# THESIS

# Optimal tomography of two-level and generalized Pauli channels


**Dániel Virosztek**


Supervisors:


**Katalin Hangos**
Research Professor, Process Control Research Group,
Computer and Automation Research Institute, Budapest, Hungary

**Dénes Petz**
Professor, Department of Mathematical Analysis,
Budapest University of Technology and Economics, Budapest, Hungary


MŰEGYETEM 1782

Budapest University of Technology and Economics

**2013**

# Contents

# Chapter 1

# Introduction

The accurate description of different quantum phenomena is a key issue in their potential use in modern IT-technologies, for example communication, security, or quantum computers. The parameter estimation of quantum channels, which is commonly called quantum process tomography [7], plays a major role in quantum information processing.

In quantum mechanics, both dynamical changes and communication is treated using *quantum channels*. These are nothing else but trace preserving completely positive mappings $\mathcal{E}$ which transform the input state $\rho$ given on the input of the channel to the output state $\mathcal{E}(\rho)$ appearing on the other side. It is a reasonable assumption that the channel belongs to a channel class, i.e., for a fixed input state the output state belongs to the parametric family $\{\mathcal{E}_\theta(\rho)\}_{\theta\in\Theta}$, where $\Theta$ denotes the parameter space. Thus, the channel estimation problem can be traced back to parameter estimation problem [10]. The channel parameter estimation problem is also called process tomography in the literature.

Direct quantum process tomography is performed by sending known quantum systems into the channel, and then estimating the output state. In quantum mechanics the measurement has a probabilistic nature [13, 15], therefore many identical copies of the input quantum system are needed, and an estimator is constructed by using statistical considerations. For achieving efficient process tomography, experiment design is necessary that consists of selecting the optimal input state, optimal measurement of the output state, and an efficient estimator of the channel from the measured data.

The field of quantum process tomography is well-established, an exhaustive description of possible tomography methods can be found in [11]. The Pauli channels - the subject of this thesis - form a relatively wide family of quantum channels. The tomography of Pauli channels has a huge literature, however, due to the level of difficulty of the topic, papers mostly deal with special cases, e.g., with the optimal parameter estimation of a depolarizing channel [17]. But there are some publications investigating the estimation of multi-parameter channels [2, 20], and the multidimensional case also appears [8, 12]. There are also some experimental results concerning the optimal estimation of the Pauli-channels [4, 5].

In contrast to the majority of the works in this area, we propose an extended problem statement: we investigate qubit Pauli channels with unknown channel directions. Despite of the novelty of the approach, there are a few papers that deal with optimally estimating qubit Pauli channels including their channel directions. In [1] the problem was examined using convex optimization methods, and a numerical method was provided for finding the optimal input - measurement pairs. In [16] we examined the optimality of the estimation problem using purely statistical considerations to achieve analytical results. However, analytical results could only be obtained for the case of known channel directions.

Therefore, the aim of this thesis is to give an analytical description of the optimal estimation of qubit Pauli channels in the case of unknown channel directions, too.

We introduce the *channel matrix* that characterizes the Pauli channel, and give an estimation method of the channel matrix. This estimation scheme uses three input qubits and three von Neumann measurements for complete channel tomography. Computing the Jordan-decomposition of the estimated channel matrix one gets the estimations of the contraction parameteres and the angle parameters that describe the channel directions. The efficiency of these estimations is measured here with three quantities: the mean squared error of the estimated contraction parameters and angle parameters, and the mean distance of the estimated and the real channel matrix are investigated.

The optimization of the different quantities needs different mathematical techniques. The mean squared error of the estimated contraction parameters and the mean distance of the estimated and the real channel matrix is minimized analytically. The mean squared error of the estimated angle parameters has a comlex form, therefore it is optimized with numerical methods in general. However, in an important special case this loss function can also be minimized analytically.

The main result of this work is that we determined the optimal measurement configurations for quantum bit Pauli channels with respect to the most relevant loss functions. For optimally estimating the contraction parameters and the channel matrix we should have input qubits and measurements in the channel directions, however, for optimally estimating the channel directions, we should use different tomography conditions: from simulation investigations and analytical results in special cases we conjecture that using input and measurement directions that are complementary to the channel directions would give a nearly optimal result.

# Chapter 2

# Basic notions

In this chapter we give a brief description of the mathematical formalism of quantum information theory and we get the hang of the *quantum channel* which is the object of the tomography problem presented in this thesis. This chapter is based on the books [13] and [15].

## 2.1 An owerview of the basics of quantum information theory

### 2.1.1 Quantum states

Every quantum system has an associated complex, separable Hilbert space. A quantum system is said to be *finite dimensional* (or simply finite) if the associated Hilbert space is $\mathbb{C}^n$ with some $n \in \mathbb{Z}^+$.

The elements of the standard basis of $\mathbb{C}^n$ are often denoted by $|0\rangle, \dots, |n-1\rangle$, and the corresponding elements of the dual space are denoted by $\langle 0|, \dots, \langle n-1|$.

The states of finite quantum systems are described by *density matrices.* $\rho \in \mathbf{M}_n(\mathbb{C})$ is a density matrix, if it satisfies the following conditions:

$$\mathrm{Tr}(\rho) = 1 \tag{2.1.1}$$

$$\rho \geq 0. \tag{2.1.2}$$

The state space of the $n$-level quantum system is denoted by $\mathcal{S}(\mathbb{C}^n)$. The set of the self-adjoint $n \times n$ matrices (denoted by $\mathbf{M}_n^{s.a.}(\mathbb{C})$) is a *real* vector space of dimension $n^2$, and it is also a Hilbert space with the inner product $\langle x, y\rangle := \mathrm{Tr} x^* y$.

It is a natural problem to characterize the quantum states. The most popular solution of this problem is the *Bloch paramertrization.* Let us fix an orthogonal basis of $\mathbf{M}_n^{s.a.}(\mathbb{C})$ containing the identity matrix: $\mathcal{F} = \{F_0 = I, F_1, F_2, \dots, F_{n^2-1}\}$. One can expand the $\rho \in \mathcal{S}(\mathbb{C}^n)$ density matrix in this basis:

$$\rho = \frac{1}{n}\left(x_0 I + x_1 F_1 + \cdots + x_{n^2-1} F_{n^2-1}\right) \ (x_0, \dots, x_{n^2-1} \in \mathbb{R}). \tag{2.1.3}$$

It follows from the condition (2.1.1) that $x_0 = 1$. The $(x_1, \dots, x_{n^2-1}) \in \mathbb{R}^{n^2-1}$ vector is called the Bloch vector of the state $\rho$.

In this thesis the most important quantum system is the quantum bit (or *qubit*). The Hilbert space of the qubit system is $\mathcal{H} = \mathbb{C}^2$, hence the states are described by $2 \times 2$ density matrices. The space of the $2 \times 2$ self-adjoint matrices is spanned by the *Pauli basis*. If $C \in \mathbf{M}_2^{s.a.}(\mathbb{C})$, then

$$C = \frac{1}{2}\left(x_0\sigma_0 + x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3\right), \tag{2.1.4}$$

where

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

are the Pauli matrices and $x_0, x_1, x_2, x_3 \in \mathbb{R}$. It is easy to check that $C$ is a density matrix if and only if $x_0 = 1$ and $x_1^2 + x_2^2 + x_3^2 \leq 1$.

Therefore by the Bloch parametrization, we can identify the state space with the closed unit ball of the space $\mathbb{R}^3$.

### 2.1.2 Quantum measurements

In quantum mechanics, the measurement has probabilistic nature. The measurable quantities, the *observables* are self-adjoint matrices, hence they have a spectral decomplosition. If $B \in \mathbf{M}_n^{s.a.}(\mathbb{C})$ then

$$B = \sum_j \lambda_j P_j, \tag{2.1.5}$$

where the $\lambda_j$-s are the different eigenvalues of $B$ and the $P_j$-s are the orthoprojections onto the eigenspace of the eigenvalue $\lambda_j$. The possible outcomes of the measurement are the eigenvalues of $B$ : if the quantum system is in the state $\rho$, we measure the eigenvalue $\lambda_j$ with probability $\mathrm{Tr}\rho P_j$.

The quantum measurement changes the state of the system. After measuring $\lambda_j$, the state of the system can be described with the density matrix

$$\rho_j = \frac{P_j \rho P_j}{\mathrm{Tr} P_j \rho P_j}.$$

Let $\{P_j\}_{j \in J}$ be the spectral projections of a self-adjoint matrix. Then naturally

$$(\forall j \in J) \ \ P_j \geq 0 \text{ and } \sum_{j \in J} P_j = I. \tag{2.1.6}$$

If a set of bounded operators $\{P_j\}_{j \in J}$ satisfies the conditions (2.1.6), then $\{P_j\}_{j \in J}$ is called a *positive operator valued measure* (the usual abbreviation is POVM).

The POVM describes a quantum measurement the following way: the possible outcomes are the elements of the POVM. If $M$ is an element of the POVM then the probability of measuring $M$ is $\mathrm{Tr}\rho M$, where $\rho$ is the state of the quantum system. The two-element POVMs form sn important special class of the POVMs. Every two-element POVM has the form $\{M, I - M\}$, where $0 \leq M \leq I$. If $M$ is a projection then the POVM $\{M, I - M\}$ is called a *von Neumann measurement.*

## 2.2 State transformations and quantum channels

### 2.2.1 State transformations

When our quantum system is not closed, the interaction with the environment causes dynamical changes. Assume that the Hilbert space of our quantum system in $\mathbb{C}^n$ and the initial state is $\rho \in \mathbf{M}_n(\mathbb{C})$. Let the Hilbert space of the environment be $\mathbb{C}^m$ and assume that the environment is in the state $\rho_e$. Before interaction the joint system has the statistical operator $\rho_e \otimes \rho \in \mathbf{M}_m\left(\mathbf{M}_n(\mathbb{C})\right)$. The effect of the interaction is implemented by unitary conjugation, the new density of the total system is $U\left(\rho_e \otimes \rho\right) U^*$ where $U \in \mathbf{M}_{m \cdot n}$ is a unitary matrix.

The new statistical operator of the quantum system we are interested in is the reduced density

$$\tilde{\rho} = \mathrm{Tr}_1\big(U\left(\rho_e \otimes \rho\right) U^*\big) \tag{2.2.1}$$

where $\mathrm{Tr}_1$ is the partial trace with respect to the first component. The map $\mathcal{E} : \rho \mapsto \mathcal{E}(\rho) := \tilde{\rho}$ is called *state transformation.*

Fortunately, there is a convenient characterization of the state transformations.

**Theorem.** *Any state transformation $\mathcal{E}: \ \mathbf{M}_n(\mathbb{C}) \to \mathbf{M}_n(\mathbb{C}); \ \rho \mapsto \mathcal{E}(\rho)$ has the form*

$$\mathcal{E}(\rho) = \sum_{k=1}^{m} E_k \rho E_k^* \tag{2.2.2}$$

*where the operators $E_k \in \mathbf{M}_n(\mathbb{C})$ satisfy*

$$\sum_{k=1}^{m} E_k^* E_k = I. \tag{2.2.3}$$

*Conversely, all linear mappings of this form are state transformations.*

The state transformations are defined on densities but they can be extended linearly to the matrix algebra $\mathbf{M}_n(\mathbb{C})$. Note that every state transformation is identity-preserving.

### 2.2.2 Examples of quantum channels

In this subsection we present some concrete examples of quantum operations.

**Example 1** (Depolarizing channel)**.** *The* depolarizing channel *describes a kind of quantum noise. Imagine that we have the statistical operator $\rho \in \mathbf{M}_n(\mathbb{C})$ and it is replaced by the compeletely mixed state $\frac{1}{n}I$ with probability $p$ and it is left untouched with probability $1-p$. The state of the quantum system after this noise is*

$$\mathcal{E}(\rho) = p\frac{1}{n}I + (1-p)\rho. \tag{2.2.4}$$

**Example 2** (Bit flip channel)**.** *The* bit flip channel *is the quantum analogue of the classical bit flip that changes the value of a bit with some probability. The bit flip channel acts on the state space of the quantum bit system. This channel is given by the equation*

$$\mathcal{E}(\rho) = E_1 \rho E_1^* + E_2 \rho E_2^* \tag{2.2.5}$$

*where*

$$E_1 = \sqrt{p}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \; E_2 = \sqrt{1-p}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

**Example 3** (Phase flip channel)**.** *The mathematical description of the* phase flip channel *is very similar to the previous example; the phase flip is given by the equation (2.2.5) with*

$$E_1 = \sqrt{p}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \; E_2 = \sqrt{1-p}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

All these state transformations belong to the family of *Pauli channels.*

# Chapter 3

# Tomography of qubit Pauli channels

The Pauli channels form a wide and well-known family of the quantum channels in the qubit case. Heuristically, the Pauli channel is a transformation of the state space that contracts in some directions. The main part of this work is about the tomography of the qubit Pauli channel, hence it seems to be useful to define it.

**Definition 1** (Qubit Pauli channel)**.** *Let* $\{\frac{1}{\sqrt{2}}I, v_1, v_2, v_3\}$ *be an arbitrary orthonormal basis of* $\mathbf{M}_2^{s.a.}(\mathbb{C})$*. Let* $\lambda_1, \lambda_2, \lambda_3$ *be real numbers satisfying the condition*

$$1 \pm \lambda_3 \geq |\lambda_1 \pm \lambda_2|. \tag{3.0.1}$$

*The mapping*

$$\mathcal{E} : \mathcal{S}(\mathbb{C}^2) \to \mathbf{M}_2^{s.a.}(\mathbb{C});\ \rho = \frac{1}{2}\left(I + \sum_{i=1}^{3} \theta_i v_i\right) \mapsto \mathcal{E}(\rho) = \frac{1}{2}\left(I + \sum_{i=1}^{3} \lambda_i \theta_i v_i\right) \tag{3.0.2}$$

*is called* qubit Pauli channel*. The* $\{\frac{1}{2}(I + tv_i) : t \in \mathbb{R}\} \subset \mathbf{M}_2^{s.a.}(\mathbb{C})$ *affine subspaces* ($i \in \{1, 2, 3\}$) *are the* channel directions*, the* $\lambda_1, \lambda_2, \lambda_3$ *numbers are the* contraction parameters.

The above defined Pauli channel is obviously trace-preserving, and the complete positivity is guaranteed via condition (3.0.1) [14]. Hence $\text{Range}(\mathcal{E}) \subset \mathcal{S}(\mathbb{C}^2) \subset \mathbf{M}_2^{s.a.}(\mathbb{C})$. Note that the condition (3.0.1) is symmetric in its three variables, and the inequality holds if and only if $(\lambda_1, \lambda_2, \lambda_3) \in \mathbf{Conv}\left((1, 1, 1), (1, -1, -1), (-1, 1, -1), (-1, -1, 1)\right) \subset \mathbb{R}^3$, where **Conv** denotes the convex hull.

The effect of the Pauli channel gets more picturesque, if we consider the representing Bloch vectors of the quantum states. In the qubit case, the Bloch parametrization is the

$$\rho : \overline{\mathbf{B}}^3 \to \mathcal{S}(\mathbb{C}^2);\ \theta = (\theta_1, \theta_2, \theta_3) \mapsto \rho(\theta) = \frac{1}{2}\left(I + \theta_1\sigma_1 + \theta_2\sigma_2 + \theta_3\sigma_3\right) \tag{3.0.3}$$

map, where $\overline{\mathbf{B}}^3 \subset \mathbb{R}^3$ is the closed unit ball (that used to be called Bloch ball in this context). In the following we introduce the *channel matrix*, that decribes the effect of the Pauli channel in the Bloch ball modell of the state space, and we characterize the channel matrices.

**Definition 2** (Channel matrix)**.** *The* $A : \overline{\mathbf{B}}^3 \to \overline{\mathbf{B}}^3$ *map is the* channel matrix *of the Pauli channel* $\mathcal{E}$*, if*

$$\mathcal{E} \circ \rho = \rho \circ A. \tag{3.0.4}$$

**Lemma 1.** *Every Pauli channel has a unique channel matrix. The channel matrices are linear transformations of* $\mathbb{R}^3$*, and every channel matrix is diagonal in an orthonormal basis of* $\mathbb{R}^3$*.*

*Proof.* The parametrization $\rho$ is a bijection, therefore the equation (3.0.4) can be written in the form

$$\rho^{-1} \circ \mathcal{E} \circ \rho = A. \tag{3.0.5}$$

This form shows the existence and the uniqueness of $A$.

$\{v_j\}_{j=1}^3$ is an orthonormal basis of the traceless subspace of $\mathbf{M}_2^{s.a.}(\mathbb{C})$, hence with the notations

$$\underline{\sigma} = \frac{1}{\sqrt{2}} \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix}, \ \underline{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \tag{3.0.6}$$

one gets

$$\underline{\sigma} = R\underline{v} \tag{3.0.7}$$

with some $R \in \mathbf{O}(3, \mathbb{R})$ matrix

Using the notation $r_{ij} = [R]_{ij}$ one has

$$\mathcal{E}(\rho(\theta)) = \mathcal{E}\left(\frac{1}{2}\left(I + \sum_{i=1}^{3} \theta_i \sigma_i\right)\right) = \mathcal{E}\left(\frac{1}{2}\left(I + \sum_{i=1}^{3} \theta_i \sum_{j=1}^{3} r_{ij} v_j\right)\right) =$$

$$= \mathcal{E}\left(\frac{1}{2}\left(I + \sum_{j=1}^{3}\left(\sum_{i=1}^{3} \theta_i r_{ij}\right) v_j\right)\right) = \frac{1}{2}\left(I + \sum_{j=1}^{3} \lambda_j \left(\sum_{i=1}^{3} \theta_i r_{ij}\right) v_j\right) =$$

$$= \frac{1}{2}\left(I + \sum_{j=1}^{3} \lambda_j \left(\sum_{i=1}^{3} \theta_i r_{ij}\right)\left(\sum_{k=1}^{3} r_{kj} \sigma_k\right)\right) = \frac{1}{2}\left(I + \sum_{k=1}^{3}\left(\sum_{i=1}^{3}\left(\sum_{j=1}^{3} \lambda_j r_{ij} r_{kj}\right)\theta_i\right)\sigma_k\right). \tag{3.0.8}$$

Therefore

$$\left((\rho^{-1} \circ \mathcal{E} \circ \rho)(\theta)\right)_k = \sum_{i=1}^{3}\left(\sum_{j=1}^{3} \lambda_j r_{ij} r_{kj}\right)\theta_i. \tag{3.0.9}$$

(3.0.9) shows that $A$ is a linear map, and its matrix elements are obtained as well:

$$[A]_{ki} = \sum_{j=1}^{3} \lambda_j r_{ij} r_{kj} = \sum_{j=1}^{3} [R]_{kj} \lambda_j [R^T]_{ji}. \tag{3.0.10}$$

Hence

$$A = R\Lambda R^T, \tag{3.0.11}$$

where

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

$\square$

Set $\mathbf{E} = \{\mathrm{Diag}\,(\varepsilon_1, \varepsilon_2, \varepsilon_3) : \ \varepsilon_i \in \{1, -1\} \ \forall i \in \{1, 2, 3\}\}$. $\mathbf{E}$ is a subgroup of $\mathbf{O}(3, \mathbb{R})$ that is isomorphic to $(\mathbb{Z}_2^3, +)$. It is easy to see that for $R_1, R_2 \in \mathbf{O}(3, \mathbb{R})$

$$R_1 \Lambda(\lambda_1, \lambda_2, \lambda_3) R_1^T \equiv R_2 \Lambda(\lambda_1, \lambda_2, \lambda_3) R_2^T \tag{3.0.12}$$

if and only if $R_2 = R_1 E$ for some $E \in \mathbf{E}$. Therefore, two orthogonal matrices determine the same channel parametrization if and only if they are in the same *left coset* of $\mathbf{E}$.

Hence, using a parametrization of $\mathbf{O}(3, \mathbb{R})/\mathbf{E}$ and the result of Lemma 1, one can parametrize the channel matrices the following way.

**Lemma 2.** *Every channel matrix $A$ has the form*

$$A(\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x) = R_z R_y R_x \Lambda R_x^{-1} R_y^{-1} R_z^{-1}, \tag{3.0.13}$$

*where*

$$R_z(\phi_z) = \begin{pmatrix} \cos\phi_z & -\sin\phi_z & 0 \\ \sin\phi_z & \cos\phi_z & 0 \\ 0 & 0 & 1 \end{pmatrix}, \ R_y(\phi_y) = \begin{pmatrix} \cos\phi_y & 0 & -\sin\phi_y \\ 0 & 1 & 0 \\ \sin\phi_z & 0 & \cos\phi_y \end{pmatrix},$$

$$R_x(\phi_x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi_x & -\sin\phi_x \\ 0 & \sin\phi_x & \cos\phi_x \end{pmatrix}, \ \Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix},$$

$0 \le \phi_z, \phi_y, \phi_x < \pi$, *and* $\phi_y = \frac{\pi}{2} \Rightarrow \phi_z = 0$, *and the real numbers* $\lambda_1 \ge \lambda_2 \ge \lambda_3$ *satisfy the condition of positivity (3.0.1).*

Note that this parametrization is surjective but not bijective. If some contraction parameters are equal, the channel matrix gets independent of some angle parameters. We will determine a parameter domain where the parametrization of the channel matrices is bijective. This domain is described in (3.1.18).

## 3.1 A tomography scheme for qubit Pauli channels

### 3.1.1 The description of the tomography scheme

For complete channel tomography, we need three input states, and three differnt measurements. Earlier investigations show that the input qubits should be *pure* states, and their Bloch vectors should be orthogonal [1, 16].

Let $\underline{\theta}_1, \underline{\theta}_2, \underline{\theta}_3$ be the Bloch vectors of the input states. The matrix $[\underline{\theta}_1, \underline{\theta}_2, \underline{\theta}_3]$ is orthogonal, hence it can be written in the form

$$[\underline{\theta}_1, \underline{\theta}_2, \underline{\theta}_3] = R_z(\vartheta_z) R_y(\vartheta_y) R_x(\vartheta_x), \tag{3.1.1}$$

where $R_z, R_y, R_x$ are the rotations defined in (3.0.13), and $0 \le \vartheta_z, \vartheta_y < \pi$, $0 \le \vartheta_x < \frac{\pi}{2}$. Therefore, the chosen input states are described by the $\vartheta_z, \vartheta_y, \vartheta_x$ angle parameters.

Now, we show that the von Neumann measurements can be parametrized in a similar way.

**Lemma 3.** *In the qubit case, the two-elemwnt POVM $\{M, I - M\}$ is a non-trivial von Neumann measurement if and only if $M$ is a rank-one projection. A density matrix is a rank-one projection if and only if its Bloch vector's length equals one.*

*Proof.* The first part of the statement is obvious. On the other hand, $M$ is a rank-one projection $\Leftrightarrow M = M^*$ and $M$ has the eigenvalues $\mu_1 = 1,\ \mu_2 = 0 \Leftrightarrow M = M^*$, $\mathrm{Tr}M = 1$, $\mathrm{Det}M = 0$. It is easy to see that the last three conditions are satisfied by the densities lying on the border of the Bloch ball. $\square$

Let $\{\{M_i, I - M_i\}\}_{i=1}^3$ denote the von Neumann measurements, and let the corresponding Bloch vectors be denoted by $\{\underline{m}_i\}_{i=1}^3$. Earlier investigations show that the measurement directions (i.e. the Bloch vectors of the measurements) should be orthogonal as well [16]. Therefore the measurement directions $\underline{m}_1, \underline{m}_2, \underline{m}_3$ can be written in the form

$$[\underline{m}_1, \underline{m}_2, \underline{m}_3] = R_z(\tau_z)R_y(\tau_y)R_x(\tau_x) \tag{3.1.2}$$

where $0 \le \tau_z, \tau_y < \pi$, $0 \le \tau_x < \frac{\pi}{2}$.

### 3.1.2 The estimation of the output states and the channel matrix

To get information about the Pauli channel, one has to estimate the output qubits. For every $j \in \{1, 2, 3\}$, we send $3N$ identical copies of the input state $\rho\left(\underline{\theta}_j\right)$ into the Pauli channel. Ater that, we perform the von Neumann measurement $\{M_i, I - M_i\}$ $N$ times in the output state $\mathcal{E}\left(\rho\left(\underline{\theta}_j\right)\right)$ $(i \in \{1, 2, 3\})$. $N$ is an important parameter of the tomography scheme, it is called the *number of measurements*. The following lemma declares a basic property of the von Neumann measurements.

**Lemma 4.** *If one performes the $\{M, I - M\}$ von Neumann measurement in the state $\rho(\underline{\theta})$, then*

$$\mathrm{Prob}(\textit{measuring } M) = \frac{1}{2}\left(1 + \underline{m} \cdot \underline{\theta}\right), \tag{3.1.3}$$

*where $\underline{m} = (m_1, m_2, m_3)$ is the Bloch vector of $M$.*

*Proof.*

$$\begin{aligned} \mathrm{Prob}(\text{measuring M}) &= \mathrm{Tr}(\rho(\theta)M) = \\ = \mathrm{Tr}\left(\frac{1}{2}\left(I + \theta_1\sigma_1 + \theta_2\sigma_2 + \theta_3\sigma_3\right)\frac{1}{2}\left(I + m_1\sigma_1 + m_2\sigma_2 + m_3\sigma_3\right)\right) &= \\ = \frac{1}{4}2\left(1 + \underline{m} \cdot \underline{\theta}\right) &= \frac{1}{2}\left(1 + \underline{m} \cdot \underline{\theta}\right), \end{aligned} \tag{3.1.4}$$

because the Pauli matrices satisfy $\mathrm{Tr}(\sigma_i\sigma_j) = 2\delta_{ij}$ $(i, j \in \{0, 1, 2, 3\})$. $\square$

Let $N_{ij}^+$ denote the number of the events when $M_i$ is measured in the state $\mathcal{E}\left(\rho\left(\underline{\theta}_j\right)\right)$, and $\underline{\xi}_j$ be the Bloch vector of $\mathcal{E}\left(\rho\left(\underline{\theta}_j\right)\right)$. Set

$$x_{ij} = \underline{m}_i \cdot \underline{\xi}_j,\ X = \{x_{ij}\}_{i,j=1}^3. \tag{3.1.5}$$

It follows from the previous lemma and from the independence of the different measurements that

$$N_{ij}^+ = \mathrm{Binom}\left(N, \frac{1 + x_{ij}}{2}\right). \tag{3.1.6}$$

The well-known properties of the binomial distribution show that

$$\mathbf{E}\left(N_{ij}^{+}\right) = \frac{N}{2}(1+x_{ij}),\ \mathbf{Var}\left(N_{ij}^{+}\right) = N\frac{1+x_{ij}}{2}\left(1-\frac{1+x_{ij}}{2}\right) = \frac{N}{4}(1-x_{ij}^2). \tag{3.1.7}$$

Now, we can state that

**Lemma 5.** *There exists an unbiased estimator of $x_{ij}$ and its variance is $\underline{\underline{O}}\left(\frac{1}{N}\right)$.*

*Proof.*

$$\hat{x}_{ij} := \frac{2}{N}N_{ij}^{+} - 1. \tag{3.1.8}$$

Then

$$\mathbf{E}\left(\hat{x}_{ij}\right) = \frac{2}{N}\frac{N}{2}(1+x_{ij}) - 1 = x_{ij},\ \mathbf{Var}\left(\hat{x}_{ij}\right) = \frac{4}{N^2}\frac{N}{4}(1-x_{ij}^2) = \frac{(1-x_{ij}^2)}{N}. \tag{3.1.9}$$

□

Let us introduce the notations

$$M = [\underline{m}_1, \underline{m}_2, \underline{m}_3],\ \Theta = [\underline{\theta}_1, \underline{\theta}_2, \underline{\theta}_3],\ \Xi = \left[\underline{\xi}_1, \underline{\xi}_2, \underline{\xi}_3\right],\ \hat{X} = \{\hat{x}_{ij}\}_{i,j=1}^{3}. \tag{3.1.10}$$

By definition,

$$X = M^T\Xi, \tag{3.1.11}$$

hence $MX = \Xi$, therefore we can estimate the Bloch vector of the output qubits:

$$\hat{\Xi} := M\hat{X}. \tag{3.1.12}$$

By the definition of the channel matrix $A$, the equation

$$\Xi = A\Theta \tag{3.1.13}$$

holds, hence let the estimator of the channel matrix be

$$\hat{A} := \hat{\Xi}\Theta^{-1} = M\hat{X}\Theta^{-1}. \tag{3.1.14}$$

### 3.1.3 The estimation of the channel parameters

To get the estimators of the channel parameters, one has to compute the inverse of the channel parametrization described in Lemma 2, and apply it to the estimated channel matrix $\hat{A}$.

Hence, we have to construct an extension of the inverse of the mapping

$$A : \mathcal{D} \to \mathbf{M}_3(\mathbb{R});\ (\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x) \mapsto A(\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x) \tag{3.1.15}$$

that is defined in (3.0.13). Here $\mathcal{D} \subset \mathbb{R}^6$ is a proper subset of the parameter domain determined in Lemma 2, where $A$ is injective. We have to find a map

$$T : \mathbf{M}_3(\mathbb{R}) \to \mathbb{R}^6;\ \hat{A} \mapsto (\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\phi}_z, \hat{\phi}_y, \hat{\phi}_x) \tag{3.1.16}$$

such that

$$T \circ A = \mathrm{Id}_{\mathcal{D}}. \tag{3.1.17}$$

The domain of $T$ should be $\mathbf{M}_3(\mathbb{R})$, because the estimation $\hat{A}$ is a random variable, its range is wider than Range$(A)$.

The construction of $T$ is the following. Let us symmetrize the estimation of the channel matrix $\hat{A}_s := \frac{1}{2}\left(\hat{A} + \hat{A}^T\right)$. $\hat{A}_s$ is symmetric, hence it is diagonizable in an orthogonal basis. The Jordan decomposition of $\hat{A}_s$ is computable, because its characteristic polynomial is qubic, hence the eigenvalues can be obtained by Cardano's formula. By simple linear algebraic methods, the eigenvectors are obtained as well. Therefore, let us assume that $\hat{A}_s$ has the eigenvalues $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \hat{\lambda}_3$ and the corresponding eigenvectors are $\mathbf{v_1} = (v_1^1, v_1^2, v_1^3), \mathbf{v_2} = (v_2^1, v_2^2, v_2^3), \mathbf{v_3} = (v_3^1, v_3^2, v_3^3)$. Without loss of generality we can assume that $v_1^3 > 0$ or $v_1^3 = 0$ & $v_1^2 > 0$ or $v_1^1 = 1$, because if the above condition is not satisfied, we may consider $-\mathbf{v_1}$.

The estimators of the contraction parameters are the $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \hat{\lambda}_3$ eigenvalues of $\hat{A}_s$. The hard task is the estimation of the angle parameters. By the investigation of the channel parametrization (3.0.13), one can see the picturesque meaning of the angle parameters:

- $\phi_z$ is the polar angle of the eigenvector $\mathbf{v_1}$
- $\phi_y$ is the azimuth angle of the eigenvector $\mathbf{v_1}$
- $\phi_z$ is the angle between the eigenvector $\mathbf{v_2}$ and the intersection of the orthocomplement subspace of $\mathbf{v_1}$ and the subspace spanned by the standard unit vectors $\mathbf{e_1}$ and $\mathbf{e_2}$.

These quantities are well-defined if and only if $\hat{\lambda}_1 > \hat{\lambda}_2 > \hat{\lambda}_3$. But we can not assume in general that $\hat{\lambda}_1 > \hat{\lambda}_2 > \hat{\lambda}_3$, hence we have to handle the following four opportunities.

1. In the general case $\hat{\lambda}_1 > \hat{\lambda}_2 > \hat{\lambda}_3$.

   In the case $v_1^3 = 1$ set $\hat{\phi}_y = \frac{\pi}{2}$, $\hat{\phi}_z = 0$. (If the azimuth angle equals $\frac{\pi}{2}$, the polar angle can not be determined, hence our choice is 0.)

   If $v_1^3 \neq 1$ and $v_1^2 = 0$, set $\hat{\phi}_y = \arccos v_1^1$ and $\hat{\phi}_z = 0$, because if $\mathbf{v_1}$ lies in the plane spanned by $\mathbf{e_1}$ and $\mathbf{e_3}$, but $\mathbf{v_1} \neq \mathbf{e_3}$, then the polar angle is zero, and the azimuth angle can be calculated by the above simple formula.

   If $v_1^2 \neq 0$, set $y = \operatorname{sgn}(v_1^2)\sqrt{(v_1^1)^2 + (v_1^2)^2}$, $z := \frac{v_1^1}{y}$ and $\hat{\phi}_y := \arccos y$, $\hat{\phi}_z := \arccos z$.

   We determined $\hat{\phi}_z$ and $\hat{\phi}_y$, hence we can write an orthonormal basis of the orthogonal subspace of $\mathbf{v_1}$ with the folowing property: the first basis vector is in the plane spanned by $\mathbf{e_1}$ and $\mathbf{e_2}$:

   $$\mathbf{s_1} = (-\sin\phi_z, \cos\phi_z, 0)^T, \ \ \mathbf{s_2} = (-\sin\phi_y\cos\phi_z, -\sin\phi_y\sin\phi_z, \cos\phi_y)^T.$$

   These two vectors are the second and the third column vectors of the matrix $R_z(\phi_z)R_y(\phi_y)$. Therefore $\mathbf{v_1}, \mathbf{s_1}, \mathbf{s_2}$ is an orthonormal basis. $\mathbf{v_2}$ and $\mathbf{v_1}$ are orthogonal, hence $\mathbf{v_2} = \langle \mathbf{v_2}, \mathbf{s_1}\rangle\, \mathbf{s_1} + \langle \mathbf{v_2}, \mathbf{s_2}\rangle\, \mathbf{s_2}$. Let us introduce the notations $q_1 = \langle \mathbf{v_2}, \mathbf{s_1}\rangle$ and $q_2 = \langle \mathbf{v_2}, \mathbf{s_2}\rangle$ . Now, we can estimate the third angle: set $x = \operatorname{sgn}(q_2)q_1$ if $|q_1| \neq 1$ and $x = 1$ if $|q_1| = 1$, and $\hat{\phi}_x := \arccos x$.

2. In the case $\hat{\lambda}_1 > \hat{\lambda}_2 = \hat{\lambda}_3$ the estimations $\hat{\phi}_y$ and $\phi_z$ can be determined as in the generel case, and we choose $\hat{\phi}_x = 0$. (If $\lambda_2 = \lambda_3$, then $R_x \Lambda R_x^{-1} \equiv \Lambda$.)

3. In the case $\hat{\lambda}_1 = \hat{\lambda}_2 > \hat{\lambda}_3$ one can calculate the eigenvector $\mathbf{v_3}$.

   If $|v_3^3| = 1$, then $\mathbf{v_1} \in \text{span}(\mathbf{e_1}, \mathbf{e_2})$, hence the azimuth angle is zero ($\hat{\phi}_y := 0$). $\text{span}(\mathbf{e_1}, \mathbf{e_2})$ is the eigensubspace of the eigenvalue $\hat{\lambda}_1$, hence the polar angle is unspecified, our choice is $\hat{\phi}_z = 0$. Obviously, $\hat{\phi}_x = 0$ in this case.

   If $|v_3^3| \neq 1$, then there is a unique eigenvector $\mathbf{v_2}$ of the eigenvalue $\hat{\lambda}_1 = \hat{\lambda}_2$ such that $v_2^3 = 0$, $v_2^1 \leq 0$ and $v_2^1 = 0 \Rightarrow v_2^2 = 1$. The polar angle can be calculated from $\mathbf{v_2}$ : $\hat{\phi}_z = \arccos v_2^2$. The azimuth angle is the angle between $\mathbf{v_3}$ and $\mathbf{e_3}$: $\hat{\phi}_y = \arccos v_3^3$, and $\hat{\phi}_x = 0$.

4. If $\lambda_1 = \lambda_2 = \lambda_3$, then set $\hat{\phi}_x = \hat{\phi}_y = \hat{\phi}_z = 0$.

Now we can determine the parameter domain, where the parametrization of the Pauli channels is bijective:

$$\begin{aligned}
\mathcal{D} = \{(\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x) \in \mathbb{R}^6 :\ 1 \pm \lambda_3 \geq |\lambda_1 \pm \lambda_2|,\ \lambda_1 \geq \lambda_2 \geq \lambda_3, \\
\phi_z, \phi_y, \phi_x \in [0, \pi), \phi_y = \frac{\pi}{2} \Rightarrow \phi_z = 0,\ \lambda_1 = \lambda_2 = \lambda_3 \Rightarrow \phi_z = \phi_y = \phi_x = 0, \\
\lambda_1 = \lambda_2 > \lambda_3 \Rightarrow (\phi_x = 0 \text{ és } \phi_y = 0 \Rightarrow \phi_z = 0), \\
\lambda_1 > \lambda_2 = \lambda_3 \Rightarrow \phi_x = 0\}.
\end{aligned} \tag{3.1.18}$$

## 3.2 The accuracy of the parameter estimation

The efficiency of the channel tomography scheme described in the previous subsection depends on the parameters of the Pauli channel, and it depends on the parameters of the estimation scheme, as well. The input states and the maesurements are described by the angle parameters $\underline{\vartheta} = (\vartheta_z, \vartheta_y, \vartheta_x)$ and $\underline{\tau} = (\tau_z, \tau_y, \tau_x)$ introduced in (3.1.1) and (3.1.2). Hence, the estimation scheme is determined by the parameters $\underline{\vartheta}$, $\underline{\tau}$ and $N$ ($N$ is the number of measurements). The Pauli channel is described by the parameters $\underline{\lambda} = (\lambda_1, \lambda_2, \lambda_3)$ and $\underline{\phi} = (\phi_z, \phi_y, \phi_x)$.

We measure the efficiency of the parameter estimation with the mean squared error of the estimated contraction parameters and angle parameters, and with the mean squared mean squared distance of the real and the estimated channel matrix.

Let us introduce the following quantities:

$$f_1(\underline{\lambda}, \underline{\phi}, \underline{\tau}, \underline{\vartheta}, N) = \mathbf{E}\left(\text{dist}(\hat{\phi}_z, \phi_z)^2 + \text{dist}(\hat{\phi}_y, \phi_y)^2 + \text{dist}(\hat{\phi}_x, \phi_x)^2\right) \tag{3.2.1}$$

$$f_2(\underline{\lambda}, \underline{\phi}, \underline{\tau}, \underline{\vartheta}, N) = \mathbf{E}\left((\hat{\lambda}_1 - \lambda_1)^2 + (\hat{\lambda}_2 - \lambda_2)^2 + (\hat{\lambda}_3 - \lambda_3)^2\right) \tag{3.2.2}$$

$$f_3(\underline{\lambda}, \underline{\phi}, \underline{\tau}, \underline{\vartheta}, N) = \mathbf{E}\left(||A(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\phi}_z, \hat{\phi}_y, \hat{\phi}_x) - A(\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x)||^2\right), \tag{3.2.3}$$

where $||\cdot||$ is the Hilbert-Schmidt norm. We identify the angle parameters that determine the same Pauli channel, therefore we define the above distance function with the formula

$$\text{dist}(\hat{\phi}_\alpha, \phi_\alpha) := \inf\{|\hat{\phi}_\alpha - (\phi_\alpha + k\pi)| :\ k \in \mathbb{Z}\}. \tag{3.2.4}$$

It is easy to see that in fact

$$\inf\{|\hat{\phi}_\alpha - (\phi_\alpha + k\pi)| :\ k \in \mathbb{Z}\} = \min\{|\hat{\phi}_\alpha - (\phi_\alpha + k\pi)| :\ k \in \{-1, 0, 1\}\}, \tag{3.2.5}$$

hence $\mathrm{dist}(\hat{\phi}_\alpha, \phi_\alpha)$ is easy to compute.

Our aim is to determine the most efficient estimation strategy for given Pauli channel and for fixed number of measurements. That is, we have to minimize the

$$(\underline{\tau}, \underline{\vartheta}) \mapsto f_i(\underline{\lambda}, \underline{\phi}, \underline{\tau}, \underline{\vartheta}, N) \ (i \in \{1,2,3\}) \tag{3.2.6}$$

functions with given $\underline{\lambda}, \underline{\phi}, N$ values.

**Proposition 6** (Rotational invariance)**.** *For any $O \in \mathbf{M}_3(\mathbb{R})$ orthogonal matrix, the estimation of the Pauli channel described by the channel matrix $A$ using the input and measurement settings $\Theta$ and $M$ (see (3.1.1), (3.1.2)) is exactly as efficient as the estimation of the Pauli channel described by $OAO^{-1}$ with the input and measurement settings $O\Theta$ and $OM$.*

Therefore, it is enough to investigate Pauli channels with channel parameters $\phi_z = \phi_y = \phi_x = 0$.

### 3.2.1 General computations

First we prove some basic statements that are useful to solve the above declared optimization problems.

Sometimes, it is convenient to expand matrices in the basis of the *matrix units,* and consider the representing vectors instead of the matrices themselves. The following lemma formulates a basic rule that can be proved by direct computations.

**Lemma 7.** *Set $B, J \in \mathbf{M}_n(\mathbb{R})$, then $L_B R_J : \mathbf{M}_n(\mathbb{R}) \to \mathbf{M}_n(\mathbb{R})$; $X \mapsto BXJ$ is a linear map, furthermore, the matrix of $L_B R_J$ is $B \otimes J^T \in \mathbf{M}_{n^2}(\mathbb{R})$. That is, if*

$$x_{11}, \dots, x_{1n}, x_{21}, \dots, x_{2n}, \dots, x_{n1}, \dots, x_{nn} \in \mathbb{R}$$

*are defined by the equation $X = \sum_{i,j=1}^n x_{ij} E_{ij}$ (where $\{E_{ij}\}_{i,j=1}^n$ is the basis of the matrix units), and*

$$y_{11}, \dots, y_{1n}, y_{21}, \dots, y_{2n}, \dots, y_{n1}, \dots, y_{nn} \in \mathbb{R}$$

*are given by the equation $BXJ = \sum_{i,j=1}^n y_{ij} E_{ij}$, then*

$$\begin{aligned} &(y_{11}, \dots, y_{1n}, y_{21}, \dots, y_{2n}, \dots, y_{n1}, \dots, y_{nn})^T = \\ &= B \otimes J^T (x_{11}, \dots, x_{1n}, x_{21}, \dots, x_{2n}, \dots, x_{n1}, \dots, x_{nn})^T. \end{aligned} \tag{3.2.7}$$

The following lemma is often cited in the further proofs.

**Lemma 8.** *The tensor product of orthogonal matrices is orthogonal as well, and the* Hadamart-square *of an orthogonal matrix is bistochastic.*

*Proof.* If $O_1, O_2 \in \mathbf{M}_n(\mathbb{R})$ are orthogonal, then

$$(O_1 \otimes O_2)(O_1 \otimes O_2)^T = (O_1 \otimes O_2)\left(O_1^T \otimes O_2^T\right) = O_1 O_1^T \otimes O_2 O_2^T = I_n \otimes I_n = I_{2n}. \tag{3.2.8}$$

A similar computation shows that

$$(O_1 \otimes O_2)^T (O_1 \otimes O_2) = I_{2n}. \tag{3.2.9}$$

Hence $O_1 \otimes O_2$ is orthogonal, therefore (among others) $(O_1 \otimes O_2)^{-1} = \left(O_1^T \otimes O_2^T\right)$.

The bistochastic property of the Hadamart-square follows from the fact that every column and row of an orthogonal matrix is a unit vector (in the Euclidean norm). $\square$

**Lemma 9** (Unbiasedness I)**.** *The estimation of the channel matrix is unbiased, that is* $\mathbf{E}(\hat{a}_{ij}) = a_{ij}$*, where* $\{\hat{a}_{ij}\}_{i,j=1}^{3}$ *are the elements of the estimation* $\hat{A}$ *defined in (3.1.14), and* $\{a_{ij}\}_{i,j=1}^{3}$ *are the elements of the channel matrix* $A(\underline{\lambda}, \underline{\phi})$.

*Proof.* (3.1.9) states that $\mathbf{E}(\hat{x}_{ij}) = x_{ij}$. Because of Lemma 7 we can write (3.1.14) in the following form:

$$\hat{A} = (R(\underline{\tau}) \otimes R(\underline{\vartheta}))\, \hat{X}. \tag{3.2.10}$$

(Now $\hat{A}$ and $\hat{X}$ denote the representing vectors.) The expected value is linear, hence

$$\mathbf{E}(\hat{A}) = (R(\underline{\tau}) \otimes R(\underline{\vartheta}))\, \mathbf{E}(\hat{X}) = (R(\tau) \otimes R(\underline{\vartheta}))\, X. \tag{3.2.11}$$

On the other hand, using the representing vectors, by the definition of $X$, we can write

$$X = \left(R(\underline{\tau})^T \otimes R(\underline{\vartheta})^T\right) A. \tag{3.2.12}$$

$\left(R(\underline{\tau})^T \otimes R(\underline{\vartheta})^T\right)^{-1} = R(\underline{\tau}) \otimes R(\underline{\vartheta})$, hence

$$A = (R(\underline{\tau}) \otimes R(\underline{\vartheta}))\, X. \tag{3.2.13}$$

(3.2.11) and (3.2.13) shows that the proof is complete. □

If $i \neq k$ or $j \neq l$, then $\hat{x}_{ij}$ and $\hat{x}_{kl}$ are independent random variables, because they are determined by independent measurements. Hence the equation (3.2.10) shows that the $\hat{a}_{ij}$ estimators are linear combinations of *independent* random variables, and the coefficients depend only on $\underline{\tau}$ and $\underline{\vartheta}$:

$$\hat{a}_k = \sum_{l \in H} c_{kl}\left(\underline{\tau}, \underline{\vartheta}\right) \hat{x}_l, \tag{3.2.14}$$

where $H = \{11, 12, 13, 21, 22, 23, 31, 32, 33\}$, $k \in H$.

**Lemma 10.** *Set* $\psi = \sum_{k \in H} d_k \hat{a}_k$ $(d_k \in \mathbb{R})$*. Then*

$$\mathbf{Var}\left(\psi\right) = \sum_{l \in H} \left( \sum_{k \in H} d_k c_{kl}\left(\tau, \vartheta\right) \right)^2 \frac{1 - x_l^2}{N}. \tag{3.2.15}$$

*Proof.* (3.2.14) shows that

$$\psi = \sum_{k \in H} d_k \hat{a}_k = \sum_{k \in H} d_k \sum_{l \in H} c_{kl}\left(\tau, \vartheta\right) \hat{x}_l = \sum_{l \in H} \left( \sum_{k \in H} d_k c_{kl}\left(\tau, \vartheta\right) \right) \hat{x}_l, \tag{3.2.16}$$

and $\mathbf{Var}(\hat{x}_{ij}) = \frac{1}{N}\left(1 - x_{ij}^2\right)$ (see (3.1.9)), hence the proof is complete. □

### 3.2.2 Linearized estimators

It is difficult to handle the functions $f_1$ and $f_2$ defined in (3.2.1) anf (3.2.2), because the estimators $\hat{\lambda}_i$ and $\hat{\phi}_\alpha$ are non-linear. These estimators are $\mathbf{M}_3(\mathbb{R}) \to \mathbb{R}$ functions (the input is the estimated channel matrix), one can approximate them with their first-order Taylor-polynomial. Let the base point of the Taylor-polynomial be the *expected value* of the input. By Lemma 10, $\mathbf{Var}\left(\hat{a}_{ij}\right) = \underline{\underline{O}}\left(\frac{1}{N}\right)$, hence the distribution of the random variable $\hat{A}$ is concentrated on a small environment of $\mathbf{E}(\hat{A})$, if $N$ is large enough.

Therefore, the linearized functions with the base point $\mathbf{E}(\hat{A})$ approximate the non-linear estimators well.

By Lemma 9, $\mathbf{E}(\hat{A}) = A$, hence we consider the following linearizations of the estimators $\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\phi}_z, \hat{\phi}_y, \hat{\phi}_x$ :

$$\tilde{\lambda}_i := \hat{\lambda}_i(A) + \left\langle \mathrm{grad}\hat{\lambda}_i(A),\, \hat{A} - A \right\rangle \ (\forall\ i \in \{1, 2, 3\}), \tag{3.2.17}$$

$$\tilde{\phi}_\alpha := \hat{\phi}_\alpha(A) + \left\langle \mathrm{grad}\hat{\phi}_\alpha(A),\, \hat{A} - A \right\rangle \ (\forall\ \alpha \in \{z, y, x\}). \tag{3.2.18}$$

Now, we can define an approximation of the functions $f_1$ and $f_2$.

$$\tilde{f}_1(\underline{\lambda}, \underline{\phi}, \underline{\tau}, \underline{\vartheta}, N) := \mathbf{E}\left((\tilde{\phi}_z - \phi_z)^2 + (\tilde{\phi}_y - \phi_y)^2 + (\tilde{\phi}_x - \phi_x)^2\right), \tag{3.2.19}$$

$$\tilde{f}_2(\underline{\lambda}, \underline{\phi}, \underline{\tau}, \underline{\vartheta}, N) := \mathbf{E}\left((\tilde{\lambda}_1 - \lambda_1)^2 + (\tilde{\lambda}_2 - \lambda_2)^2 + (\tilde{\lambda}_3 - \lambda_3)^2\right) \tag{3.2.20}$$

**Lemma 11** (Unbiasedness II)**.** *The linearized estimators $\tilde{\lambda}_i$ $(i \in \{1, 2, 3\})$ and $\tilde{\phi}_\alpha$ $(\alpha \in \{z, y, x\})$ are unbiased.*

*Proof.*

$$\mathbf{E}(\tilde{\lambda}_i) = \mathbf{E}(\hat{\lambda}_i(A)) + \mathbf{E}\left(\left\langle \mathrm{grad}\hat{\lambda}_i(A),\, \hat{A} - A \right\rangle\right) = \hat{\lambda}_i(A) + \left\langle \mathrm{grad}\hat{\lambda}_i(A),\, \mathbf{E}(\hat{A} - A) \right\rangle =$$

$$= \hat{\lambda}_i(A) + 0 = \lambda_i. \tag{3.2.21}$$

It was used that the estimation of the channel matrix is unbiased (see Lemma 9), and the parameter estimation is the (left) inverse of the channel parametrization, hence

$$\hat{\lambda}_i\left(A(\underline{\lambda}, \underline{\phi})\right) = \lambda_i\ (i \in \{1, 2, 3\}). \tag{3.2.22}$$

Similarly, $\hat{\phi}_\alpha\left(A(\underline{\lambda}, \underline{\phi})\right) = \phi_\alpha$, therefore

$$\mathbf{E}(\tilde{\phi}_\alpha) = \mathbf{E}(\hat{\phi}_\alpha(A)) + \mathbf{E}\left(\left\langle \mathrm{grad}\hat{\phi}_\alpha(A),\, \hat{A} - A \right\rangle\right) = \phi_\alpha\ (\alpha \in \{z, y, x\}). \tag{3.2.23}$$

□

It follows that the loss functions defined by the *mean squared error* can be written in a simpler form:

$$\tilde{f}_1(\underline{\lambda}, \underline{\phi}, \underline{\tau}, \underline{\vartheta}, N) = \mathbf{Var}(\tilde{\phi}_z) + \mathbf{Var}(\tilde{\phi}_y) + \mathbf{Var}(\tilde{\phi}_x), \tag{3.2.24}$$

$$\tilde{f}_2(\underline{\lambda}, \underline{\phi}, \underline{\tau}, \underline{\vartheta}, N) = \mathbf{Var}(\tilde{\lambda}_1) + \mathbf{Var}(\tilde{\lambda}_2) + \mathbf{Var}(\tilde{\lambda}_3). \tag{3.2.25}$$

(3.2.17) and (3.2.18) show that the linearized estimators are linear combinations of the the estimated channel matrix elements, the $\hat{a}_{ij}$-s (up to an additive constant). By Lemma 10, the variance of the random variables of this type is computable, hence $\tilde{f}_1$ and $\tilde{f}_2$ can be written in a closed form.

The parameter estimation

$$T : \mathbf{M}_3(\mathbb{R}) \to \mathbb{R}^6;\ \hat{A} \mapsto (\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\phi}_z, \hat{\phi}_y, \hat{\phi}_x) \tag{3.2.26}$$

is the (left) inverse of the channel parametrization

$$A : \mathcal{D} \to \mathbf{M}_3(\mathbb{R});\ (\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x) \mapsto A(\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x), \tag{3.2.27}$$

hence

$$\mathrm{d}T\left(A(\underline{\lambda}, \underline{\phi})\right) = \left(\mathrm{d}A(\underline{\lambda}, \underline{\phi})\right)^{-1}. \tag{3.2.28}$$

Because of the rotational invariance, it is enough to investigate the channels with angle parameters $\phi_z = \phi_y = \phi_x = 0$. It follows from (3.0.13) that

$$A(\lambda_1, \lambda_2, \lambda_3, \phi_z, 0, 0) = \begin{pmatrix} \lambda_1 \cos^2 \phi_z + \lambda_2 \sin^2 \phi_z & (\lambda_1 - \lambda_2) \sin \phi_z \cos \phi_z & 0 \\ (\lambda_1 - \lambda_2) \sin \phi_z \cos \phi_z & \lambda_1 \sin^2 \phi_z + \lambda_2 \cos^2 \phi_z & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{3.2.29}$$

$$A(\lambda_1, \lambda_2, \lambda_3, 0, \phi_y, 0) = \begin{pmatrix} \lambda_1 \cos^2 \phi_y + \lambda_3 \sin^2 \phi_y & 0 & (\lambda_1 - \lambda_3) \sin \phi_y \cos \phi_y \\ 0 & 1 & 0 \\ (\lambda_1 - \lambda_3) \sin \phi_y \cos \phi_y & 0 & \lambda_1 \sin^2 \phi_y + \lambda_3 \cos^2 \phi_y \end{pmatrix}, \tag{3.2.30}$$

$$A(\lambda_1, \lambda_2, \lambda_3, 0, 0, \phi_x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda_2 \cos^2 \phi_x + \lambda_3 \sin^2 \phi_x & (\lambda_2 - \lambda_3) \sin \phi_x \cos \phi_x \\ 0 & (\lambda_2 - \lambda_3) \sin \phi_x \cos \phi_x & \lambda_2 \sin^2 \phi_x + \lambda_3 \cos^2 \phi_x \end{pmatrix}, \tag{3.2.31}$$

therefore

$$\frac{\partial A}{\partial \phi_z}(\lambda_1, \lambda_2, \lambda_3, 0, 0, 0) = \begin{pmatrix} 0 & (\lambda_1 - \lambda_2) & 0 \\ (\lambda_1 - \lambda_2) & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{3.2.32}$$

$$\frac{\partial A}{\partial \phi_y}(\lambda_1, \lambda_2, \lambda_3, 0, 0, 0) = \begin{pmatrix} 0 & 0 & (\lambda_1 - \lambda_3) \\ 0 & 0 & 0 \\ (\lambda_1 - \lambda_3) & 0 & 0 \end{pmatrix}, \tag{3.2.33}$$

$$\frac{\partial A}{\partial \phi_x}(\lambda_1, \lambda_2, \lambda_3, 0, 0, 0) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & (\lambda_2 - \lambda_3) \\ 0 & (\lambda_2 - \lambda_3) & 0 \end{pmatrix}. \tag{3.2.34}$$

On the other hand $A(\underline{\lambda}, \underline{0}) = \mathrm{Diag}(\lambda_1, \lambda_2, \lambda_3)$, hence

$$\frac{\partial A}{\partial \lambda_1}(\underline{\lambda}, \underline{0}) = \mathrm{Diag}(1, 0, 0),\ \frac{\partial A}{\partial \lambda_2}(\underline{\lambda}, \underline{0}) = \mathrm{Diag}(0, 1, 0),\ \frac{\partial A}{\partial \lambda_3}(\underline{\lambda}, \underline{0}) = \mathrm{Diag}(0, 0, 1). \tag{3.2.35}$$

Let the derivative of $A$ be written in the form

$$\mathrm{d}A = \begin{pmatrix} \frac{\partial a_{11}}{\partial \lambda_1} & \frac{\partial a_{11}}{\partial \lambda_2} & \frac{\partial a_{11}}{\partial \lambda_3} & \frac{\partial a_{11}}{\partial \phi_z} & \frac{\partial a_{11}}{\partial \phi_y} & \frac{\partial a_{11}}{\partial \phi_x} \\ \frac{\partial a_{22}}{\partial \lambda_1} & \frac{\partial a_{22}}{\partial \lambda_2} & \frac{\partial a_{22}}{\partial \lambda_3} & \frac{\partial a_{22}}{\partial \phi_z} & \frac{\partial a_{22}}{\partial \phi_y} & \frac{\partial a_{22}}{\partial \phi_x} \\ \frac{\partial a_{33}}{\partial \lambda_1} & \frac{\partial a_{33}}{\partial \lambda_2} & \frac{\partial a_{33}}{\partial \lambda_3} & \frac{\partial a_{33}}{\partial \phi_z} & \frac{\partial a_{33}}{\partial \phi_y} & \frac{\partial a_{33}}{\partial \phi_x} \\ \frac{\partial a_{12}}{\partial \lambda_1} & \frac{\partial a_{12}}{\partial \lambda_2} & \frac{\partial a_{12}}{\partial \lambda_3} & \frac{\partial a_{12}}{\partial \phi_z} & \frac{\partial a_{12}}{\partial \phi_y} & \frac{\partial a_{12}}{\partial \phi_x} \\ \frac{\partial a_{13}}{\partial \lambda_1} & \frac{\partial a_{13}}{\partial \lambda_2} & \frac{\partial a_{13}}{\partial \lambda_3} & \frac{\partial a_{13}}{\partial \phi_z} & \frac{\partial a_{13}}{\partial \phi_y} & \frac{\partial a_{13}}{\partial \phi_x} \\ \frac{\partial a_{23}}{\partial \lambda_1} & \frac{\partial a_{23}}{\partial \lambda_2} & \frac{\partial a_{23}}{\partial \lambda_3} & \frac{\partial a_{23}}{\partial \phi_z} & \frac{\partial a_{23}}{\partial \phi_y} & \frac{\partial a_{23}}{\partial \phi_x} \end{pmatrix}. \tag{3.2.36}$$

(3.2.32), (3.2.33), (3.2.34) and (3.2.35) show that

$$\mathrm{d}A(\lambda_1, \lambda_2, \lambda_3, 0, 0, 0) = \mathrm{Diag}(1, 1, 1, \lambda_1 - \lambda_2, \lambda_1 - \lambda_3, \lambda_2 - \lambda_3), \tag{3.2.37}$$

It follows from (3.2.28) and (3.2.37) that

$$\mathrm{d}T\left(A(\underline{\lambda},\underline{0})\right)=\mathrm{Diag}(1,1,1,\frac{1}{\lambda_1-\lambda_2},\frac{1}{\lambda_1-\lambda_3},\frac{1}{\lambda_2-\lambda_3}). \tag{3.2.38}$$

By the construction of $T$,

$$T=T|\mathbf{M}_3^s(\mathbb{R})\circ S, \tag{3.2.39}$$

where $\mathbf{M}_3^s(\mathbb{R})\subset\mathbf{M}_3(\mathbb{R})$ is the subspace of the symmetric matrices and $S$ is the orthogonal projection onto this subspace ($S:\ \hat{A}\mapsto\hat{A}_s:=\frac{1}{2}\left(\hat{A}+(\hat{A})^T\right)$). Set

$$\hat{a}_{12,s}=\frac{1}{2}\left(\hat{a}_{12}+\hat{a}_{21}\right),\ \hat{a}_{13,s}=\frac{1}{2}\left(\hat{a}_{13}+\hat{a}_{31}\right),\ \hat{a}_{23,s}=\frac{1}{2}\left(\hat{a}_{23}+\hat{a}_{32}\right). \tag{3.2.40}$$

With this notation, the result of (3.2.38) is the following:

$$\frac{\partial\hat{\lambda}_1}{\partial\hat{a}_{11}}=\frac{\partial\hat{\lambda}_2}{\partial\hat{a}_{22}}=\frac{\partial\hat{\lambda}_3}{\partial\hat{a}_{33}}=1,\ \frac{\partial\hat{\phi}_z}{\partial\hat{a}_{12,s}}=\frac{1}{\lambda_1-\lambda_2},\ \frac{\partial\hat{\phi}_y}{\partial\hat{a}_{13,s}}=\frac{1}{\lambda_1-\lambda_3},\ \frac{\partial\hat{\phi}_x}{\partial\hat{a}_{23,s}}=\frac{1}{\lambda_2-\lambda_3}, \tag{3.2.41}$$

if $\hat{A}=\mathrm{Diag}(\lambda_1,\lambda_2,\lambda_3)$, and all the other partial derivatives vanish. Applying the chain rule to the equation (3.2.39) one gets the gradient of the parameter estimations.

**Lemma 12.** *The non-vanishing partial derivatives of the parameter estimations $\hat{\lambda}_i,\hat{\phi}_\alpha$ ($i\in\{1,2,3\},\alpha\in\{z,y,x\}$) in the point $\hat{A}=\mathrm{Diag}(\lambda_1,\lambda_2,\lambda_3)$ are the following:*

$$\frac{\partial\hat{\lambda}_1}{\partial\hat{a}_{11}}=\frac{\partial\hat{\lambda}_2}{\partial\hat{a}_{22}}=\frac{\partial\hat{\lambda}_3}{\partial\hat{a}_{33}}=1,\ \frac{\partial\hat{\phi}_z}{\partial\hat{a}_{12}}=\frac{\partial\hat{\phi}_z}{\partial\hat{a}_{21}}=\frac{1}{2(\lambda_1-\lambda_2)},$$

$$\frac{\partial\hat{\phi}_y}{\partial\hat{a}_{13}}=\frac{\partial\hat{\phi}_y}{\partial\hat{a}_{31}}=\frac{1}{2(\lambda_1-\lambda_3)},\ \frac{\partial\hat{\phi}_x}{\partial\hat{a}_{23}}=\frac{\partial\hat{\phi}_x}{\partial\hat{a}_{32}}=\frac{1}{2(\lambda_2-\lambda_3)}. \tag{3.2.42}$$

By Lemma 12, the variance of the linearized parameter estimations defined in (3.2.17) and (3.2.18) has the form

$$\mathbf{Var}(\tilde{\lambda}_1)=\mathbf{Var}(\hat{a}_{11}),\ \mathbf{Var}(\tilde{\lambda}_2)=\mathbf{Var}(\hat{a}_{22}),\ \mathbf{Var}(\tilde{\lambda}_3)=\mathbf{Var}(\hat{a}_{33}) \tag{3.2.43}$$

and

$$\mathbf{Var}(\tilde{\phi}_z)=\frac{1}{4(\lambda_1-\lambda_2)^2}\mathbf{Var}(\hat{a}_{12}+\hat{a}_{21}),\ \mathbf{Var}(\tilde{\phi}_y)=\frac{1}{4(\lambda_1-\lambda_3)^2}\mathbf{Var}(\hat{a}_{13}+\hat{a}_{31}), \tag{3.2.44}$$

$$\mathbf{Var}(\tilde{\phi}_x)=\frac{1}{4(\lambda_2-\lambda_3)^2}\mathbf{Var}(\hat{a}_{23}+\hat{a}_{32}). \tag{3.2.45}$$

$A(\hat{\lambda}_1,\hat{\lambda}_2,\hat{\lambda}_3,\hat{\phi}_z,\hat{\phi}_y,\hat{\phi}_x)=\hat{A}_s$, hence the distance of the real and the estimated channel matrix is the following:

$$\begin{aligned}||A(\lambda_1,\lambda_2,\lambda_3,\phi_z,\phi_y,\phi_x)-A(\hat{\lambda}_1,\hat{\lambda}_2,\hat{\lambda}_3,\hat{\phi}_z,\hat{\phi}_y,\hat{\phi}_x)||^2=||A(\lambda_1,\lambda_2,\lambda_3,\phi_z,\phi_y,\phi_x)-\hat{A}_s||^2=\\ =(\hat{a}_{11}-a_{11})^2+(\hat{a}_{22}-a_{22})^2+(\hat{a}_{33}-a_{33})^2+\\ +2\left(\frac{\hat{a}_{12}-a_{12}}{2}+\frac{\hat{a}_{21}-a_{21}}{2}\right)^2+2\left(\frac{\hat{a}_{13}-a_{13}}{2}+\frac{\hat{a}_{31}-a_{31}}{2}\right)^2+2\left(\frac{\hat{a}_{23}-a_{23}}{2}+\frac{\hat{a}_{32}-a_{32}}{2}\right)^2.\end{aligned} \tag{3.2.46}$$

$\hat{A}$ is unbiased, therefore

$$\mathbf{E}\left(||A(\lambda_1,\lambda_2,\lambda_3,\phi_z,\phi_y,\phi_x)-A(\hat{\lambda}_1,\hat{\lambda}_2,\hat{\lambda}_3,\hat{\phi}_z,\hat{\phi}_y,\hat{\phi}_x)||^2\right)=$$

$$=\mathbf{Var}(\hat{a}_{11})+\mathbf{Var}(\hat{a}_{22})+\mathbf{Var}(\hat{a}_{33})+\frac{1}{2}\left(\mathbf{Var}(\hat{a}_{12}+\hat{a}_{21})+\mathbf{Var}(\hat{a}_{13}+\hat{a}_{31})+\mathbf{Var}(\hat{a}_{23}+\hat{a}_{32})\right). \tag{3.2.47}$$

**Summary.** *The loss functions $\tilde{f}_1, \tilde{f}_2$ and $f_3$ defined in (3.2.19), (3.2.20) and (3.2.3) have the form*

$$\tilde{f}_1(\underline{\lambda},\underline{0},\underline{\tau},\underline{\vartheta},N)=$$

$$=\frac{1}{4(\lambda_1-\lambda_2)^2}\mathbf{Var}(\hat{a}_{12}+\hat{a}_{21})+\frac{1}{4(\lambda_1-\lambda_3)^2}\mathbf{Var}(\hat{a}_{13}+\hat{a}_{31})+\frac{1}{4(\lambda_2-\lambda_3)^2}\mathbf{Var}(\hat{a}_{23}+\hat{a}_{32}), \tag{3.2.48}$$

$$\tilde{f}_2(\underline{\lambda},\underline{0},\underline{\tau},\underline{\vartheta},N)=\mathbf{Var}(\hat{a}_{11})+\mathbf{Var}(\hat{a}_{22})+\mathbf{Var}(\hat{a}_{33}), \tag{3.2.49}$$

$$f_3(\underline{\lambda},\underline{0},\underline{\tau},\underline{\vartheta},N)=$$

$$=\mathbf{Var}(\hat{a}_{11})+\mathbf{Var}(\hat{a}_{22})+\mathbf{Var}(\hat{a}_{33})+\frac{1}{2}\left(\mathbf{Var}(\hat{a}_{12}+\hat{a}_{21})+\mathbf{Var}(\hat{a}_{13}+\hat{a}_{31})+\mathbf{Var}(\hat{a}_{23}+\hat{a}_{32})\right). \tag{3.2.50}$$

## 3.3 Optimal tomography settings

In this section we determine the optimal tomography settings by minimizing the quantities in the above summary.

### 3.3.1 Optimal estimation of the channel matrix

**Theorem 13.**

$$f_3(\underline{\lambda},\underline{0},\underline{\tau},\underline{\vartheta},N)\geq\frac{1}{N}\left(6-(\lambda_1^2+\lambda_2^2+\lambda_3^2)\right), \tag{3.3.1}$$

*and (3.3.1) holds with equality, if $\underline{\tau}=\underline{\vartheta}=\underline{0}$.*

*Proof.* Observe that

$$f_3(\underline{\lambda},\underline{0},\underline{\tau},\underline{\vartheta},N)=-\frac{1}{2}\Bigg(\mathbf{Var}(\hat{a}_{12}-\hat{a}_{21})+\mathbf{Var}(\hat{a}_{13}-\hat{a}_{31})+\mathbf{Var}(\hat{a}_{23}-\hat{a}_{32})\Bigg)+\sum_{i,j\in\{1,2,3\}}\mathbf{Var}(\hat{a}_{ij}). \tag{3.3.2}$$

Using the representing vectors one can see that

$$\hat{A}=\left(R(\underline{\tau})\otimes R(\underline{\vartheta})\right)\hat{X}, \tag{3.3.3}$$

where $R(\underline{\zeta})=R_z(\zeta_z)R_y(\zeta_y)R_x(\zeta_x)$, $(\zeta\in\{\tau,\vartheta\})$. The $\hat{x}_{ij}$-s are independent estimators, hence

$$\mathbf{Var}\left(\hat{A}\right)=\left(R(\underline{\tau})\otimes R(\underline{\vartheta})\right)^{\mathbf{2}}\mathbf{Var}\left(\hat{X}\right), \tag{3.3.4}$$

where $(\cdot)^{\mathbf{2}}$ denotes the Hadamart-square. By Lemma 8, $(R(\underline{\tau}) \otimes R(\underline{\vartheta}))^{\mathbf{2}}$ is bistochastic and $\mathbf{Var}(\hat{x}_{ij}) = \frac{1-x_{ij}^2}{N}$, hence

$$\sum_{i,j\in\{1,2,3\}} \mathbf{Var}(\hat{a}_{ij}) = \frac{1}{N}\left(9 - \sum_{i,j\in\{1,2,3\}} x_{ij}^2\right). \tag{3.3.5}$$

On the other hand, by Lemma 10

$$\mathbf{Var}(\hat{a}_{ij} - \hat{a}_{ji}) = \sum_{l\in H} \left(c_{ij,l}\left(\underline{\tau}, \underline{\vartheta}\right) - c_{ji,l}\left(\underline{\tau}, \underline{\vartheta}\right)\right)^2 \frac{1-x_l^2}{N}. \tag{3.3.6}$$

If $i \neq j$, then

$$\sum_{l\in H} \left(c_{ij,l}\left(\underline{\tau}, \underline{\vartheta}\right) - c_{ji,l}\left(\underline{\tau}, \underline{\vartheta}\right)\right)^2 = 2, \tag{3.3.7}$$

because it is the squared norm of the sum of two orthogonal unit vectors. $\frac{1}{N}(1-x_l^2) \leq \frac{1}{N}$ $(\forall l \in H)$, hence

$$\sum_{l\in H} \left(c_{ij,l}\left(\underline{\tau}, \underline{\vartheta}\right) - c_{ji,l}\left(\underline{\tau}, \underline{\vartheta}\right)\right)^2 \frac{1-x_l^2}{N} \leq \frac{2}{N}. \tag{3.3.8}$$

It follows that

$$\mathbf{Var}(\hat{a}_{12} - \hat{a}_{21}) \leq \frac{2}{N},\ \mathbf{Var}(\hat{a}_{13} - \hat{a}_{31}) \leq \frac{2}{N},\ \mathbf{Var}(\hat{a}_{23} - \hat{a}_{32}) \leq \frac{2}{N}. \tag{3.3.9}$$

The last step is that

$$\sum_{i,j\in\{1,2,3\}} x_{ij}^2 = \mathrm{Tr} XX^T = \mathrm{Tr}\left(R(\underline{\tau})^T \Lambda(\lambda_1,\lambda_2,\lambda_3) R(\underline{\vartheta}) R(\underline{\vartheta})^T \Lambda(\lambda_1,\lambda_2,\lambda_3)^T R(\underline{\tau})\right) = \lambda_1^2+\lambda_2^2+\lambda_3^2, \tag{3.3.10}$$

and the inequality (3.3.1) is proved. It is easy to check that if $\underline{\tau} = \underline{\vartheta} = \underline{0}$, then

$$\mathbf{Var}(\hat{a}_{ij}) = \frac{1}{N}(1-\delta_{ij}\lambda_i^2),\ \mathbf{Var}(\hat{a}_{12} - \hat{a}_{21}) = \mathbf{Var}(\hat{a}_{13} - \hat{a}_{31}) = \mathbf{Var}(\hat{a}_{23} - \hat{a}_{32}) = \frac{2}{N},$$

where $i, j \in \{1,2,3\}$ and $\delta_{ij}$ is the Kronecker-symbol, hence the minimum is obtained. □

### 3.3.2 Optimal estimation of the contraction parameters

**Theorem 14.**

$$\tilde{f}_2(\underline{\lambda}, \underline{0}, \underline{\tau}, \underline{\vartheta}, N) \geq \frac{1}{N}\left(3 - (\lambda_1^2 + \lambda_2^2 + \lambda_3^2)\right), \tag{3.3.11}$$

*and (3.3.11) holds with equality, if $\underline{\tau} = \underline{\vartheta} = 0$.*

*Proof.*

$$\tilde{f}_2(\underline{\lambda}, \underline{0}, \underline{\tau}, \underline{\vartheta}, N) = \sum_{i\in\{1,2,3\}} \mathbf{Var}(\hat{a}_{ii}) = \sum_{i,j\in\{1,2,3\}} \mathbf{Var}(\hat{a}_{ij}) - \sum_{i\neq j} \mathbf{Var}(\hat{a}_{ij}). \tag{3.3.12}$$

(3.3.4) shows that $\mathbf{Var}\left(\hat{a}_{ij}\right) \leq \frac{1}{N}$ $(\forall i, j)$, because $\frac{1-x_{ij}^2}{N} \leq \frac{1}{N}$ for every $i, j$, and $(R(\underline{\tau}) \otimes R(\underline{\vartheta}))^{\mathbf{2}}$ is double stochastic. Therefore

$$\sum_{i\neq j} \mathbf{Var}(\hat{a}_{ij}) \leq \frac{6}{N}. \tag{3.3.13}$$

If we compare the results of (3.3.5) and (3.3.10) with (3.3.12), we get that

$$\mathbf{Var}(\hat{a}_{11}) + \mathbf{Var}(\hat{a}_{22}) + \mathbf{Var}(\hat{a}_{33}) \geq \frac{1}{N}\left(9 - (\lambda_1^2 + \lambda_2^2 + \lambda_3^2)\right) - \frac{6}{N}, \tag{3.3.14}$$

hence the inequality (3.3.11) is proved. It is easy to check that if $\underline{\tau} = \underline{\vartheta} = \underline{0}$, then

$$\mathbf{Var}(\hat{a}_{11}) = \frac{1}{N}(1-\lambda_1^2),\ \mathbf{Var}(\hat{a}_{22}) = \frac{1}{N}(1-\lambda_2^2),\ \mathbf{Var}(\hat{a}_{33}) = \frac{1}{N}(1-\lambda_3^2), \tag{3.3.15}$$

hence the minimum is obtained. □

**Practical consequence.** *The estimation of the channel matrix and the contraction parameters is optimal, if the input states, the von Neumann measurements and the Pauli channel have the same directions.*

**Note.** *In this work, a tomography scheme for Pauli channels with* unknown *channel directions is presented. Nonetheless, one can apply this scheme for Pauli channels with known channel directions. Assume that the channel directions are* $\left\{\frac{1}{2}(I + t\sigma_i)\right\}_{i=1}^{3}$. *Now, we do not estimate any angle parameters and the estimators of the contraction parameters get simpler:* $\hat{\lambda}_i = \hat{a}_{ii}$ $(i \in \{1,2,3\})$. *Recall that*

$$\sum_{i=1}^{3} \mathbf{Var}\left(\hat{a}_{ii}\right) \geq \frac{1}{N}\left(3 - (\lambda_1^2 + \lambda_2^2 + \lambda_3^2)\right) \tag{3.3.16}$$

*and if* $\underline{\tau} = \underline{\vartheta} = \underline{0}$, *then this inequality holds with equality. Therefore, the tomography of a Pauli channel with known directions is optimal if the input states, the von Neumann measurements and the Pauli channel have the same directions. Hence the proof of Theorem 14 is a new verification of the result that appiers in [16].*

### 3.3.3 Optimal estimation of the angle parameters

(3.3.3) shows that the quantities $\mathbf{Var}(\hat{a}_{ij} + \hat{a}_{ji})$ $(i \neq j)$ can be written as explicit functions of $\underline{\tau}, \underline{\vartheta}, X$ and $N$, and by (3.1.11), $X$ is an explicit function of $\underline{\lambda}, \underline{\tau}$ and $\underline{\vartheta}$ as well. Therefore the quantity

$$\tilde{f}_1(\underline{\lambda}, \underline{0}, \underline{\tau}, \underline{\vartheta}, N) =$$
$$= \frac{1}{4(\lambda_1-\lambda_2)^2}\mathbf{Var}(\hat{a}_{12} + \hat{a}_{21}) + \frac{1}{4(\lambda_1-\lambda_3)^2}\mathbf{Var}(\hat{a}_{13} + \hat{a}_{31}) + \frac{1}{4(\lambda_2-\lambda_3)^2}\mathbf{Var}(\hat{a}_{23} + \hat{a}_{32}) \tag{3.3.17}$$

can be expressed in a closed formula. With fix $\underline{\lambda}$ and $N$, the optimization of $\tilde{f}_1$ is an extremal value problem with six variables $(\tau_z, \tau_y, \tau_x, \vartheta_z, \vartheta_y, \vartheta_x)$ that can not be solved analytically. Therefore we seek the best estimation strategy by numerical optimization with some fixed contraction parameters.

For example, set $\lambda_1 = 0.8, \lambda_2 = 0.65, \lambda_3 = 0.5$ and $N = 1000$. In this case, the optimal angles $\left(\underline{\tau}_{opt}, \underline{\vartheta}_{opt}\right)$ that can be calculated numerically do not show any regularity except $\underline{\tau}_{opt} = \underline{\vartheta}_{opt}$. However, $\tilde{f}_1$ is nearly minimal in two special points: $\tilde{f}_1 = 0.03676$, if $\tau_z = \vartheta_z = \frac{\pi}{4}, \tau_y = \vartheta_y = \frac{\pi}{4}, \tau_x = \vartheta_x = 0$ or $\tau_z = \vartheta_z = \frac{\pi}{4}, \tau_y = \vartheta_y = 0, \tau_x = \vartheta_x = \frac{\pi}{4}$, while the minimum is

$$\min_{\underline{\tau},\, \underline{\vartheta}} \tilde{f}_1 = 0.03634. \tag{3.3.18}$$

The difference can be considered small, for comparison, $\tilde{f}_1(\underline{\tau} = \underline{0}, \underline{\vartheta} = \underline{0}) = 0.05$.

We get a similar result, if we fix the contraction parameters $\lambda_1 = 0.9, \lambda_2 = 0.67, \lambda_3 = 0.6$ and the number of measurements $N = 1000$. In this case $\underline{\tau}_{opt} = \underline{\vartheta}_{opt}$ holds as well, and in the above mentioned points $\underline{\tau} = \underline{\vartheta} = (\frac{\pi}{4}, \frac{\pi}{4}, 0)$ and $\underline{\tau} = \underline{\vartheta} = (\frac{\pi}{4}, 0, \frac{\pi}{4})$ $\tilde{f}_1$ is nearly minimal: $\tilde{f}_1 = 0.01675$, while

$$\min_{\underline{\tau},\ \underline{\vartheta}} \tilde{f}_1 = 0.01659. \tag{3.3.19}$$

For comparison, $\tilde{f}_1(\underline{\tau} = \underline{0}, \underline{\vartheta} = \underline{0}) = 0.02446$.

There are some results of empirical simulations in the *Appendix.* These simulation results are useful to illustrate and verify the theoretical results and we can formulate conjectures based on these investigations as well. Figure 5.1 and Figure 5.2 in the Appendix show the graph of the function $\tilde{f}_1$ with $\lambda_1 = 0.8,\ \lambda_2 = 0.65,\ \lambda_3 = 0.5,\ N = 1000,\ \underline{\tau} = \underline{\vartheta}$, and with fixed $\vartheta_x = \tau_x = 0$ and $\vartheta_x = \tau_x = \frac{\pi}{4}$, respectively.

Now, we can formulate a conjecture based on the numerical optimization computations.

**Conjecture.** *For any fixed parameters* $\underline{\lambda}, N$

- *if* $\tilde{f}_1(\underline{\lambda}, \underline{0}, \underline{\tau}, \underline{\vartheta}, N)$ *is minimal at* $(\underline{\tau}_{opt}, \underline{\vartheta}_{opt})$*, then* $\underline{\tau}_{opt} = \underline{\vartheta}_{opt}$*,*
- *the estimation strategies described by parameters* $\underline{\tau}_1 = \underline{\vartheta}_1 = (\frac{\pi}{4}, \frac{\pi}{4}, 0)$ *and* $\underline{\tau}_2 = \underline{\vartheta}_2 = (\frac{\pi}{4}, 0, \frac{\pi}{4})$ *are* nearly *optimal.*

## 3.4 Pauli channel with known parameters in one direction

In the previous subsection we did not succeed to find the optimal tomography setting to estimate the angle parameters analytically. However, if we assume some a priori knowledge about the Pauli channel, this optimization problem has an analytical solution.

Let us assume that we have the following informations about the orthonormal system $\{v_j\}_{j=1}^3 \subset \mathbf{M}_2^{s.a.}(\mathbb{C})$ and about the contraction parameters $\lambda_1, \lambda_2, \lambda_3$ that determine the Pauli channel:

$$v_3 = \sigma_3 \text{ and } \lambda_3 = 0. \tag{3.4.1}$$

In this case the channel matrix has the following simplified form

$$A(\lambda_1, \lambda_2, \phi) = R(\phi)\Lambda(\lambda_1, \lambda_2)R(\phi)^{-1}, \tag{3.4.2}$$

that is

$$A(\lambda_1, \lambda_2, \phi) = \begin{pmatrix} \lambda_1 \cos\phi^2 + \lambda_2 \sin\phi^2 & (\lambda_1 - \lambda_2)\sin\phi\cos\phi & 0 \\ (\lambda_1 - \lambda_2)\sin\phi\cos\phi & \lambda_1 \sin\phi^2 + \lambda_2 \cos\phi^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{3.4.3}$$

Now it is easy to show that the input states and the von Neumann measurements should be orthogonal vectors in the plane spanned by $\sigma_1$ and $\sigma_2$. Hence, the input states and the measurements can be parametrized using a single angle parameter $\vartheta$ and $\tau$, respectively:

$$\Theta = \begin{pmatrix} \theta_{11} & \theta_{21} & 0 \\ \theta_{12} & \theta_{22} & 0 \\ 0 & 0 & 1 \end{pmatrix} = R(\vartheta), \quad M = \begin{pmatrix} m_{11} & m_{21} & 0 \\ m_{12} & m_{22} & 0 \\ 0 & 0 & 1 \end{pmatrix} = R(\tau).$$

Let us linearize the estimation of the angle parameter. The parameter estimation $T : \mathbf{M}_2(\mathbb{R}) \to \mathbb{R}^3;\ \hat{A} \mapsto (\hat{\lambda}_1, \hat{\lambda}_2, \hat{\phi})$ is the left inverse of the channel parametrization

$$A : \mathbb{R}^3 \supset \mathcal{D} \to \mathbf{M}_2(\mathbb{R});\ (\lambda_1, \lambda_2, \phi) \mapsto A(\lambda_1, \lambda_2, \phi), \tag{3.4.4}$$

therefore the derivatives satisfy

$$\mathrm{d}T\left(A(\lambda_1, \lambda_2, \phi)\right) = \left(\mathrm{d}A(\lambda_1, \lambda_2, \phi)\right)^{-1}. \tag{3.4.5}$$

Let us write the derivative of $A$ in the matrix form

$$\mathrm{d}A = \begin{pmatrix} \frac{\partial a_{11}}{\partial \lambda_1} & \frac{\partial a_{11}}{\partial \lambda_2} & \frac{\partial a_{11}}{\partial \phi} \\ \frac{\partial a_{22}}{\partial \lambda_1} & \frac{\partial a_{22}}{\partial \lambda_2} & \frac{\partial a_{22}}{\partial \phi} \\ \frac{\partial a_{12}}{\partial \lambda_1} & \frac{\partial a_{12}}{\partial \lambda_2} & \frac{\partial a_{12}}{\partial \phi} \end{pmatrix}. \tag{3.4.6}$$

It follows from (3.4.3) that

$$\mathrm{d}A(\lambda_1, \lambda_2, \phi) = \begin{pmatrix} \cos^2\phi & \sin^2\phi & (\lambda_2 - \lambda_1)\sin 2\phi \\ \sin^2\phi & \cos^2\phi & (\lambda_1 - \lambda_2)\sin 2\phi \\ \frac{1}{2}\sin 2\phi & -\frac{1}{2}\sin 2\phi & (\lambda_1 - \lambda_2)\cos 2\phi \end{pmatrix}. \tag{3.4.7}$$

Because of the rotational invariance it is enough to investigate the channels with angle parameter $\phi = 0$, hence it is useful to observe that

$$\mathrm{d}A(\lambda_1, \lambda_2, 0) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (\lambda_1 - \lambda_2) \end{pmatrix}. \tag{3.4.8}$$

Because of the construction of $T$,

$$T = T|\mathbf{M}_2^s(\mathbb{R}) \circ S, \tag{3.4.9}$$

where $\mathbf{M}_2^s(\mathbb{R}) \subset \mathbf{M}_2(\mathbb{R})$ is the subspace of the symmetric matrices and $S$ is the orthoprojection onto this subspace ($S : \hat{A} \mapsto \hat{A}_s := \frac{1}{2}\left(\hat{A} + (\hat{A})^T\right)$). Let us introduce the notation $\hat{a}_{12,s} = \frac{1}{2}(\hat{a}_{12} + \hat{a}_{21})$.

If we calculate the inverse of the matrix in (3.4.8), we get that in the point $\hat{A} = \mathrm{Diag}(\lambda_1, \lambda_2, 0)$ we have

$$\frac{\partial \hat{\lambda}_1}{\partial \hat{a}_{11}} = \frac{\partial \hat{\lambda}_2}{\partial \hat{a}_{22}} = 1,\ \frac{\partial \hat{\phi}}{\partial \hat{a}_{12,s}} = \frac{1}{\lambda_1 - \lambda_2}, \tag{3.4.10}$$

and all the other partial derivatives vanish. By applying the chain rule to the equation (3.4.9), we get the derivatives of the parameter estimations.

**Lemma 15.** *With the usual notations*

$$\frac{\partial \hat{\lambda}_1}{\partial \hat{a}_{11}} = \frac{\partial \hat{\lambda}_2}{\partial \hat{a}_{22}} = 1,\ \frac{\partial \hat{\phi}}{\partial \hat{a}_{12}} = \frac{\partial \hat{\phi}}{\partial \hat{a}_{21}} = \frac{1}{2(\lambda_1 - \lambda_2)} \tag{3.4.11}$$

*and the other components of* $\mathrm{d}T$ *are zeros in the point* $\hat{A} = \mathrm{Diag}(\lambda_1, \lambda_2, 0)$.

Because of the unbiasedness of $\hat{A}$, the linearized estimator of the angle parameter with the base point $\mathbf{E}(\hat{A})$ is the following:

$$\tilde{\phi} = \hat{\phi}(A) + \frac{\partial \hat{\phi}}{\partial \hat{a}_{11}}(A) \cdot (\hat{a}_{11} - a_{11}) + \cdots + \frac{\partial \hat{\phi}}{\partial \hat{a}_{22}}(A) \cdot (\hat{a}_{22} - a_{22}). \tag{3.4.12}$$

By Lemma 15, if the angle parameter of the channel is $\phi = 0$, then

$$\tilde{\phi} = \hat{\phi}\left(A(\lambda_1, \lambda_2, 0)\right) + \frac{1}{2(\lambda_1 - \lambda_2)}\left(\hat{a}_{12} - a_{12} + \hat{a}_{21} - a_{21}\right). \tag{3.4.13}$$

The mean squared error of this estimation has the form

$$\tilde{g}(\lambda_1, \lambda_2, \phi, \tau, \vartheta, N) = \mathbf{E}\left(\tilde{\phi} - \phi\right)^2, \tag{3.4.14}$$

The estimation $\hat{A}$ is unbiased and the linearized estimation $\tilde{\phi}$ is unbiased as well, hence it follows from (3.4.13) that

$$\tilde{g}(\lambda_1, \lambda_2, \phi, \tau, \vartheta, N) = \mathbf{Var}(\tilde{\phi}) = \frac{1}{4(\lambda_1 - \lambda_2)^2}\mathbf{Var}\left(\hat{a}_{12} + \hat{a}_{21}\right). \tag{3.4.15}$$

By (3.2.15), the expression $\mathbf{Var}\left(\hat{a}_{12} + \hat{a}_{21}\right)$ can be written as an explicit function of $\lambda_1, \lambda_2, \tau, \vartheta$ and $N$. We have to solve the minimization problem

$$\tilde{g}(\lambda_1, \lambda_2, \phi, \tau, \vartheta, N) \to \min$$

with two variables $(\tau, \vartheta)$ for fixed $\lambda_1, \lambda_2, N$ values. The result is formulated in the following theorem.

**Theorem 16.** *Assume that $\lambda_1 > \lambda_2$, $\lambda_2 \neq 0$ and $\lambda_1 \neq -\lambda_2$.*

1. *If $(\lambda_1 + \lambda_2)^2 \geq 2(\lambda_1 - \lambda_2)^2$, then $\tilde{g}(\lambda_1, \lambda_2, 0, \tau, \vartheta, N)$ is minimal if and only if*

$$\tau_{opt} = \vartheta_{opt} = \frac{\pi}{4} \quad \left(\mathrm{mod}\ \frac{\pi}{2}\right).$$

*The minimal value of the loss function is*

$$\tilde{g}(\lambda_1, \lambda_2, 0, \frac{\pi}{4}, \frac{\pi}{4}, N) = \frac{1}{4(\lambda_1 - \lambda_2)^2}\frac{1}{2N}\left(4 - (\lambda_1 + \lambda_2)^2\right). \tag{3.4.16}$$

2. *If $(\lambda_1 + \lambda_2)^2 < 2(\lambda_1 - \lambda_2)^2$, then $\tilde{g}(\lambda_1, \lambda_2, 0, \tau, \vartheta, N)$ is minimal if and only if*

$$\tau_{opt} = \vartheta_{opt} = x \ \text{or}\ \frac{\pi}{2} - x \quad \left(\mathrm{mod}\ \frac{\pi}{2}\right),$$

*where $x = \frac{1}{4}\arccos\left(-\frac{(\lambda_1+\lambda_2)^2}{2(\lambda_1-\lambda_2)^2}\right)$. The minimal value is now*

$$\tilde{g}(\lambda_1, \lambda_2, 0, \tau_{opt}, \vartheta_{opt}, N) = \frac{1}{4(\lambda_1 - \lambda_2)^2}\frac{1}{2N}\left(4 - (\lambda_1^2 + \lambda_2^2) - \frac{1}{8}\frac{(\lambda_1 + \lambda_2)^4}{(\lambda_1 - \lambda_2)^2}\right). \tag{3.4.17}$$

*Proof.*

$$\tilde{g}(\lambda_1,\lambda_2,0,\tau,\vartheta,N) = \frac{1}{4(\lambda_1-\lambda_2)^2}\mathbf{Var}\,(\hat{a}_{12}+\hat{a}_{21}) = \frac{1}{4(\lambda_1-\lambda_2)^2}\frac{1}{8N}\times$$

$$\times\left(16-3(\lambda_1^2+\lambda_2^2)-2\lambda_1\lambda_2+(\lambda_1+\lambda_2)^2(\cos 4\tau+\cos 4\vartheta)+(\lambda_1-\lambda_2)^2\cos 4(\tau+\vartheta)\right). \quad (3.4.18)$$

Computing the partial derivatives with respect to $\tau$ and $\vartheta$, one gets that if $\tilde{g}$ has a local extremum in $(\tau,\vartheta)$, then

$$-(\lambda_1+\lambda_2)^2\sin(4\vartheta) = (\lambda_1-\lambda_2)^2\sin(4(\vartheta+\tau)), \quad (3.4.19)$$

$$-(\lambda_1+\lambda_2)^2\sin(4\tau) = (\lambda_1-\lambda_2)^2\sin(4(\vartheta+\tau)). \quad (3.4.20)$$

$(\lambda_1+\lambda_2)^2 \neq 0$, hence (3.4.19) and (3.4.20) show that $\sin(4\vartheta)=\sin(4\tau)$. By a trigonometrical fact, if additionally $4\vartheta \neq 4\tau \pmod{2\pi}$, then $\sin(4\vartheta+4\tau)=0$. Therefore by (3.4.20) and (3.4.20) $\tau=0,\ \vartheta=\frac{\pi}{4} \pmod{\frac{\pi}{2}}$ or $\tau=\frac{\pi}{4},\ \vartheta=0 \pmod{\frac{\pi}{2}}$. It is clear form (3.4.18) that $\tilde{g}$ is strictly greater in these points than in the point $\tau=\vartheta=\frac{\pi}{4} \pmod{\frac{\pi}{2}}$, because

$$-2(\lambda_1+\lambda_2)^2+(\lambda_1-\lambda_2)^2 \leq -(\lambda_1-\lambda_2)^2, \quad (3.4.21)$$

and (3.4.21) holds with equality if and only if $\lambda_2=0$, but we assumed that $\lambda_2\neq 0$. Hence the minimum of $\tilde{g}$ is obtained on the line $\tau=\vartheta$. Therefore, let us assume that $\tau=\vartheta$, and our problem turns into an optimization problem with one variable. It is clear from (3.4.18) that we have to minimize

$$F_{\lambda_1,\lambda_2}(\tau) := 2(\lambda_1+\lambda_2)^2\cos(4\tau)+(\lambda_1-\lambda_2)^2\cos(8\tau) \to \min. \quad (3.4.22)$$

$F_{\lambda_1,\lambda_2}$ is $\frac{\pi}{2}$-periodic in its variable $\tau$-ban, hence we may assume that $\tau\in\left[0,\frac{\pi}{2}\right)$.

1. $F'_{\lambda_1,\lambda_2}(\tau)=0$ if and only if $(\lambda_1+\lambda_2)^2\sin(4\tau) = -2(\lambda_1-\lambda_2)^2\sin(4\tau)\cos(4\tau)$, that is

$$\tau\in\left\{0,\frac{\pi}{4}\right\} \text{ or } \cos(4\tau) = -\frac{(\lambda_1+\lambda_2)^2}{2(\lambda_1-\lambda_2)^2}. \quad (3.4.23)$$

2. $$F''_{\lambda_1,\lambda_2}(\tau) = -32\left((\lambda_1+\lambda_2)^2\cos(4\tau)+2(\lambda_1-\lambda_2)^2\cos(8\tau)\right). \quad (3.4.24)$$

   This expression is negative in $\tau=0$, hence there is a local maximum in $\tau=0$. $F''_{\lambda_1,\lambda_2}$ is positive in $\tau=\frac{\pi}{4}$ if and only if $(\lambda_1+\lambda_2)^2>2(\lambda_1-\lambda_2)^2$.

Therefore if $(\lambda_1+\lambda_2)^2>2(\lambda_1-\lambda_2)^2$, then by (3.4.23) and by the bounded property of the cosine function, $\tilde{g}$ can be extremal only in the points $\tau=0$ and $\tau=\frac{\pi}{4}$. $(\lambda_1+\lambda_2)^2 > 2(\lambda_1-\lambda_2)^2$, hence $\tilde{g}$ has a local minimum in $\tau=\frac{\pi}{4}$. (Recall that in $\tau=0$ there is a local maximum.) $\tilde{g}$ does not have any other local minimum, hence there is a global minimum in $\tau=\frac{\pi}{4}$.

If $(\lambda_1+\lambda_2)^2 = 2(\lambda_1-\lambda_2)^2$, then (3.4.23) shows that if $\tilde{g}$ has a local extremum in $\tau$, then $\tau\in\{0,\frac{\pi}{4}\}$. $\tilde{g}$ has a local maximum in $\tau=0$, hence by the Weierstrass theorem (that states that every continuous peridic function has a minimum) $\tilde{g}$ has a (global) minimum in $\tau=\frac{\pi}{4}$.

If $(\lambda_1+\lambda_2)^2 < 2(\lambda_1-\lambda_2)^2$, then $\tilde{g}$ has local maximum in $\tau = 0$ and in $\tau = \frac{\pi}{4}$. (3.4.23) shows that the minimum can be obtained in

$$\left\{\tau:\ \cos(4\tau) = -\frac{(\lambda_1+\lambda_2)^2}{2(\lambda_1-\lambda_2)^2},\ \tau\in\left[0,\frac{\pi}{2}\right)\right\} =$$

$$= \left\{\frac{1}{4}\arccos\left(-\frac{(\lambda_1+\lambda_2)^2}{2(\lambda_1-\lambda_2)^2}\right), \frac{\pi}{2}-\frac{1}{4}\arccos\left(-\frac{(\lambda_1+\lambda_2)^2}{2(\lambda_1-\lambda_2)^2}\right)\right\}. \qquad (3.4.25)$$

It is clear form (3.4.22) that $F_{\lambda_1,\lambda_2}(\tau) \equiv F_{\lambda_1,\lambda_2}(\frac{\pi}{2}-\tau)$, hence by the Weierstrass theorem, the global minimum is obtained in both of the above mentioned points.

The minimal values of the loss function $\tilde{g}$ (see (3.4.16) and (3.4.17)) are calculated by direct computations from (3.4.18). □

**Practical consequence.** *To optimally estimate the angle parameter, the directions of the input states and the von Neumann measurements have to be the same, but these directions differ from the channel directions and depend on the contraction parameters $\lambda_1$ and $\lambda_2$.*

**Example 4.** *If $\lambda_1 = 0.8$ and $\lambda_2 = 0.2$ the optimal input and measurement directions are $\tau_{opt} = \vartheta_{opt} = \frac{\pi}{4}$. These are the optimal angles in most cases, however, if $\lambda_1 = 1$ and $\lambda_2 = 0$ then $\tau_{opt} = \vartheta_{opt} = \frac{\pi}{3}$ or $\frac{\pi}{6}$.*

# Chapter 4

# Tomography of generalized Pauli channels

As a generalization of the quantum bit Pauli channel, *Petz* and *Ohno* introduced the *generalized* Pauli channels for finite dimensional quantum systems [14]. None of the papers related to this topic define the channel directions (see e.g. [8] and [12]). Our aim is to define the *channel directions* and the *angle parameters* for generalized Pauli channels.

A class of the generalized Pauli channels is in one-to-one correspondence with the *mutually unbiased bases* of the underlying Hilbert space. By this bijection and by a parametrization of the unitary matrices we can define angle parameters for the Pauli channels of this class.

## 4.1 Pauli channels given by Abelian subalgebras

### 4.1.1 Mutually unbiased bases

**Definition 3** (MUB)**.** *Let $\mathcal{F}_1 = \{f_1^1, \dots f_1^n\}, \dots, \mathcal{F}_r = \{f_r^1, \dots f_r^n\}$ be orthonormal bases of the space $\mathbb{C}^n$. If*

$$\left|\left\langle f_k^i, f_l^j \right\rangle\right|^2 = \frac{1}{n} \tag{4.1.1}$$

*holds for every $1 \le k \ne l \le r$ and $i, j \in \{1, \dots, n\}$, then $\mathcal{F}_1, \dots, \mathcal{F}_r$ are called* mutually unbiased bases *(the often used abbreviation is* MUB*).*

If $n = p^M$ with some prime $p$ and $M \in \mathbb{Z}^+$, then there exists a MUB of $\mathbb{C}^n$ with $n+1$ elements [3]. The construction is the following. Let us identify the elements of the standard basis of $\mathbb{C}^n$ with the elements of the finite field $F(n)$. Let us define the function

$$\chi(\theta) := \exp\left(\frac{2\pi i}{p}\left(\theta + \theta^p + \theta^{p^2} + \dots + \theta^{p^{M-1}}\right)\right)$$

and the matrices

$$X_q := \sum_{s \in F(n)} |s+q\rangle \langle s|\,,\ Z_r := \sum_{s \in F(n)} \chi(rs)\, |s\rangle \langle s|\,. \tag{4.1.2}$$

Then $\{\{Z_r:\ r \in F(n)\}\} \cup \{\{X_q Z_{qr}:\ q \in F(n)\}:\ r \in F(n)\}$ is a system of sets with $n+1$ elements. Every element of this system is a set of $n$ *commuting* matrices. One can map the common eigenbasis to every set of commuting matrices. This way we get a MUB with $n+1$ elements. This MUB is called the *fundamental* MUB of $\mathbb{C}^n$.

**Example 5.** *Set $p=2, M=1$. Then $F(n)=\mathbb{Z}_2=\{0,1\}$. By (4.1.2),*

$$X_0 = |0\rangle\langle 0| + |1\rangle\langle 1| = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},\ X_1 = |1\rangle\langle 0| + |0\rangle\langle 1| = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{4.1.3}$$

$$Z_0 = 1\cdot|0\rangle\langle 0| + 1\cdot|1\rangle\langle 1| = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$Z_1 = 1\cdot|0\rangle\langle 0| + \exp\left(\frac{2\pi i}{2}\cdot 1\right)|1\rangle\langle 1| = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{4.1.4}$$

*Therefore*

$$\{X_0Z_0, X_1Z_0\} = \left\{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right\},\ \{X_0Z_0, X_1Z_1\} = \left\{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right\},$$

$$\{Z_0, Z_1\} = \left\{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right\}. \tag{4.1.5}$$

*The common eigenbases of the above three sets of matrices are*

$$\mathcal{F}_1 = \left\{\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\right\},$$

$$\mathcal{F}_2 = \left\{\frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle), \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle)\right\},\ \mathcal{F}_3 = \{|0\rangle, |1\rangle\}, \tag{4.1.6}$$

*respectively. This is the fundamental MUB of $\mathbb{C}^2$.*

### 4.1.2 The connection between MUBs and complementary subalgebras

**Definition 4.** *Let $\mathcal{A}_1, \mathcal{A}_2 \subset \mathbf{M}_n(\mathbb{C})$ be $*$-subalgebras with unit. $\mathcal{A}_1$ and $\mathcal{A}_2$ are called* complementary *(or* quasi-orthogonal*) if their traceless parts ($\mathcal{A}_1 \ominus \mathbb{C}I$ and $\mathcal{A}_2 \ominus \mathbb{C}I$) are orthogonal with respect to the Hilbert-Schmidt inner product.*

**Lemma 17.** *Let $\mathcal{F}_1, \dots, \mathcal{F}_{n+1} \subset \mathbb{C}^n$ be mutually unbiased bases and $\mathcal{A}_i$ be the subalgebra of the matrices that are diagonal in the basis $\mathcal{F}_i$ ($i \in \{1,\dots,n+1\}$). Then $\{\mathcal{A}_i\}_{i=1}^{n+1}$ are pairwise complementary maximal Abelian subalgebras (MASAs) that linearly span $\mathbf{M}_n(\mathbb{C})$.*

*Proof.* By definition, $\mathcal{A}_i$ is a maximal Abelian subalgebra. We have to prove the quasi-orthogonality. Set $1 \le k \ne l \le n+1$, $X \in \mathcal{A}_k \ominus \mathbb{C}I,\ Y \in \mathcal{A}_l \ominus \mathbb{C}I$, that is

$$X = \left(f_k^1, \dots, f_k^n\right)\mathrm{Diag}(x_1,\dots,x_n)\begin{pmatrix} (f_k^1)^* \\ \vdots \\ (f_k^n)^* \end{pmatrix},\ Y = \left(f_l^1, \dots, f_l^n\right)\mathrm{Diag}(y_1,\dots,y_n)\begin{pmatrix} (f_l^1)^* \\ \vdots \\ (f_l^n)^* \end{pmatrix}, \tag{4.1.7}$$

where $x_i, y_i \in \mathbb{C},\ \sum_{i=1}^n x_i = \sum_{i=1}^n y_i = 0$. Then

$$\mathrm{Tr} XY^* = \mathrm{Tr}\left( \left(f_k^1, \ldots, f_k^n\right) \mathrm{Diag}(x_1, \ldots, x_n) \begin{pmatrix} (f_k^1)^* \\ \vdots \\ (f_k^n)^* \end{pmatrix} \left(f_l^1, \ldots, f_l^n\right) \mathrm{Diag}(\overline{y}_1, \ldots, \overline{y}_n) \begin{pmatrix} (f_l^1)^* \\ \vdots \\ (f_l^n)^* \end{pmatrix} \right) =$$

$$= \mathrm{Tr}\left( \begin{pmatrix} x_1 (f_k^1)^* \\ \vdots \\ x_n (f_k^n)^* \end{pmatrix} \left(\overline{y}_1 f_l^1, \ldots, \overline{y}_n f_l^n\right) \begin{pmatrix} (f_l^1)^* \\ \vdots \\ (f_l^n)^* \end{pmatrix} \left(f_k^1, \ldots, f_k^n\right) \right) =$$

$$= \sum_{d=1}^n \sum_{c=1}^n x_d \overline{y}_c \underbrace{\left\langle f_k^d,\, f_l^c \right\rangle \left\langle f_l^c,\, f_k^d \right\rangle}_{\frac{1}{n}} = \frac{1}{n} \sum_{d=1}^n \sum_{c=1}^n x_d \overline{y}_c = \frac{1}{n} \left( \sum_{d=1}^n x_d \right) \left( \sum_{c=1}^n \overline{y}_c \right) = 0. \quad (4.1.8)$$

The last part of the satetement can be proved by counting dimensions. $\dim(\mathcal{A}_i) = n$, hence $\dim(\mathcal{A}_i \ominus \mathbb{C}I) = n-1$. $\mathbb{C}I, \mathcal{A}_1 \ominus \mathbb{C}I, \ldots, \mathcal{A}_{n+1} \ominus \mathbb{C}I$ are orthogonal subspaces, therefore

$$\dim(\mathrm{span}(\mathcal{A}_1, \ldots, \mathcal{A}_{n+1})) = \dim(\mathbb{C}I \oplus (\mathcal{A}_1 \ominus \mathbb{C}I) \oplus \cdots \oplus (\mathcal{A}_{n+1} \ominus \mathbb{C}I)) =$$

$$= 1 + (n+1)(n-1) = n^2 = \dim(\mathbf{M}_n(\mathbb{C})), \text{ hence } \mathrm{span}(\mathcal{A}_1, \ldots, \mathcal{A}_{n+1}) = \mathbf{M}_n(\mathbb{C}). \quad (4.1.9)$$

□

If $\mathcal{F}_i$ and $\mathcal{A}_i$ are the same as in Lemma 17 and $E_i:\ \mathbf{M}_n(\mathbb{C}) \to \mathbf{M}_n(\mathbb{C})$ is the orthogonal projection onto the subalgebra $\mathcal{A}_i$, then by the definition of Petz and Ohno [14], the map

$$\mathcal{E}: \mathbf{M}_n(\mathbb{C}) \to \mathbf{M}_n(\mathbb{C});\ A \mapsto \mathcal{E}(A) := \left(1 - \sum_{i=1}^{n+1} \lambda_i\right) \frac{\mathrm{Tr} A}{n} I + \sum_{i=1}^{n+1} \lambda_i E_i(A) \qquad (4.1.10)$$

is a generalized Pauli channel if the numbers $\lambda_i \in \mathbb{R}$ are chosen such that $\mathcal{E}$ is completely positive. $\mathcal{A}_1, \ldots, \mathcal{A}_{n+1}$ are pairwise complementary MASAs, therefore the above defined $\mathcal{E}$ is completely positive if and only if

$$1 + n\lambda_i \geq \sum_{j=1}^{n+1} \lambda_j \geq -\frac{1}{n-1}\ \forall\ i \in \{1, \ldots, n+1\}. \qquad (4.1.11)$$

(See [14].)

## 4.2 Channel directions and generalized angle parameters

**Definition 5** (Channel directions)**.** *The* channel directions *of the generalized Pauli channel given by the pairwise complementary subalgebras $\mathcal{A}_1, \ldots, \mathcal{A}_r$ are the affine subspaces*

$$\mathcal{D}_i = \{\sigma \in \mathcal{A}_i:\ \sigma^* = \sigma,\ \mathrm{Tr}(\sigma) = 1\} \subset \mathbf{M}_n(\mathbb{C}). \qquad (4.2.1)$$

**Unitary transformations of MUBs** Given a MUB $\mathcal{F}_1, \dots, \mathcal{F}_{n+1}$ and a unitary matrix $X$, the bases $X\mathcal{F}_1, \dots, X\mathcal{F}_{n+1}$ given by the definition $X\mathcal{F}_i := \{Xf_i^1, \dots, Xf_i^n\}$ are mutually unbiased bases as well. If $X \in SU(n)$ then there are

$$\alpha_1, \dots, \alpha_{n-1}, \theta_1, \dots, \theta_{\frac{n(n-1)}{2}}, \beta_1, \dots, \beta_{\frac{n(n-1)}{2}}$$

real parameters such that

$$X = D(\alpha_1, \dots, \alpha_{n-1}) \prod_{1 \le j < k \le n} U_{j,k}(\theta_{jk}, \beta_{jk}) \tag{4.2.2}$$

where

$$D(\alpha_1, \dots, \alpha_{n-1}) = \mathrm{Diag}\left(e^{i\alpha_1}, e^{i\alpha_2}, \dots, e^{i\alpha_{n-1}}, e^{-i\sum_{l=1}^{n-1}\alpha_l}\right), \tag{4.2.3}$$

and $U_{j,k}$ is the rotation of the subspace spanned by the standard basis vectors $|j-1\rangle$ and $|k-1\rangle$: its effect on the mentioned subspace is given by the matrix

$$\tilde{U}(\theta, \beta) = \begin{pmatrix} \cos\theta & -\sin\theta e^{-i\beta} \\ \sin\theta e^{i\beta} & \cos\theta \end{pmatrix} \tag{4.2.4}$$

and it equals with the identity on the orthocomplement (see [6]).

Now we can define *angle parameters* for Pauli channels given by MASAs.

**Definition 6.** *Let $n = p^M$ ($p$ is a prime, $M \in \mathbb{Z}^+$), $\mathcal{F}_1, \dots, \mathcal{F}_{n+1}$ be the fundamental MUB of $\mathbb{C}^n$ and $X$ be the unitary matrix determined by the parameters*

$$\alpha_1, \dots, \alpha_{n-1}, \theta_1, \dots, \theta_{\frac{n(n-1)}{2}}, \beta_1, \dots, \beta_{\frac{n(n-1)}{2}}.$$

*Let $\mathcal{A}_i$ be the subalgebra of the matrices that are diagonal in the basis $X\mathcal{F}_i$, $E_i$ be the orthoprojection onto $\mathcal{A}_i$. Assume that the numbers $\{\lambda_i\}_{i=1}^{n+1}$ satisfy the condition (4.1.11). Then the Pauli channel*

$$\mathcal{E} : \mathbf{M}_n(\mathbb{C}) \to \mathbf{M}_n(\mathbb{C});\ A \mapsto \mathcal{E}(A) := \left(1 - \sum_{i=1}^{n+1} \lambda_i\right) \frac{\mathrm{Tr}A}{n} I + \sum_{i=1}^{n+1} \lambda_i E_i(A) \tag{4.2.5}$$

*has the* generalized angle parameters $\left(\alpha_1, \dots, \alpha_{n-1}, \theta_1, \dots, \theta_{\frac{n(n-1)}{2}}, \beta_1, \dots, \beta_{\frac{n(n-1)}{2}}\right)$.

**Example 6.** *In the previous example we computed the fundamental MUB of $\mathbb{C}^2$ (4.1.6). It is easy to see that the corresponding subalgebras are the following:*

$$\mathcal{A}_1 = \mathrm{span}\{I, \sigma_1\},\ \mathcal{A}_2 = \mathrm{span}\{I, \sigma_2\},\ \mathcal{A}_3 = \mathrm{span}\{I, \sigma_3\}. \tag{4.2.6}$$

*If $\mathcal{E}$ is the Pauli channel with the generalized angle parameters $(\alpha, \theta, \beta)$, then the MUB of $\mathcal{E}$ is $\{X\mathcal{F}_1, X\mathcal{F}_2, X\mathcal{F}_3\}$, where*

$$X = X(\alpha, \theta, \beta) = \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta e^{-i\beta} \\ \sin\theta e^{i\beta} & \cos\theta \end{pmatrix}. \tag{4.2.7}$$

*Therefore, the corresponding subalgebras are*

$$\mathcal{A}'_1 = \mathrm{span}\{I, X\sigma_1 X^*\},\ \mathcal{A}'_2 = \mathrm{span}\{I, X\sigma_2 X^*\},\ \mathcal{A}'_3 = \mathrm{span}\{I, X\sigma_3 X^*\}. \tag{4.2.8}$$

*It follows that the channel matrix of $\mathcal{E}$ given by the equation $A = \rho^{-1} \circ \mathcal{E} \circ \rho$ ($\rho$ is the Bloch-parametrization) has the form*

$$A = \tilde{R}(\alpha, \theta, \beta)\mathrm{Diag}(\lambda_1, \lambda_2, \lambda_3)\tilde{R}^{-1}(\alpha, \theta, \beta), \tag{4.2.9}$$

*where*

$$\tilde{R}(\alpha, \theta, \beta) = \left\{ \left\langle \frac{1}{\sqrt{2}}\sigma_i,\ X(\alpha, \theta, \beta)\frac{1}{\sqrt{2}}\sigma_j X^*(\alpha, \theta, \beta) \right\rangle \right\}_{i,j=1}^{3}. \tag{4.2.10}$$

*The generalized angle parameters are not equal with the angle parameters introduced in Lemma 2 in general, for example*

$$\tilde{R}(\alpha, \theta, \beta) = \begin{pmatrix} \cos 2\alpha \cos^2\theta - \cos 2(\alpha - \beta)\sin^2\theta & \sin 2\alpha \cos^2\theta + \sin 2(\alpha - \beta)\sin^2\theta & \cos(2\alpha - \beta)\sin 2\theta \\ -\sin 2\alpha \cos^2\theta + \sin 2(\alpha - \beta)\sin^2\theta & \cos 2\alpha \cos^2\theta + \cos 2(\alpha - \beta)\sin^2\theta & -\sin(2\alpha - \beta)\sin 2\theta \\ -\cos\beta \sin 2\theta & -\sin\beta \sin 2\theta & \cos 2\theta \end{pmatrix}, \tag{4.2.11}$$

*but*

$$R_z(\phi_z)R_y(\phi_y)R_x(\phi_x) = \begin{pmatrix} \cos\phi_z \cos\phi_y & -\cos\phi_z \sin\phi_y \sin\phi_x - \sin\phi_z \cos\phi_x & -\cos\phi_z \sin\phi_y \cos\phi_x + \sin\phi_z \sin\phi_x \\ \sin\phi_z \cos\phi_y & -\sin\phi_z \sin\phi_y \sin\phi_x + \cos\phi_z \cos\phi_x & -\sin\phi_z \sin\phi_y \cos\phi_x - \cos\phi_z \sin\phi_x \\ \sin\phi_y & \cos\phi_y \sin\phi_x & \cos\phi_y \cos\phi_x \end{pmatrix}, \tag{4.2.12}$$

*where $R_z, R_y, R_x$ are the rotations defined in Lemma 2. However, these parametrizations show some analogies in some cases. For example,*

$$\tilde{R}(\alpha, 0, 0) = \begin{pmatrix} \cos 2\alpha & \sin 2\alpha & 0 \\ -\sin 2\alpha & \cos 2\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \equiv R_z(-2\alpha), \tag{4.2.13}$$

*and*

$$\tilde{R}(0, \theta, 0) = \begin{pmatrix} \cos 2\theta & 0 & \sin 2\theta \\ 0 & 1 & 0 \\ -\sin 2\theta & 0 & \cos 2\theta \end{pmatrix} \equiv R_y(-2\theta). \tag{4.2.14}$$

# Chapter 5

# Conclusion and future work

In this thesis we investigated the tomography of quantum Pauli channels which form a wide class of the state transformations.

In Chapter 3 we considered the Pauli channels acting on two-level quantum systems. For quantum bit Pauli channels we defined the *channel directions* rigorously and introduced the *channel matrix* that describes the effect of the channel. By a parametrization of the rotation group $\mathbf{O}(3, \mathbb{R})$ we defined angle parameters that describe the channel dirctions.

We developed a tomography scheme including efficient estimations of the channel matrix, the contraction parameters and the angle parameters. The accuracy of an estimation method can be measured with various quantities. In this thesis we considered the *mean squared error* of the estimated channel matrix, contraction parameters and angle parameters. We constructed optimal tomography settings with respect to the above mentioned quantities by minimizing the corresponding loss functions.

We proved that the estimation of the channel matrix is optimal if we let the input and measurement directions be the same as the channel directions. A similar result appears if the aim is the efficient estimation of the contraction parameters. The proofs are analytical, this is a new concept in the study of the Pauli channels with *unknown* channel directions (see [1, 16]).

In the general case we did not succed to minimize the mean squared error of the estimated angle parameters, hence we formulate conjectures about the optimal settings based on numerical optimization and empirical simulations. However, if we assume some a priori knowledge about the Pauli channel, we can find the optimal settings to estimate the angle parameter analytically.

In Chapter 4 we investigated the *generalized* Pauli channels that act on arbitrary finite-dimensional quantum systems. We defined the channel directions and introduced *angle parametrization* for a wide class of generalized Pauli channels.

These are the first steps of an interesting further work that is nothing else but the generalization of the tomography scheme developed in this thesis for $n$-level quantum systems.

## Acknowledgement

I would like to thank Professor Dénes Petz for great conversations and especially Professor Katalin Hangos for her support and help in all areas of the scientific work. I would like to thank László Ruppert for introducing me in this topic and for his great advices.

# Bibliography

[1] G. Balló, K. M. Hangos and D. Petz. Convex Optimization-Based Parameter Estimation and Experiment Design for Pauli Channels. *IEEE Trans. on Automatic Control*, 99, 2012.

[2] A. Bendersky, F. Patawski, and J.P. Paz. Selective and efficient estimation of parameters for quantum process tomography. *Phys. Rev. Lett.* **100**, 190403, 2008.

[3] I. Bengtsson. Three ways to look at mutually unbiased bases. *arXiv:quant-ph/0610216*, 2006.

[4] M.P.A. Branderhorst, J. Nunn, I.A. Walmsley, and R.L. Kosut. Simplified quantum process tomography. *New J. Phys.* **11**, 115010, 2009.

[5] A. Chiuri, V. Rosati, G. Vallone, S. Padua, H. Imai, S. Giacomini, C. Macchiavello, and P. Mataloni. Experimental realization of optimal noise estimation for a general Pauli channel. *Phys. Rev. Lett.*, **107**, 253602, 2011.

[6] D. D'Alessandro. *Introduction to Quantum Control and Dynamics.* Applied Mathematics and Nonlinear Science. Chapman and Hall/CRC, 2008.

[7] G. M. D'Ariano, M. G. A. Paris, and M. F. Sacchi. Quantum tomography. *Advances in Imaging and Electron Physics*, 128:205, 2003.

[8] A. Fujiwara and H. Imai. Quantum parameter estimation of a generalized Pauli channel. *J. Phys. A*, **36**, 8093–8103,

[9] F. Hiai and D. Petz. *Introduction to Matrix Analysis and Applications.* To appear in Hindustan Book Agency.

[10] Z. Ji, G. Wang, R. Duan, Y. Feng, and M. Ying. Parameter estimation of quantum channels. *IEEE Trans. Inf. Theory*, 54:5172–5185, 2008.

[11] M. Mohseni, A. T. Rezakhani, and D. A. Lidar. Quantum process tomography: Resource analysis of different strategies. *Physical Review A*, 77:032322, 2008.

[12] M. Nathanson and M. B. Ruskai. Pauli diagonal channels constant on axes. *J. Phys. A*, **40**, 8171-8204, 2007.

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

[14] D. Petz and H. Ohno. Generalizations of Pauli channels. *Acta Math. Hungar.*, 124:165-177, 2009.

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

[16] L. Ruppert, D. Virosztek and K. M. Hangos. Optimal parameter estimation of Pauli channels. *J. Phys. A: Math. Theor.*, 45:265305, 2012.

[17] M. Sasaki, M. Ban, and S. M. Barnett. Optimal parameter estimation of a depolarizing channel. *Phys. Rev. A*, **66**(2), 022308, 2002.

[18] D. Virosztek, L. Ruppert and K. M. Hangos. Pauli channel tomography with unknown channel directions. *arXiv:quant-ph/1304.4492*, 2013.

[19] Wolfram Research, Inc., Champaign, Illinois. *Mathematica Edition: Version 8.0.* 2010.

[20] 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*, **79**(6), 062301, 2009.

# Appendix

## Empirical simulations

In the subsection 3.3.3 we did not succeed in minimizing the mean squared error of the estimated angle parameters analytically. Therefore we performed empirical simulations to formulate conjectures based on them. (On the other hand, empirical simulations are useful to check the analytical results.)

The parameter estimation method for qubit Pauli channel described in this work can be performed by computer, because

1. one can generate the random variables $N_{ij}^{+}$ (see (3.1.6)) with the built-in packages of symbolic mathematical programming languages,
2. with given $N_{ij}^{+}$-s, one gets the parameter estimations by direct computations.

Performing the parameter estimation method $K$ times, we get the estimations

$$\hat{\lambda}_1^j, \hat{\lambda}_2^j, \hat{\lambda}_3^j, \hat{\phi}_z^j, \hat{\phi}_y^j, \hat{\phi}_x^j \ (j \in \{1 \dots K\}). \tag{5.0.1}$$

Let us define the quantities

$$\hat{f}_1 = \frac{1}{K} \sum_{j=1}^{K} \mathrm{dist}(\hat{\phi}_z^j, \phi_z)^2 + \mathrm{dist}(\hat{\phi}_y^j, \phi_y)^2 + \mathrm{dist}(\hat{\phi}_x^j, \phi_x)^2, \tag{5.0.2}$$

$$\hat{f}_2 = \frac{1}{K} \sum_{j=1}^{K} (\hat{\lambda}_1^j - \lambda_1)^2 + (\hat{\lambda}_2^j - \lambda_2)^2 + (\hat{\lambda}_3^j - \lambda_3)^2, \tag{5.0.3}$$

$$\hat{f}_3 = \frac{1}{K} \sum_{j=1}^{K} ||A(\hat{\lambda}_1^j, \hat{\lambda}_2^j, \hat{\lambda}_3^j, \hat{\phi}_z^j, \hat{\phi}_y^j, \hat{\phi}_x^j) - A(\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x)||^2. \tag{5.0.4}$$

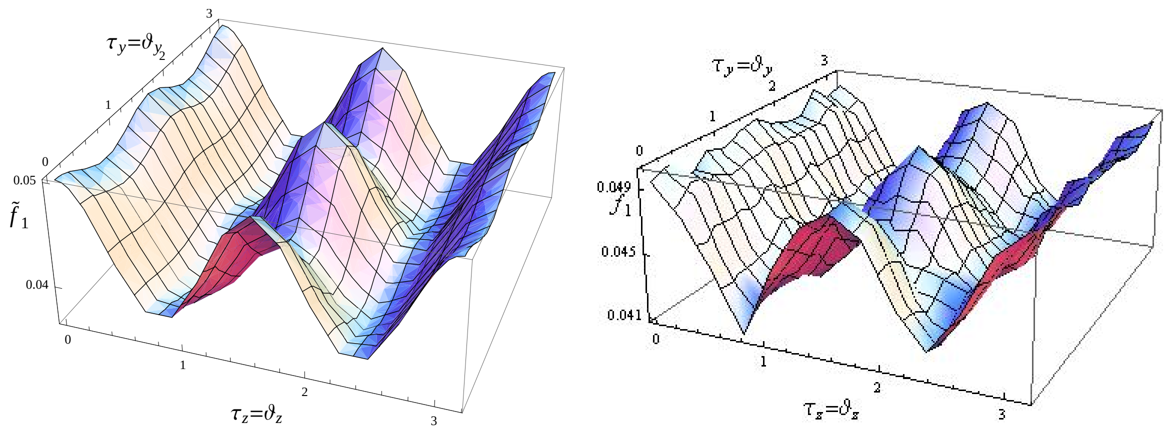
This is nothing else but calculating the *sample means* instead of the expected values that define the loss functions $f_1$, $f_2$ and $f_3$ (see (3.2.1), (3.2.2) and (3.2.3)).

It is a well-known property of the sample mean that

$$\mathbf{E}\hat{f}_i = f_i \text{ and } \mathbf{Var}\left(\hat{f}_i\right) = \underline{\underline{O}}\left(\frac{1}{K}\right) \ (\forall\, i \in \{1, 2, 3\}). \tag{5.0.5}$$

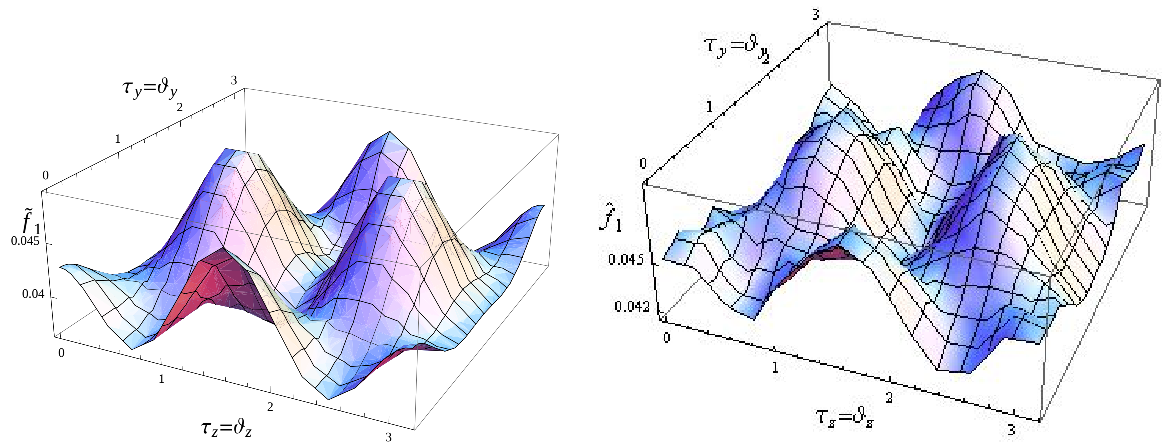
Therefore, $\hat{f}_i$ is a good approximation of $f_i$ if $K$ is large enough.

The empirical simulations were performed by *Mathematica 8* [19]. The following figures are to make the empirical results picturesque.




$(a)$ $(b)$

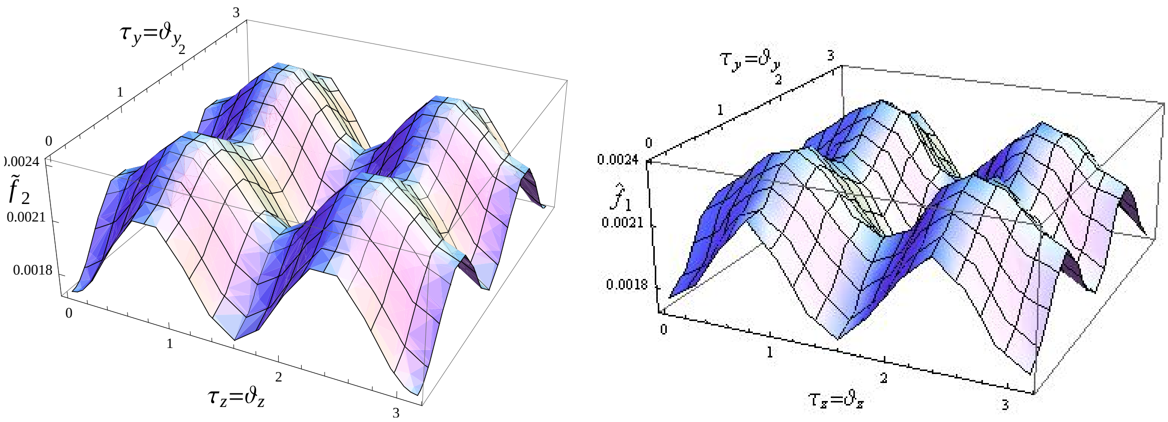
Figure 5.1: (a) $\tilde{f}_1$ and (b) $\hat{f}_1$. $\vartheta_x = \tau_x = 0,\ 0 \leq \vartheta_z = \tau_z, \vartheta_y = \tau_y \leq \pi$




$(a)$ $(b)$

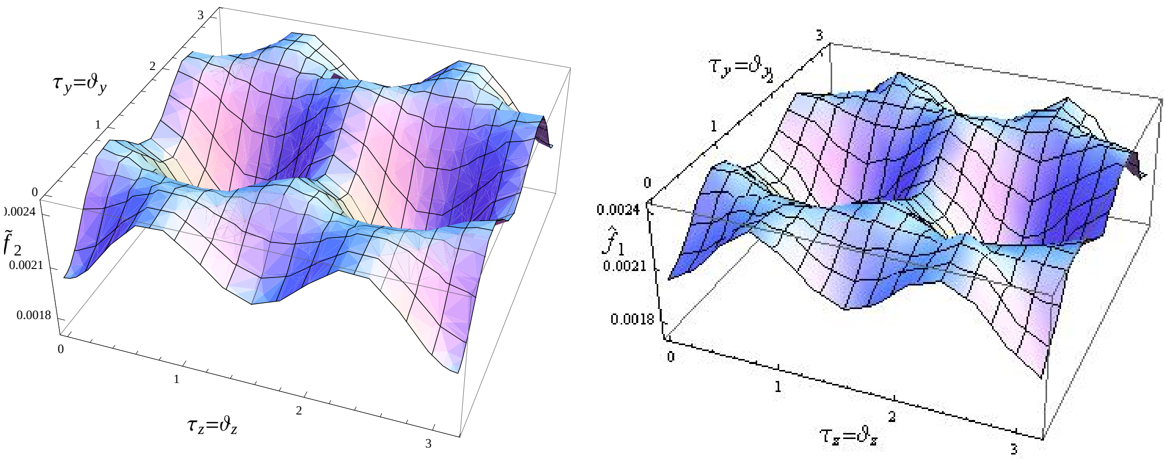
Figure 5.2: (a) $\tilde{f}_1$ and (b) $\hat{f}_1$. $\vartheta_x = \tau_x = \frac{\pi}{4},\ 0 \leq \vartheta_z = \tau_z, \vartheta_y = \tau_y \leq \pi$.

Set $\lambda_1 = 0.8, \lambda_2 = 0.65, \lambda_3 = 0.5, N = 1000$ and $\underline{\tau} = \underline{\vartheta}$. Then we can plot the functions $\tilde{f}_1$ and $\hat{f}_1$ with fixed parameters $\vartheta_x = \tau_x = 0$ and $\vartheta_x = \tau_x = \frac{\pi}{4}$ (Figure 5.1 and Figure 5.2, respectively). To check the analytical results, one may plot the functions $\tilde{f}_2$ and $\hat{f}_2$ with fixed $\vartheta_x = \tau_x = 0$ and $\vartheta_x = \tau_x = \frac{\pi}{4}$ (Figure 5.3 and Figure 5.4).




$(a)$ $(b)$

Figure 5.3: (a) $\tilde{f}_2$ and (b) $\hat{f}_2$. $\vartheta_x = \tau_x = 0,\ 0 \leq \vartheta_z = \tau_z, \vartheta_y = \tau_y \leq \pi$




$(a)$ $(b)$

Figure 5.4: (a) $\tilde{f}_2$ and (b) $\hat{f}_2$. $\vartheta_x = \tau_x = \frac{\pi}{4},\ 0 \leq \vartheta_z = \tau_z, \vartheta_y = \tau_y \leq \pi$.