matrix.

The resulting estimations, and characteristic quantities are plotted in Fig. 3. The chosen (α, β, γ) vector was $(0.4, 0.1, -0.5)$. For more examples, see [14].

As we can see, the parameter estimation in this convex case is very simple and accurate.

D. The generalized Pauli channel in 3 dimension

The generalized Pauli channel is discussed in [13]. It is a generalization of the Pauli channels for qubits into higher dimensions. Its definition for a general Hermitian operator A taken from the operator algebra $\mathcal{L}(\mathcal{H})$:

$$\mathcal{E}(A) = \left(1 - \sum_{i=1}^{u} \lambda_i\right) \frac{\mathrm{Tr}(A)}{d} \hat{\mathbb{1}} + \sum_{i=1}^{u} \lambda_i E_i(A) \,,$$

where d is the dimension of the input quantum system, and u is the number of *complementary subalgebras*[4] of $\mathcal{L}(\mathcal{H})$ used in the channel construction. Let $\mathcal{A}_1, \ldots, \mathcal{A}_u$ be pairwise complementary subalgebras, which linearly span the whole

---

[4] Two subalgebras $\mathcal{A}_1$ and $\mathcal{A}_2$ of $\mathcal{L}(\mathcal{H})$ are called complementary if the traceless subspaces of $\mathcal{A}_1$ and $\mathcal{A}_2$ are orthogonal with respect to the Hilbert–Schmidt inner product [16].
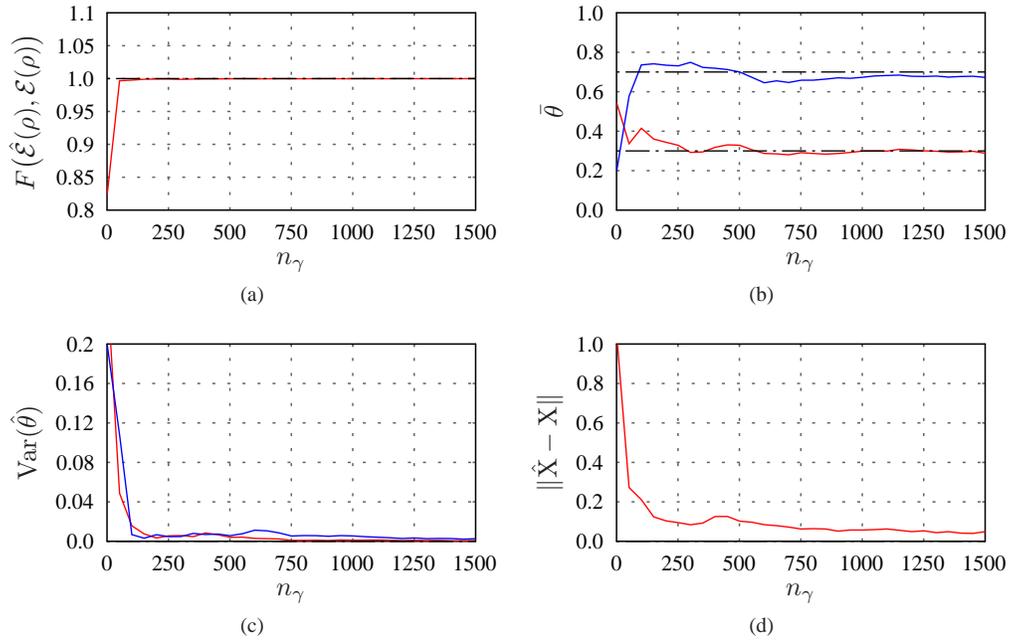
Fig. 2: The parameters are $\gamma = 0.7$ and $p = 0.3$. (a) shows the fidelity of the estimated and ideal channel output, (b) shows the expected value of the estimated parameters, (c) shows the variance of the estimated parameters, and (d) shows the Hilbert–Schmidt norm of the estimation error.
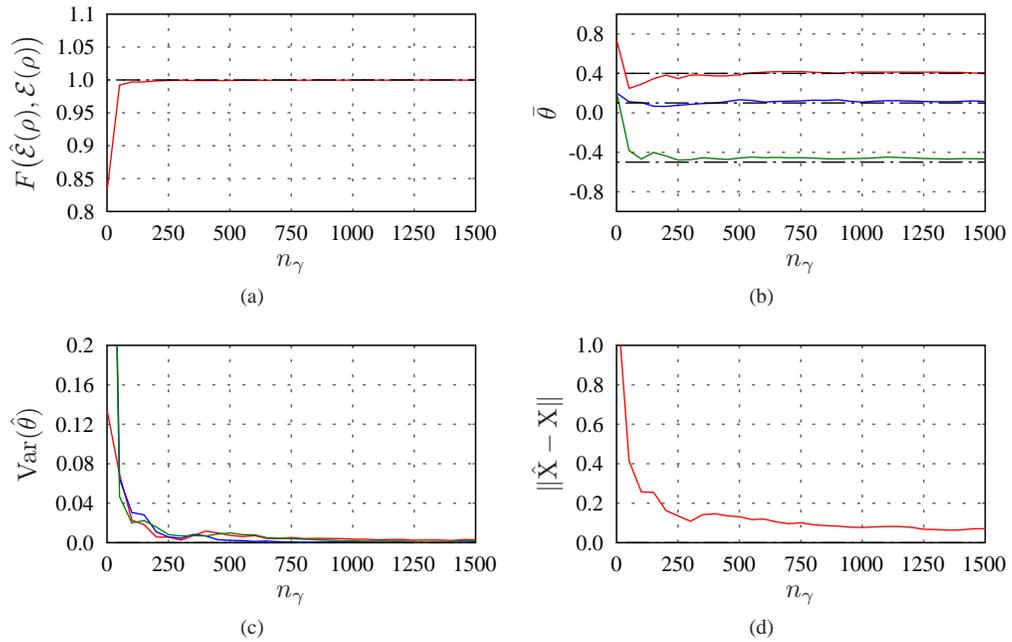

Fig. 3: The parameters are $\alpha = 0.4$, $\beta = 0.1$ and $\gamma = -0.5$. (a) shows the fidelity of the estimated and ideal channel output, (b) shows the expected value of the estimated parameters, (c) shows the variance of the estimated parameters, and (d) shows the Hilbert–Schmidt norm of the estimation error.

algebra $\mathcal{L}(\mathcal{H})$. Then E_i is the orthogonal projection with respect to the Hilbert–Schmidt inner product onto $\mathcal{A}_i$.

Here, we consider the case when all of the complementary subalgebras are maximal Abelian. In this case the set of subalgebras correspond to a set of orthonormal bases called *mutually unbiased bases* (MUB). Two orthonormal bases $\{|\psi_{1,j}\rangle\}$ and $\{|\psi_{2,j}\rangle\}$ are mutually unbiased if and only if the maximal Abelian subalgebras $\mathcal{A}_1$ and $\mathcal{A}_2$ containing operators diagonal in the bases $\{|\psi_{1,j}\rangle\}$ and $\{|\psi_{2,j}\rangle\}$ are complementary.

Let $\{|\psi_{1,j}\rangle\}, \ldots, \{|\psi_{u,j}\rangle\}$ be u sets of bases which are pairwise mutually unbiased. Then the orthogonal projection E_i can be constructed as

$$E_i(A) = \sum_{j=1}^{d} \langle\psi_{i,j}|A|\psi_{i,j}\rangle|\psi_{i,j}\rangle\langle\psi_{i,j}| \ .$$

The effect of the generalized Pauli channel is thus the following. It projects the input state onto each subalgebra $\mathcal{A}_i$, then scales the results with the corresponding channel parameter λ_i. Trivially, if $\lambda_i = 0$ for all i, then the result is the completely mixed state $\frac{1}{d}\hat{1}$. In contrast, if $\lambda_i = 1$ for all i, then we get back the input state as expected.

The channel is trace preserving by construction, and completely positive if and only if the parameters λ_i satisfy

$$1 + d\lambda_i \geq \sum_{j} \lambda_j \geq \frac{-1}{d-1} \ , \qquad (10)$$

A set of MUB can be constructed, for example, as follows [16]. Let $|\varphi_0\rangle, \ldots, |\varphi_{d-1}\rangle$ be a basis. Let X and Z be unitary operators such that

$$X|\varphi_i\rangle = |\varphi_{i+1 \bmod (d-1)}\rangle \ ,$$
$$Z|\varphi_i\rangle = \mathrm{e}^{\mathrm{i}k\frac{2\pi}{d}}|\varphi_i\rangle \ .$$

Then the unitaries $S_{j,k} := Z^j X^k$ $(j,k = 0, \ldots, d-1)$ are pairwise orthogonal with respect to the Hilbert–Schmidt inner product. If we select the matrices X, Z, and also each one in the form ZX^k, $k = 1, \ldots, d-1$, then the eigenvectors of each of these unitaries form a set of MUB. Moreover, if d is prime, then this procedure results in exactly $d+1$ MUB, which is the maximum number of MUB possible.

In the present example, we choose $d = 3$, thus the used MUB will be the bases obtained using the eigenvectors of the matrices X, Z, ZX, ZX^2. Then the Choi matrix of the channel will be the following:

$$X_{\mathcal{E}} = \frac{1}{3}\begin{bmatrix} f_1 & 0 & 0 & 0 & f_3 & 0 & 0 & 0 & f_3 \\ 0 & f_2 & 0 & 0 & 0 & f_4 & f_4^* & 0 & 0 \\ 0 & 0 & f_2 & f_4^* & 0 & 0 & 0 & f_4 & 0 \\ 0 & 0 & f_4 & f_2 & 0 & 0 & 0 & f_4^* & 0 \\ f_3 & 0 & 0 & 0 & f_1 & 0 & 0 & 0 & f_3 \\ 0 & f_4^* & 0 & 0 & 0 & f_2 & f_4 & 0 & 0 \\ 0 & f_4 & 0 & 0 & 0 & f_4^* & f_2 & 0 & 0 \\ 0 & 0 & f_4^* & f_4 & 0 & 0 & 0 & f_2 & 0 \\ f_3 & 0 & 0 & 0 & f_3 & 0 & 0 & 0 & f_1 \end{bmatrix} ,$$

where

$$f_1 = 1 + 2\lambda_2, \quad f_2 = 1 - \lambda_2, \quad f_3 = \lambda_1 + \lambda_3 + \lambda_4,$$
$$f_4 = \lambda_1 - \frac{\lambda_3}{2}(1 + \mathrm{i}\sqrt{3}) - \frac{\lambda_4}{2}(1 - \mathrm{i}\sqrt{3}),$$

Note that the $n = 2$ case using the same MUB selection method is exactly the same as the T-representation of Pauli channels in subsection IV-C.

The decomposition (6) of the Choi matrix can be calculated easily based on the previous examples, given that the optimization variables are exactly the channel parameters:

$$h_i = \lambda_i, \quad i = 1, \ldots, 4$$

This, and the construction of the channel shows that the problem of parameter estimation is convex, and solvable using only (5) in any such – arbitrarily high – dimension, in which the channel itself can be defined.

The input state used in this experiment was the pure state

$$|\Psi\rangle = \frac{1}{\sqrt{6}}\big(|\psi_{1,1}\rangle + |\psi_{2,1}\rangle + |\psi_{3,1}\rangle + |\psi_{4,1}\rangle\big)$$

which is constructed from the first vector $|\psi_{j,1}\rangle$ of each basis of the MUB, on the analogy of the Bloch vector $\frac{1}{\sqrt{3}}[1,1,1]^{\mathrm{T}}$ in the two dimensional case. Based on this analogy, this state is expected to be sufficient for the characterisation of the channel parameters.

As we have discussed before, each channel parameter λ_i affects the length of the projected input state $E_i(\rho)$ in the subalgebra $\mathcal{A}_i$ independently. The channel has no other effect. Thus, we can measure the effect of the channel by focusing only on the subalgebras, and estimate the parameters independently. For this purpose, any Hermitian operator belonging to the subalgebra $\mathcal{A}_i$ can be considered as observable, and can be used for the estimation of λ_i. The reason for this is that the effect of the channel is the same on the whole subalgebra, so it is enough to measure any direction inside $\mathcal{A}_i$ to get information on λ_i. Moreover, the independence of the measurements on different subalgebras result in a diagonal covariance matrix.

For example, valid observables for each subalgebra can be constructed using the MUB in the following way:

$$A_i = \sum_{j=1}^{d} j|\psi_{i,j}\rangle\langle\psi_{i,j}|$$

This way we get observables with d different eigenvalues. Beyond that, the specific value of the eigenvalues are irrelevant, as we are interested only in the outcome probabilities.

As it can be seen from the above, a total of four configurations were used in the experiments. The resulting estimations, and characteristic quantities are plotted in Fig. 4. The chosen $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ vector was $(-0.3, -0.2, -0.1, 0.1)$. See [14] for more examples.

We can see that in the case of a qutrit channel with a high number of parameters, this method can provide very accurate parameter estimation, even on the border of the parameter domain [14]. In some of the images, the Hilbert–Schmidt norm does not seem to decrease over the given range
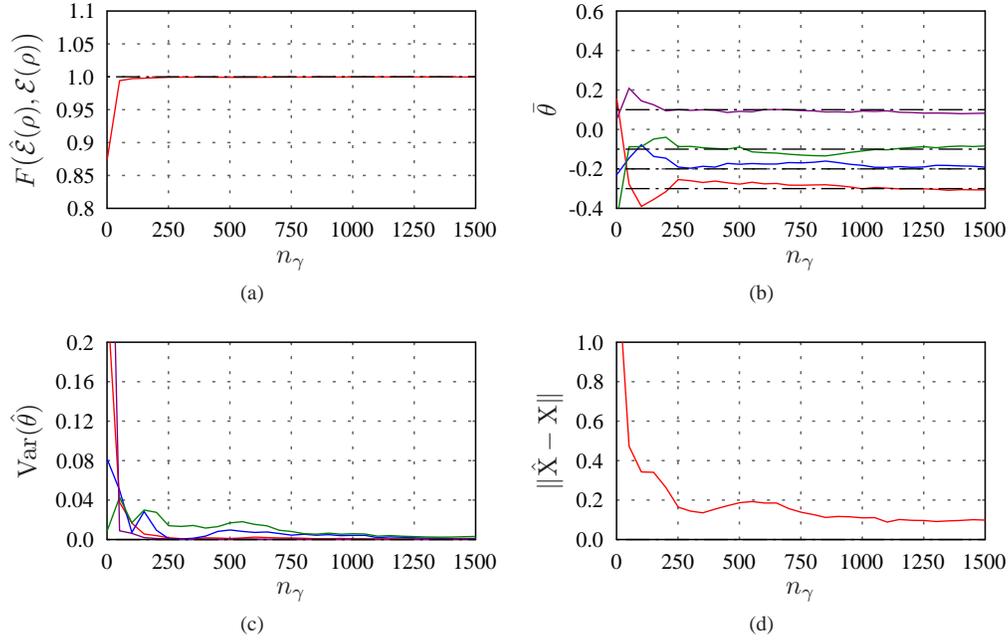
Fig. 4: $\lambda_1 = -0.3$, $\lambda_2 = -0.2$, $\lambda_3 = -0.1$ and $\lambda_4 = 0.1$. These saturate the second inequality in (10), so this is again a border point. (a) shows the fidelity of the estimated and ideal channel output, (b) shows the expected value of the estimated parameters, (c) shows the variance of the estimated parameters, and (d) shows the Hilbert–Schmidt norm of the estimation error.

of the measurement number n_γ. Results from experiments performed with higher measurement numbers show that the reason of this is slower convergence.

V. CONCLUSION

In this work, we introduced a method for the parameter estimation of quantum channel model families, based on convex optimization.

Using simulation case studies it has been shown, that the proposed method of affine decomposition of the Choi matrix can significantly increase the accuracy of the parameter estimation for the case of a known channel model structure. In addition, with performing auxiliary optimization problems, convex relations between optimization variables can also be exploited to improve the estimation.

In some of the cases, particularly that of Pauli channels, this method results in a purely convex optimization problem, thus we can obtain a globally optimal estimation with relatively simple numerical algorithms. In the general case however, beside the convex part, the method may need a nonconvex optimization step as well.

ACKNOWLEDGEMENT

REFERENCES

[1] M. A. NIELSEN, I. L. CHUANG, *Quantum Computation and Quantum Information*, Cambridge University Press (2000)

[2] G. M. D'ARIANO, M. G. A. PARIS, M. F. SACCHI, Quantum Tomography, *Advances in Imaging and Electron Physics*, **128**: 205 (2003)

[3] M. G. A. PARIS, J. ŘEHÁČEK (eds.), *Quantum State Estimation*, *Lecture Notes in Physics, Berlin Springer Verlag*, vol. 649 (2004)

[4] M. MOHSENI, A. T. REZAKHANI, D. A. LIDAR, Quantum Process Tomography: Resource Analysis of Different Strategies, *Physical Review A*, **77**: 032322 (2008)

[5] M. F. SACCHI, Maximum-likelihood reconstruction of completely positive maps, *Phys. Rev. A*, **63, 5**: 054104 (2001)

[6] R. KOSUT, I. A. WALMSLEY, H. RABITZ, Optimal Experiment Design for Quantum State and Process Tomography and Hamiltonian Parameter Estimation (2004), `quant-ph/0411093`

[7] M. SASAKI, M. BAN, S. M. BARNETT, Optimal parameter estimation of a depolarizing channel, *Phys. Rev. A*, **66, 2**: 022308 (2002)

[8] A. JAMIOLKOWSKI, Linear transformations which preserve trace and positive semidefiniteness of operators, *Reports on Mathematical Physics*, **3**: 275–278 (1972)

[9] L. VANDENBERGHE, S. BOYD, Semidefinite Programming, *SIAM Review*, **38, 1**: 49–95 (1996)

[10] K. AUDENAERT, B. DE MOOR, Optimizing completely positive maps using semidefinite programming, *Phys. Rev. A*, **65, 3**: 030302 (2002)

[11] J. LÖFBERG, YALMIP : A Toolbox for Modeling and Optimization in MATLAB, in *Proceedings of the CACSD Conference*, Taipei, Taiwan (2004)

[12] K. C. TOH, M. TODD, R. TÜTÜNCÜ, SDPT3 - a MATLAB software package for semidefinite programming, *Optimization Methods and Software*, **11**: 545–581 (1998)

[13] D. PETZ, H. OHNO, Generalizations of Pauli channels, *Acta Math. Hungar*, **124**: 165–177 (2009), `0812.2668`

[14] G. BALLÓ, K. M. HANGOS, Parameter estimation of quantum processes using convex optimization, *ArXiv e-prints* (2010), `1004.5209`

[15] D. PETZ, *Quantum Information Theory and Quantum Statistics*, Theoretical and Mathematical Physics, Springer-Verlag (2008)

[16] D. PETZ, Complementarity in quantum systems, *Reports on Mathematical Physics*, **59**: 209–224 (2007), `arXiv:quant-ph/0610189`