Chapter 1
Optimal approximation of non-physical maps via maximum likelihood estimation

Vladimír Bužek, Mário Ziman, and **Martin Plesch**
Research Center for Quantum Information, SAS, Dúbravská cesta 9, 845 11 Bratislava, Slovakia
and
QUNIVERSE, Líščie Údolie 116, 841 04 Bratislava, Slovakia

1.1 Introduction

Any quantum dynamics [1,2], i.e., the process that is described by a completely positive (CP) map of a quantum-mechanical system, can be probed in two different ways: Either we use a single entangled state of a bi-partite system [3], or we use a collection of linearly independent single-particle test states [4,5] (forming a basis of the vector space of all Hermitian operators). Given the fragility of entangled states in this Chapter we will focus our attention on the process reconstruction using only single-particle states.

The task of a process reconstruction is to determine an unknown quantum channel (a "black box") using correlations between known input states and results of measurements performed on states that have been transformed by the channel (see Fig. 1.1). The linearity of quantum dynamics implies that the channel $\mathcal{E}$ is exhaustively described by its action $\varrho_j \to \varrho'_j = \mathcal{E}[\varrho_j]$ on a set of basis states, i.e., a collection of linearly independent states ϱ_j, that play a role of *test states*. Therefore, to perform a reconstruction of the channel $\mathcal{E}$ we have to perform a complete state tomography [1,6,7] of ϱ'_j. The number of test states equals d^2, where $d = \dim \mathcal{H}$ is the dimension of the Hilbert space $\mathcal{H}$ associated with the system. Consequently, in order to reconstruct a channel we have to determine $d^2(d^2-1)$ real parameters, i.e., 12 numbers in the case of qubit ($d = 2$).

In what follows we will assume that test states can be prepared on demand perfectly. Nevertheless, the reconstruction of the channel $\mathcal{E}$ can be affected by the

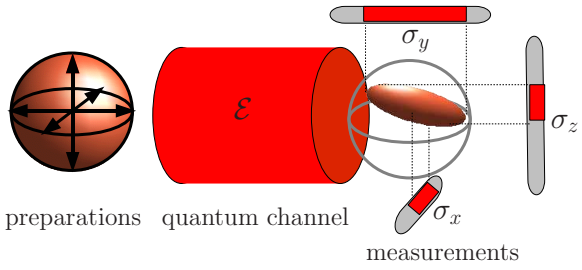

Figure 1.1 A schematic representation of a reconstruction of a single-qubit channel. Input (test) states ϱ_k of the single-qubit channels are represented by the Bloch sphere (the state space of a single qubit). At the output of the single-qubit channel (modelled as a transformation of the Bloch sphere into an ellipsoid, i.e., the Bloch sphere is "deformed" by the action of the channel) a complete measurement of test states is performed. The complete measurement is performed via the projective measurement of σ operators. Based on correlations between input and output states of the test qubits the action of the quantum channel (a CP map) is determined (reconstructed).

lack of required information, as each test state is represented by a finite ensemble of identically prepared test particles (e.g., qubits). Correspondingly, measurements performed at the output can result in an approximate estimation of transformed test states. This situation is typical for experiments - one cannot prepare an infinite ensemble of identically prepared particles, so frequencies of the measured outcomes are only approximations of probability distributions. Consequently, a straightforward reconstruction of output states $\varrho_j^{(out)}$ might lead to non-physical conclusions about the action of the quantum channel. As a result we can find a negative operator $\varrho_j^{(out)}$, or a channel $\mathcal{E}$, which is not completely positive (CP) [1,2].

We will recall the method of *maximum likelihood* (MML) to perform an estimation of an unknown channel [8]. We will use this method to perform a reconstruction based on numerical simulation of the anti-unitary *Universal-NOT gate* (the spin-flip operation) [9–14]. This is a linear, but not a CP map and we will show how our estimation will result in the optimal physical approximations of the U-NOT operation. In order to demonstrate the power of this approach we will also apply it to obtain an approximation of non-linear quantum-mechanical maps: the so called *non-linear polarization rotation* (NPR) [15] and a highly non-linear transformation $\varrho \longrightarrow \varrho^2$.

Having the result of the MML method, one may always analyze the question how reliable this result is. As any numerical method, MML may fail by falling to (or rather climbing up to) a non-local maxima. On the other hand, data provided by experiment may lead to non-physical results. In this case one shall not wonder that any physical approximation based on inconsistent data will result in an inconsistent map, e.g. a map that does not reproduce the experimental data perfectly. To deal with such cases, we will examine the resulting value of the MML functional. We will show that in certain cases it may be used as an indicator of whether the

whole reconstruction scheme is consistent (e.g., whether we have a proper a priori knowledge about input test states, etc.).

This Chapter is organized as follows: In Sec. 2 we present some basic facts about quantum states and quantum channels. This should help a reader who is not familiar with the problem of quantum state and process estimation (reconstruction) to understand some technical details that are presented later in the Chapter. Then, in Sec. 3 we describe properties of single-qubit channels. In Sec. 4 we briefly introduce the method of maximum likelihood. In Sec. 5 we apply the MML for an estimation of the Universal-NOT gate, while in Sec. 6 we will present two physical approximations of non-linear maps, namely, the non-linear polarization rotation and the map $\varrho \longrightarrow \varrho^2$. In Sec. 7 we will examine how to use the value of the MML functional to verify consistency of the reconstruction scheme - namely, whether the prior knowledge about the input states is correct. In Sec. 8 we will summarize our results.

1.2 Quantum states and quantum channels

1.2.1 States of quantum systems and density operators

In quantum theory a state of a physical system is described by a state vector or a density operator ϱ that acts on a Hilbert space $\mathcal{H}$ (of a specific dimension d) associated with the system. Density operators are mathematical objects that allow us to describe states of physical systems that have not been prepared in a unique way. Specifically, a state of a quantum system that has been prepared in a unique way is described by a state vector $|\Psi\rangle$. One can associate with the state vector $|\Psi\rangle$ a density operator (projector) $\varrho = |\Psi\rangle\langle\Psi|$. In many situations though the preparation is not under full control and the system is prepared in one state from a specific set of states $\{|\Psi_j\rangle\langle\Psi_j|\}$ with corresponding probabilities p_j ($p_j \geq 0$ and $\sum_j p_j = 1$). Such a "mixed" states of a quantum system is described by a density operator

$$\varrho = \sum_j p_j |\Psi_j\rangle\langle\Psi_j|. \tag{1.1}$$

The density operator has to have non-negative eigenvalues and its trace has to be equal to unity, i.e., $\mathrm{Tr}\varrho = 1$. The purity of the density operator can be quantified in terms of a von Neumann entropy

$$S = -\mathrm{Tr}\left(\varrho \ln \varrho\right). \tag{1.2}$$

If we assume that the state vectors $|\Psi_j\rangle$ are mutually orthogonal ($|\Psi_j\rangle$ are basis vectors of the Hilbert space $\mathcal{H}$ such that $\langle\Psi_k|\Psi_j\rangle = \delta_{j,k}$), then the eigenvalues of the density operator (1.1) are equal to probabilities p_j and the von Neumann entropy (1.2) reads

$$S = -\sum_{j=0}^{d} p_j \ln p_j. \tag{1.3}$$

It is clear now that if the system has been prepared in a unique (pure) state $|\Psi_k\rangle$, i.e., $p_k = \delta_{j,k}$, then the entropy S is equal to zero. On the other hand, if the system is prepared in a mixture of mutually orthogonal states $|\Psi_j\rangle$ (basis vectors of d-dimensional Hilbert space $\mathcal{H}$) with the equal probability, i.e., $p_j = 1/d$, then the von Neumann entropy of the corresponding (total) mixture achieves its maximum value $S = \ln d$.

Alternatively a concept of a mixed state can appear when our quantum system is considered to be quantum-mechanically correlated with another quantum system. To be more specific, let us consider two quantum systems A and B with Hilbert spaces $\mathcal{H}_A$ ($\dim\mathcal{H}_A = d_A$) and $\mathcal{H}_B$ ($\dim\mathcal{H}_B = d_B$), respectively. The two Hilbert spaces are spanned by basis vectors $|\Psi_k^{(A)}\rangle$ and $|\Psi_k^{(B)}\rangle$, respectively. For simplicity, let us assume that $d_A = d_B$. Now we consider a pure bi-partite state $|\Phi\rangle_{AB}$ ($S_{AB} = 0$). If this pure state can be expressed in a factorized form, i.e., as a product of two state vectors

$$|\Phi\rangle_{AB} = |\phi^{(A)}\rangle_A \otimes |\phi^{(B)}\rangle_B, \tag{1.4}$$

then the two systems under consideration are each in a pure state. The two systems are not correlated, which can be seen by performing a partial trace over one of the sub-system (let say, B) and from the state vector $|\Phi\rangle_{AB}$ we obtain a pure state $|\phi^{(A)}\rangle_A$ ($S_A = 0$). On the other hand, if the bi-partite state $|\Phi\rangle_{AB}$ cannot be expressed in a factorized form (1.4) then we say, the the two sub-systems are mutually entangled (quantum-mechanically correlated). Any such state can be expressed in a form of the so-call Schmidt decomposition

$$|\Phi\rangle_{AB} = \sum_j \lambda_j |\xi_j^{(A)}\rangle_A |\xi_j^{(B)}\rangle_B, \tag{1.5}$$

where λ_j are non-negative real numbers satisfying the normalization condition $\sum_j \lambda_j^2 = 1$ and $|\xi_j^{(A)}\rangle_A$ and $|\xi_j^{(B)}\rangle_B$ are orthonormal states from $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. From the Schmidt decomposition it follows that states of the two sub-systems are described the density operators

$$\varrho_{(A)} = \sum_j \lambda_j^2 |\xi_j^{(A)}\rangle\langle\xi_j^{(A)}| \,; \quad \varrho_{(B)} = \sum_j \lambda_j^2 |\xi_j^{(B)}\rangle\langle\xi_j^{(B)}| \,, \tag{1.6}$$

from where we see that eigenvalues of the density operators $\varrho_{(A)}$ and $\varrho_{(B)}$ are equal to λ_j^2. Consequently, von Neumann entropies of both sub-systems are identical and equal to $S_A = S_B = -\sum_j \lambda_j^2 \, \ln \lambda_j^2$.

1.2.2 State tomography

A state of a quantum system can be completely reconstructed when meanvalues of all system observables (the so-called "quorum") are known from a measurement. A typical example would be a state of a qubit - a spin-1/2 particle. This systems has

a 2-dimensional Hilbert space and in general its state can be described by a density operator

$$\varrho = \frac{1}{2}\left(I + \vec{n}\cdot\vec{\sigma}\right), \tag{1.7}$$

where $\vec{\sigma} = \{\sigma_x, \sigma_y, \sigma_z\}$ with $\sigma_{x,y,z}$ being Pauli matrices. Components of the vector $\vec{n} = \{n_x, n_y, n_z\}$ are the meanvalues of the Pauli operators in the state ϱ, e.g., $n_x = \mathrm{Tr}(\varrho\sigma_x)$. These meanvalues can be determined when probabilities $p_\pm$ on the eigenvectors $|\pm x\rangle$ of the operator σ_x (we have $\sigma_x|\pm x\rangle = \pm|\pm x\rangle$) are obtained from a measurement. In order to obtain a probability distributions $p_\pm = \langle\pm x|\varrho|\pm x\rangle$ a measurement on a sufficiently large (infinite) ensemble of identically prepared qubits has to be performed. Once the measurement of all observables belonging to the quorum is done the complete tomography (reconstruction) of the state of a qubit can be performed.

It may happen that for one or other reason meanvalues of just a subset of observables from the quorum are known from the measurement. In this case a complete tomography of the state cannot be performed because of missing information. Nevertheless, a partial reconstruction (an estimation) of the measured state can be performed. One of possible approaches in this case is the state-reconstruction based on the principle of Maximum Entropy as introduced by E.T. Jaynes (see, e.g., Ref. [7] and references therein). This method works as follows: Let us assume that a subset $\{G_\nu\}$ of system operators has been measured in the experiment, and that the corresponding mean values are

$$\langle G_\nu\rangle = \mathrm{Tr}\left(\varrho G_\nu\right). \tag{1.8}$$

The question is how to select from a set of all density operators ϱ_G that satisfy the condition (1.8) the one that is considered to be the "best" estimation of the true density operator. Jaynes has suggested that among all density operators ϱ_G that satisfy the condition (1.8) one should select that operator ϱ_{est} that maximizes the von Neumann entropy.

In real physical situation it happens that infinite ensembles of identically prepared particles are not available. Therefore, strictly speaking the reconstruction methods that are based on the "perfect" knowledge of probability distributions are not applicable. In these cases, when only finite set of registered "clicks" are obtained from measurements one should use more appropriate methods. In particular, the method based on Bayesian inversion (for details see Ref. [16], and references therein) results in a reliable estimation of states of quantum systems. Alternatively, the method of Maximum Likelihood can be employed (see, e.g., Ref. [17] and references therein).

1.2.3 Quantum processes

If a quantum system is isolated from its environment, then it evolves unitarily, its dynamics is described by a unitary operator U, and the density operator is trans-

formed in a usual way $\varrho^{(out)} = U\varrho U^\dagger$. On the other hand if the system interacts with its environment, then its dynamics is described by a transformation

$$\varrho^{(out)} = \mathcal{E}[\varrho]\,, \tag{1.9}$$

where $\mathcal{E}$ is a map that fulfills three conditions:

- The output density operator $\varrho^{(out)}$ has to describe a physical state. $\mathrm{Tr}\left(\mathcal{E}[\varrho]\right)$ is the probability that the process represented by $\mathcal{E}$ occurs when the system is originally in the state ϱ. Therefore $0 \leq \mathrm{Tr}\left(\mathcal{E}[\varrho]\right) \leq 1$ for all input states ϱ. That is, the map $\mathcal{E}$ is positive.

- $\mathcal{E}$ is a convex-linear map on a set of density operators. Therefore, if the initial state ϱ is expressed as a mixture $\varrho = \sum_j p_j \varrho_j$ (such that $p_j \geq 0$ and $\sum_j p_j = 1$), then
$$\mathcal{E}\left[\sum_j p_j \varrho_j\right] = \sum_j p_j \mathcal{E}[\varrho_j]\,. \tag{1.10}$$

- The map $\mathcal{E}$ has to complete positive.

In order to understand the condition of complete positivity, let us consider two systems A and B. The state of this composite system is described by a density operator ϱ_{AB}. Let us assume that this state is transformed in a following way: the sub-system A undergoes a transformation described by the map $\mathcal{E}_A$ while the sub-system B is left unchanged, that is, we have

$$\varrho_{AB}^{(out)} = \left(\mathcal{E}_A \otimes I_B\right)\left[\varrho_{AB}\right]\,, \tag{1.11}$$

where I_B is the identity operator on the system B. The condition of complete positivity requires that the operator $\varrho_{AB}^{(out)}$ is positive for any input state ϱ_{AB}.

We note, that not all positive maps $\mathcal{E}$ are also completely positive. A typical example of a positive but not completely positive operator is the so called Universal-NOT gate $\mathcal{E}_{\mathtt{NOT}}$, which is an anti-unitary transformation, that "flips" an input state of a qubit $|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle$ into its antipode $|\Psi^\perp\rangle = \beta^*|0\rangle - \alpha^*|1\rangle$, such that $\langle\Psi|\Psi^\perp\rangle = 0$ for all possible input states $|\Psi\rangle$. It is easy to show the operation $\mathcal{E}_{\mathtt{NOT}}$ that is supposed to act on a single qubit state as

$$|\Psi\rangle\langle\Psi| \to \mathcal{E}_{\mathtt{NOT}}\left[|\Psi\rangle\langle\Psi|\right] = |\Psi^\perp\rangle\langle\Psi^\perp| \tag{1.12}$$

is positive, since both the input and the output of the channel (1.12) are positive density operators. In order to show that the operation is not completely positive let us consider the action of the transformation $\mathcal{E}_{\mathtt{NOT}} \otimes I$ on the singlet state $|\Xi\rangle_{AB}$ which we defined as

$$|\Xi\rangle_{AB} = \frac{1}{\sqrt{2}}\left(|\Psi\rangle_A|\Psi^\perp\rangle_B - |\Psi^\perp\rangle_A|\Psi\rangle_B\right). \tag{1.13}$$

This is a maximally entangled state which is $SU(2) \otimes SU(2)$ invariant and it violates Bell inequalities (for more details see Ref. [1]). The operator $\varrho_{AB} = |\Xi\rangle\langle\Xi|$ has eigenvalues $\lambda = \{1, 0, 0, 0\}$. The output state $\varrho_{AB}^{(out)}$ of the action of the operator $\mathcal{E}_{\mathtt{NOT}} \otimes I$ on the singlet state (1.13) has the eigenvalues $\lambda = \{-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\}$. From here it follows that the operator $\varrho_{AB}^{(out)}$ has the trace equal to unity, but it does not describe a physical state since its eigenvalues are negative. From above it follows that the map $\mathcal{E}_{\mathtt{NOT}}$ is positive but not complete positive (see also Sec. 5).

Every map $\mathcal{E}$ satisfies the three conditions specified above if and only if it can be expressed as

$$\mathcal{E}[\varrho] = \sum_j E_j \varrho E_j^\dagger \tag{1.14}$$

via the set of operators $\{E_j\}$ that map the input Hilbert space to the output Hilbert space and that satisfy the condition

$$\sum_j E_j^\dagger E_j = I. \tag{1.15}$$

1.2.4 Processes tomography

In order to determine the channel $\mathcal{E}$ we have to find out the operators E_j that define this map via the decomposition Eq. (1.14). In order to determine operators E_j it is useful to express them via a set of "fixed" operators $\tilde{E}_k$ which form a basis for the set of all operators on the state space. Therefore one can express arbitrary E_j as a superposition of basis operators $\tilde{E}_k$:

$$E_j = \sum_{k=0}^{d^2-1} e_{jk} \tilde{E}_k \tag{1.16}$$

with e_{jk} being complex numbers. Now the channel $\mathcal{E}$ can be expressed as

$$\mathcal{E}[\varrho] = \sum_{mn=0}^{d^2-1} \chi_{mn} \tilde{E}_m \varrho \tilde{E}_n^\dagger \,, \tag{1.17}$$

where by construction the matrix $\chi_{mn} = \sum_j e_{jm} e_{nj}^*$ us positive Hermitian. This matrix completely describes the channel $\mathcal{E}$ (once the set of basis operators $\tilde{E}_j$ is fixed). In general the matrix χ contains $d^2(d^2-1)$ independent real parameters. To see this we just remind ourselves that $\mathcal{E}$ implements transformation between $d \times d$ complex matrices (input and output states of the quantum system under consideration). Therefore χ is a $d^2 \times d^2$ matrix. If we take into account that in- and out states are described by Hermitian operators with the trace equal to unity [which results into d^2 constraints that are expressed by the completeness relation (1.15)] then we see that the matrix χ is specified by $d^2(d^2-1)$ real parameters. Complete reconstruction of a quantum channel is equivalent to determination of these parameters.

For this we need to have a capacity to prepare input test states and after the action of the channel to perform complete reconstruction of the output states.

From above it follows that in order to perform a complete reconstruction of an unknown CP map $\mathcal{E}$ we need a set d^2 linearly independent test states $\{\varrho_k\}_{k=0}^{d^2-1}$ that are transformed by the channel. The states $\varrho_k^{(out)} = \mathcal{E}[\varrho_k]$ are supposed to be reconstructed (completely) and out of the correlations between the in and out states, ϱ_k and $\varrho_k^{(out)}$, respectively, the map (channel) can be reconstructed.

Since the map $\mathcal{E}$ is linear it can be reconstructed by the following inversion procedure: Let the set of operators $\{\varrho_k\}_{k=0}^{d^2-1}$ form a linearly independent basis for the space of $d \times d$ matrices (density operators). This means that any $d \times d$ matrix can be written in a form of a linear combination of the basis operators ϱ_k. Usually (see, e.g., Ref. [1]) these basis operators are taken in the form $\varrho_k = |k_1\rangle\langle k_2|$, where $\{k_j\}_{j=0}^{d-1}$ is an orthonormal basis set in the Hilbert space $\mathcal{H}$ of the system under consideration*. We note that the transformation of the operator $\varrho_k = |k_1\rangle\langle k_2|$ can be determined via performing preparation and measurement of four projectors, namely:

$$|k_1\rangle\langle k_2| = |+\rangle\langle +| + i|-\rangle\langle -| - \frac{1+i}{2}\left(|k_1\rangle\langle k_1| + |k_2\rangle\langle k_2|\right) , \tag{1.18}$$

where $|+\rangle = (|k_1\rangle + |k_2\rangle)/\sqrt{2}$ and $|-\rangle = (|k_1\rangle + i|k_2\rangle)/\sqrt{2}$, respectively. The linearity of the map $\mathcal{E}$ implies that it is sufficient to determine the output states $\mathcal{E}[|+\rangle\langle +|]$, $\mathcal{E}[|-\rangle\langle -|]$, $\mathcal{E}[|k_1\rangle\langle k_1|]$, and $\mathcal{E}[|k_2\rangle\langle k_2|]$ in order to determine the transformation on the basis operators ϱ_k, i.e.,

$$\mathcal{E}[|k_1\rangle\langle k_2|] = \mathcal{E}[|+\rangle\langle +|] + i\mathcal{E}[|-\rangle\langle -|] - \frac{1+i}{2}\left(\mathcal{E}[|k_1\rangle\langle k_2|] + \mathcal{E}[|k_1\rangle\langle k_2|]\right) . \tag{1.19}$$

Moreover, because the operators ϱ_k form a basis, we have

$$\mathcal{E}[\varrho_k] = \sum_{l=0}^{d^2-1} \lambda_{kl}\varrho_l , \tag{1.20}$$

where the parameters λ_{kl} are determined by the measurement results in the operator basis, i.e.,

$$\lambda_{kl} = \mathrm{Tr}\left(\tilde{E}_k\, \mathcal{E}[\varrho_l]\right) , \tag{1.21}$$

where $\tilde{E}_k = \varrho_k$. This is a usual choice of the basis operators $\tilde{E}_k$ because ϱ_k are Hermitian operators, so they can play a role of genuine observables. To proceed we use the expression

$$\tilde{E}_m\varrho_k\tilde{E}_n^\dagger = \sum \beta_{mn,kl}\varrho_l , \tag{1.22}$$

*We assume that the number k is represented by a pair $\{k_1, k_2\}$ such that $k = 0, \ldots, d^2-1$. We can think of k being expressed in a d-nary form via k_1k_2.

where $\beta_{mn,kl}$ are complex numbers that are determined from a simple algebraic relation between $\tilde{E}_m$ and ϱ_k. If $\tilde{E}_m = \varrho_m$ then given the fact that $\varrho_x = |x_1\rangle\langle x_2|$ ($x = l, m, n, k$ with $0 \leq x \leq d^2 - 1$) we find

$$\beta_{mn,kl} = \delta_{l_1,m_1}\delta_{n_1,l_2}\delta_{m_2,k_1}\delta_{k_2,n_2} \,. \tag{1.23}$$

When we combine Eqs. (1.21) and (1.22) we obtain the equation

$$\sum_{k=0}^{d^2-1}\sum_{m,n=0}^{d^2-1} \chi_{mn}\beta_{mn,kl}\varrho_k = \sum_{j=0}^{d^2-1} \lambda_{lj}\varrho_j \,. \tag{1.24}$$

From the linear independence of the operators ϱ_k it follows that for each k

$$\sum_{m,n=0}^{d^2-1} \chi_{mn}\beta_{mn,kl} = \lambda_{kl} \,. \tag{1.25}$$

This relation is necessary and sufficient condition for the matrix χ to give the correct quantum channel $\mathcal{E}$. Formally we can express χ as

$$\chi_{mn} = \sum_{j,k=0}^{d^2-1} \kappa_{mn,kl}\lambda_{kl} \,, \tag{1.26}$$

where the matrix κ is the inverse of the $d^4 \times d^4$ matrix β. From here one obtains the reconstructed expression for the map $\mathcal{E}$ (for technical details see, e.g., Ref. [1] and references therein).

1.3 Structure of qubit channels

As discussed in the previous section (physical) quantum channels are described by linear trace-preserving CP maps $\mathcal{E}$ defined on a set of density operators [1,2,18]. Any qubit channel $\mathcal{E}$ can be imagined as an affine transformation of a three-dimensional Bloch vector $\vec{r}$ (representing a qubit state), i.e., $\vec{r} \to \vec{r}' = T\vec{r} + \vec{t}$, where T is a real 3x3 matrix and $\vec{t}$ is a translation [18]. This form guarantees that the transformation $\mathcal{E}$ is Hermitian and trace preserving. The CP condition defines (non-trivial) constraints on possible values of the involved parameters. In fact, the set of all CP trace-preserving maps forms a specific convex subset of all affine transformations. Representing the qubit states by four-dimensional vectors $\vec{v}_\varrho = (1, \vec{r})$, where the first element corresponds to the normalization of the state ($\mathrm{Tr}\varrho = 1$), one can express the action of the channel $\mathcal{E}$ in a more compact matrix form:

$$\mathcal{E}[\varrho] = \begin{pmatrix} 1 & \vec{0} \\ \vec{t} & T \end{pmatrix} \begin{pmatrix} 1 \\ \vec{r} \end{pmatrix} = \begin{pmatrix} 1 \\ \vec{t} + T\vec{r} \end{pmatrix} \,. \tag{1.27}$$

In other words the qubit channels form 4x4 matrices of the affine form.

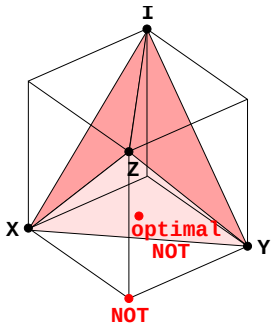

Figure 1.2 Unital CP maps are embeded in the set of all positive unital maps (cube). The CP maps form a tetrahedron with four unitary transformations in its vertices (extremal points) I, x, y, z corresponding to the Pauli σ-matrices. The non-physical U-NOT operation ($\lambda_1 = \lambda_2 = \lambda_3 = -1$) and its optimal completely positive approximation quantum Universal-NOT gate ($\lambda_1 = \lambda_2 = \lambda_3 = -1/3$) are shown (for more details see Sec. 5).

The matrix T can be written in the so-called singular-value decomposition, i.e., $T = R_U D R_V$ with R_U, R_V corresponding to orthogonal rotations and $D = \mathrm{diag}\{\lambda_1, \lambda_2, \lambda_3\}$ being diagonal where λ_k are the singular values of T. This means that any map $\mathcal{E}$ is a member of a less-parametric family of maps of the "diagonal form" $\Phi_{\mathcal{E}}$, i.e., $\mathcal{E}[\varrho] = U\Phi_{\mathcal{E}}[V\varrho V^\dagger]U^\dagger$ where U, V are unitary operators. The reduction of parameters is very helpful, and most of the properties (including complete positivity) of $\mathcal{E}$ are reflected by the properties of $\Phi_{\mathcal{E}}$. The map $\mathcal{E}$ is CP only if $\Phi_{\mathcal{E}}$ is. Let us note that $\Phi_{\mathcal{E}}$ is determined not only by the matrix D, but also by a new translation vector $\vec{\tau} = R_U\vec{t}$, i.e., under the action of the map $\Phi_{\mathcal{E}}$ the Bloch sphere transforms as follows $r_j \to r'_j = \lambda_j r_j + \tau_j$.

A special class of CP maps are the unital maps, that transform the total mixture ($\varrho = I/2$) into itself, $\mathcal{E}[I/2] = I/2$. In this case $\vec{t} = \vec{\tau} = \vec{0}$, and the corresponding map $\Phi_{\mathcal{E}}$ is uniquely specified by just three real parameters. The positivity of the transformation $\Phi_{\mathcal{E}}$ results into conditions $|\lambda_k| \leq 1$, while to fulfill the CP condition we need that the four inequalities $|\lambda_1 \pm \lambda_2| \leq |1 \pm \lambda_3|$ are satisfied. These conditions specify a tetrahedron lying inside a cube of all positive unital maps. In this case the extreme points represent four unitary transformations $I, \sigma_x, \sigma_y, \sigma_z$ (see Fig.1.2).

1.4 Method of maximum likelihood

The MML is a general estimation scheme [17,19] that has already been considered for a reconstruction of quantum operations from incomplete data. It has been studied by Hradil and Fiurášek [20], and by Sachci [21] (criticized in Ref. [22]). The task of the maximum likelihood in the process reconstruction is to find out a map $\mathcal{E}$, for which the *likelihood* is maximal. By the definition we assume that the estimated map has to be CP. Let us now briefly describe the main idea of this method in more details.

Given the measured data represented by couples ϱ_k, F_k (ϱ_k is one of the test states and F_k is a positive operator corresponding to the outcome of an individual

measurement used in the kth run of the experiment) the likelihood functional is defined by the formula

$$L(\mathcal{E}) = -\log \prod_{k=1}^{N} p(k|k) = -\sum_{k=1}^{N} \log \mathrm{Tr}\mathcal{E}[\varrho_k]F_k \ , \tag{1.28}$$

where N is the total number of measurement events ("clicks") and we used $p(j|k) = \mathrm{Tr}\mathcal{E}[\varrho_k]F_j$ for conditional probability of using the test state ϱ_k and observing the outcome F_j. The aim is to find a physical map $\mathcal{E}_{est}$ that maximizes this function, i.e., $L(\mathcal{E}_{est}) = \max_{\mathcal{E}} L(\mathcal{E})$. This variational task is usually performed numerically.

Our approach (see also Ref. [8]) is different from those described in Refs. [20–22] in the way that we find the maximum of the functional defined in Eq.(1.28). The parametrization of $\mathcal{E}$ itself guarantees the trace-preserving condition. Hence only the CP condition must be checked separately during the numerical maximization. Instead of using the Lagrange multipliers (and increasing thereby the number of parameters for the numerical procedure), we introduce the CP condition as an external boundary for a Nelder-Mead simplex algorithm. The maximization itself is performed by the Mathematica 5.0 built-in function with following parameters:

- **Method = Nelder-Mead**
  We chose the *Simplex algorithm* because it gives the most stable results with the smallest memory requirements.

- **Shrink ratio and Contract ratio = 0.95**
  These parameters are normally taken somewhere around 0.5. Their values close to unity induce a rather slow "cooling" of the process and prevents from falling into a local maxima. So the global minimum can be determined reliably. The price to pay is, as usual, a longer time for a numerical search.

- **Reflect ratio = 1.5**
  This parameter is bigger than the standard choice but it helps us to enhance the probability of finding the global maximum.

In what follows we will analyze different examples of non-physical operations and reconstructions that were obtained via MML method.

1.5 Universal-NOT gate

Let us assume the Universal-NOT (spin-flip) operation discussed in Sec. 2. This operation corresponds to the inversion of the Bloch sphere (see Fig. 1.3). It is well known that this inversion preserves angles (which is related to the scalar product $|\langle\Phi, \Psi\rangle|$ of rays). Therefore, by the arguments of the Wigner theorem the ideal spin-flip operation must be implemented either by a unitary or by an anti-unitary operation. Unitary operations correspond to proper rotations of the Bloch sphere,

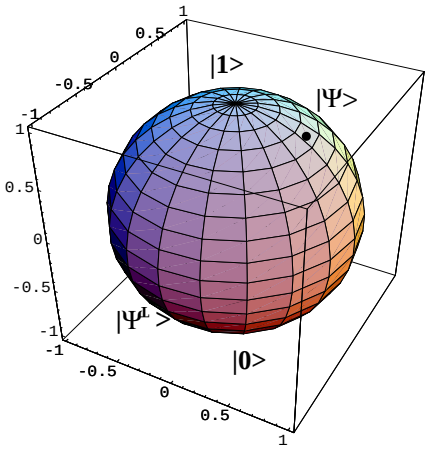

Figure 1.3 The state space of a qubit is the Bloch sphere. Pure states are represented by points on the sphere, while statistical mixtures are points inside the sphere. The Universal-NOT operation corresponds to the inversion of the sphere, since the states $|\Psi\rangle$ and $|\Psi^{\perp}\rangle$ are antipodes.

whereas anti-unitary operations correspond to orthogonal transformations with determinant -1. The spin-flip operation is an anti-unitary operation, i.e., it is not completely positive (see Sec. 2).

Due to the fact that the tensor product of an anti-linear and a linear operator is not correctly defined the Universal-NOT gate cannot be applied to a qubit while the rest of the world is governed by unitary evolution.[†] Therefore the ideal (perfect) Universal-NOT gate that would flip a qubit initially prepared in an *arbitrary* state does not exist.

Obviously, if the state of the qubit is known, then we can always perform a flip operation. In this situation the classical and quantum operations share many similar features, since the knowledge of the state is a classical information, which can be manipulated according to the rules of classical information processing (e.g. known states can be copied, flipped, etc). But, the universality of the operation is lost. That is, the gate which would flip the state $|0\rangle \rightarrow |1\rangle$, is not able to perform a flip $|(0\rangle + |1\rangle)/\sqrt{2} \rightarrow (|0\rangle - |1\rangle)/\sqrt{2}$.

Since it is not possible to realize a perfect Universal-NOT gate [9] which would flip an arbitrary (unknown) qubit state, it is of interest to study, what is the best approximation to the perfect Universal-NOT gate. Here one can consider two possible scenarios. The first one is based on the measurement of input qubit(s) – using the results of an optimal measurement one can manufacture an orthogonal qubit, or any desired number of them. Obviously, the fidelity of the NOT operation in this case is equal to the fidelity of estimation of the state of the input qubit(s). The sec-

[†]In fact, exactly this property makes the spin-flip operation so important in all criteria of inseparability for two-qubit systems [23,24].

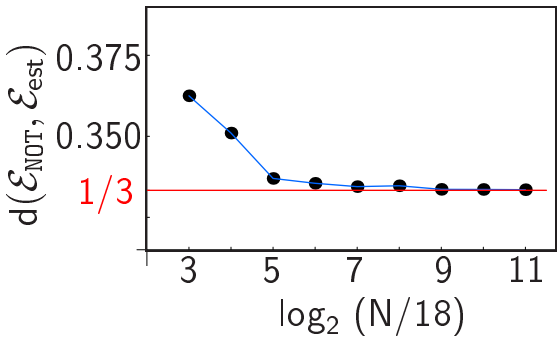

Figure 1.4 The distance $d(\mathcal{E}_{\texttt{NOT}}, \mathcal{E}_{est})$ as a function of the number of measured outcomes N in a logarithmic scale. We used 6 input states (eigenvectors of $\sigma_x, \sigma_y, \sigma_z$) and measured $\sigma_x, \sigma_y, \sigma_z$. The distance converges to the theoretical value $1/3$ that corresponds to the optimal Universal-NOT.

ond scenario would be to approximate an anti-unitary transformation on a Hilbert space of the input qubit(s) by a unitary transformation on a larger Hilbert space which describes the input qubit(s) and ancillas.

It has been shown recently, that the best achievable fidelity of both flipping scenarios is the same [10–12]. That is, the fidelity of the optimal Universal-NOT gate is equal to the fidelity of the best state-estimation performed on input qubits [25–27] (one might say, that in order to flip a qubit we have to transform it into a bit). More detailed description of the unitary transformation realizing the quantum scenario for the spin-flip operation can be found in Ref. [10–12]. The experimental realization of the optimal Universal-NOT gate has been reported in Ref. [14]. In this experiment, qubits were encoded in polarization states of photons.

As we have said, the Universal-NOT gate ($\mathcal{E}_{\texttt{NOT}} : |\Psi\rangle \rightarrow |\Psi^{\perp}\rangle$) is associated with the inversion of the Bloch sphere, i.e., $\vec{r} \rightarrow -\vec{r}$, which is not a CP map. It represents an non-physical transformation specified by $\lambda_1 = \lambda_2 = \lambda_3 = -1$. The distance (see Fig. 1.2) between this map and the tetrahedron of completely positive maps is extremal, i.e., it is the most non-physical map among linear transformations of a single qubit and can be performed only approximatively. A quantum "machine" that optimally implements an approximation of the Universal-NOT is represented by the map $\tilde{\mathcal{E}}_{\texttt{NOT}} = \mathrm{diag}\{1, -1/3, -1/3, -1/3\}$. The distance [28] between the U-NOT gate and its optimal physical approximation reads

$$d(\tilde{\mathcal{E}}_{\texttt{NOT}}, \mathcal{E}_{\texttt{NOT}}) = \int_{states} \mathrm{d}\varrho \mathrm{Tr}|(\mathcal{E}_{\texttt{NOT}} - \tilde{\mathcal{E}}_{\texttt{NOT}})[\varrho]| = 1/3\,. \tag{1.29}$$

This channel corresponds to the best CP approximation of the Universal-NOT gate (the spin-flip operation).

For the estimation of the Universal-NOT gate via our numerical Gedanken experiment we choose as inputs six test states - the eigenstates of $\sigma_x, \sigma_y, \sigma_z$. The

data are generated as (random) results of three projective measurements $\sigma_x, \sigma_y, \sigma_z$ applied in order to perform the output state reconstruction. In order to analyze the convergence of the method we have performed the reconstruction for different number of detected events ("clicks") and compared the distance between the original map $\mathcal{E}_{\texttt{NOT}}$ and the estimated map $\mathcal{E}_{est}$. The result is plotted in Fig.1.4, where we can see that the distance converges to $1/3$ as calculated in Eq.(1.29). For $N = 100 \times 18$ clicks, i.e., each measurement is performed 100 times per particular input state, the algorithm leads us to the map

$$\mathcal{E}_{est} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -0.0002 & -0.3316 & -0.0074 & 0.0203 \\ 0.0138 & -0.0031 & -0.3334 & 0.0488 \\ -0.0137 & 0.0298 & -0.0117 & -0.3336 \end{pmatrix}, \tag{1.30}$$

which is very close $[d(\mathcal{E}_{est}, \tilde{\mathcal{E}}_{\texttt{NOT}}) = 0.0065]$ to the best approximation of the Universal-NOT operation, i.e., $\tilde{\mathcal{E}}_{\texttt{NOT}} = \mathrm{diag}\{1, -1/3, -1/3, -1/3\}$.

We conclude that for sufficiently large N the MML reconstruction gives us the same result as a theoretical prediction derived in Ref. [10]. In order to illustrate the power of this approach we will find approximations of non-linear quantum mechanical transformations.

1.6 Non-linear transformations

Quantum mechanics is intrinsically linear theory and therefore non-linear transformations cannot be considered as legitimate quantum maps. Nevertheless, one can consider a toy model in which one looks for optimal physical approximations of non-linear quantum transformations. Such maps are sometimes used as an "effective" description of specific processes. In what follows we will consider two specific examples: The non-linear polarization rotation and the transformation generating powers of an input density operator.

1.6.1 Non-linear polarization rotation

Let us consider a non-linear transformation of a qubit defined by the relation [15]:

$$\mathcal{E}_\theta[\varrho] = e^{i\frac{\theta}{2}\langle\sigma_z\rangle_\varrho \sigma_z} \varrho e^{-i\frac{\theta}{2}\langle\sigma_z\rangle_\varrho \sigma_z}. \tag{1.31}$$

Unlike the Universal-NOT gate this map is non-linear. Four test states are not sufficient to allow us to determine the action of non-linear maps. Therefore in our Gedanken experiment we have to consider all possible input states (that cover the whole Bloch sphere) as test states, but still we use only three different measurements performed on the outcomes of the channel. These measurement data are sufficient for the channel reconstruction.

Firstly, we present an analytic derivation of a physical approximation of $\mathcal{E}_\theta$. This approximation is the closest physical map $\tilde{\mathcal{E}}_\theta$, i.e., $d(\tilde{\mathcal{E}}_\theta, \mathcal{E}_\theta) = \min$. The map

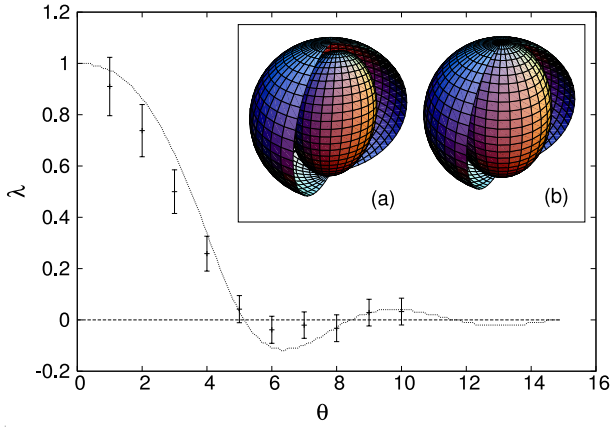

Figure 1.5 We present analytical as well as numerical results of an approximation of a non-linear map $\mathcal{E}_\theta$ for different values of the parameter θ (measured in radians). The numerical ("experimental") results shown in the graph in terms of a set of discrete points with error bars are obtained via the MML. The theoretical approximation $\tilde{\mathcal{E}}_\theta$ of the non-linear NPR map is characterized by the parameter λ that is plotted (solid line) in the figure as a function of the parameter θ. In the inset (a) is the Bloch sphere transformation for $\theta = 3$ obtained by MML and in (b) the same transformation obtained analytically.

$\mathcal{E}_\theta$ exhibits two symmetries: the continuous U(1) symmetry (rotations around the z-axis) and the discrete σ_x symmetry (rotation around the x-axis by π). The physical approximation $\tilde{\mathcal{E}}_\theta$ should possess these properties as well. Exploiting the two symmetries possible transformations of the Bloch vector are restricted as follows $x \to \lambda x, y \to \lambda y, z \to pz$. In the process of minimization the parameter p behaves trivially and equals to unity. It means that $\tilde{\mathcal{E}}_\theta$ is of the form $\mathcal{E}_\lambda = \mathrm{diag}\{1, \lambda, \lambda, 1\}$. Our task is to minimize the distance $d(\mathcal{E}_\theta, \mathcal{E}_\lambda) = \int d\varrho |\mathcal{E}_\theta[\varrho] - \mathcal{E}_\lambda[\varrho]|$ in order to find the physical approximation $\tilde{\mathcal{E}}_\theta$, i.e., the functional dependence of λ on θ.

We plot the parameter λ that specifies the best physical approximation of the NPR map in Fig. 1.5. In the same figure we also present a result of the maximum likelihood estimation of the NPR map based on a finite number of "measurements". Here, for every point (θ), the non-linear operation was applied to 1800 input states that have been chosen randomly (via a Monte Carlo method so they uniformly cover the whole Bloch sphere). These input states have been transformed according to the non-linear transformation (1.31). Subsequently simulations of random projective measurements have been performed. With these "experimental" data a maximization procedure was performed as described in the previous section. The resulting approximation specified by a value of λ (error bars shown in the graph represent the variance in outcomes for subsequent runs with different test states, but the same procedure parameters) transforms the original Bloch sphere as it is shown in an inset for particular value $\theta = 3$. The figure (a) corresponds to result obtained by MML, and the figure (b) has been obtained via analytic calculations. We see that the original Bloch sphere is transformed into an ellipsoid, one axis of

which is significantly longer than the remaining two axes, that are of a comparable length. The mean of these two lengths corresponds to the parameter λ that specifies the map. We conclude that the MML is in an excellent agreement with our analytical calculations.

1.6.2 Powers of density operator: $\varrho \longrightarrow \varrho^2$

Similar to the previous example, the map $\varrho \longrightarrow \varrho^2$ is intrinsically non-linear and we have to define the map for all possible inputs. The most general state of a qubit can be written as $\varrho(\vec{r}) = 1/2(I + \vec{r}.\vec{\sigma})$ where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are Pauli matrices and $\vec{r}$ is a real vector and $\vec{r} \leq 1$. In such a notation we can define the map under consideration as follows:

$$\mathcal{E}[\varrho] = \frac{1}{2}\left(I + |\vec{r}|\ \vec{r}.\vec{\sigma}\right) = |\vec{r}|\varrho + (1 - |\vec{r}|)\, I/2. \tag{1.32}$$

Since the action of the map cannot be written as a matrix independent of ϱ times ϱ the map is not linear. Hence it is a very interesting example to study reconstruction schemes, as it is does not change the typical test states (pure states and complete mixture state).

Due to the high symmetry of the map one would expect that the best physical approximation is a contraction of the whole Bloch sphere

$$\mathcal{E}[\varrho]_{physical} = k\varrho + (1 - k)\, I/2 \tag{1.33}$$

with $k \leq 1$. Indeed the result of the MML method shows within the precision given by the finite number of test-states a result of the expected form

$$\mathcal{E}[\varrho]_{physical} = 0.85\, \varrho + (1.00 - 0.85)\, I/2. \tag{1.34}$$

Even thought here we have considered only the second power of the density operator ϱ using the same arguments one can approximate channels that generate an arbitrary power of the original density operator. In addition, taking into account all symmetries associated with the transformation $\varrho \longrightarrow \varrho^k$ one can derive an analytical expression for its physical (CP) approximation.

1.7 Analysis of results

With the three examples presented in previous section we have demonstrated the power of the MML method to find physical approximations of operations that are truly non-physical. In reality (when analyzing data of real experiments) one always expects to have physical operations. However, if the data indicate a non-physical operation, this may be a consequence of errors in experiment, wrong interpretation of data, but also a failure of the reconstruction method.

To rule out the last possibility, one has always to analyze precisely the outcome of the procedure. By performing the maximization numerically, the functional itself is a rather simple function of the input data. The only challenges are boundary conditions imposed by the CP condition. These may cause the search-engine to "stick" in a point, which is not a local maxima of the functional, but is confined by the boundary conditions. This case is easy to detect by calculating the CP condition of the resulting operation and to check if the result is on the boundary of the CP maps. If so, it is worthwhile to run the maximization procedure again with different starting conditions. However, for parameters of the maximization procedure specified in Sec. 3 for every testing case the MML resulted in the proper maxima at the first attempt.

If the result of the MML method is correct and the resulting operation is yet on the boundary of CP maps, there is a strong evidence that the incoming data were biased by some kind of errors. To analyze this problem closer, we have to take into account not only the resulting operations, but also the value of L in Eq. (1.28). This value defines the logarithm of the probability to obtain, for specified input states (used in the MML method as test states), the same results as the experimental results. For proper data, this probability should be comparable to the probability of a sequence of measured data produced directly by the reconstructed operation. However, for improper data (in our examples these data are produced by non-physical operations) the reconstructed operations may reproduce these data with a much smaller probability.

We define the $L_{data}(\mathcal{E})$ to be the value of the functional (1.28) for the resulting approximation. For the same set of test states as used in the original experiments (denoted by ϱ_k) we perform a Gedanken-experiment. We apply the reconstructed operation on every such state and then apply the measurement in the same direction as in the original experiment (the resulting positive operators we denote $\bar{F}_k$). Then we define

$$L_{test}(\mathcal{E}) = -\sum_{k=1}^{N} \log \mathrm{Tr}\mathcal{E}[\varrho_k]\bar{F}_k. \tag{1.35}$$

The same procedure may be repeated sufficiently many times to obtain a typical value of the functional (calculated as the mean of all runs and denoted by $\bar{L}_{test}(\mathcal{E})$) and the typical variance of this value $\sigma(L_{test}(\mathcal{E}))$.

In Tab. 1.1 results of the calculations for three examples presented in this Chapter are shown. As one can clearly see, for the physical approximation of the Universal-NOT gate the difference between $L_{data}(\mathcal{E})$ and $\bar{L}_{test}(\mathcal{E})$ is rather big, showing a clear evidence that the input data originated in an non-physical operation. This is, however, not the case for the non-linear polarization rotation and for the transformation $\varrho \longrightarrow \varrho^2$. In these cases the typical sequence (as a whole) of the experimental data has a comparable probability to appear as any other sequence produced by the reconstructed operation. So we may conclude that this method

gives us a partial tool (a necessary condition) to identify data that would lead to non-physical operations.

	$L_{data}(\mathcal{E})$	$\bar{L}_{test}(\mathcal{E})$	$\sigma(L_{test}(\mathcal{E}))$
U NOT	-1371	-1472	6.79
NPR	-1423	-1426	12.64
$\varrho \longrightarrow \varrho^2$	-1343	-1335	17.44

Table 1.1 Values of the functional L for different examples of non-physical operations. For the Universal-NOT gate it is clear that the probability to obtain sequence similar to the sequence given by the experimental data is much lower than to obtain a typical sequence of data produced by the reconstructed operation. For other two examples the difference is not significant.

1.8 Conclusions

In this Chapter we have shown that the method of maximum likelihood can be efficiently used for derivation of physical approximations of non-physical maps (both non-CP linear maps as well as non-linear quantum-mechanical transformations). We have applied this method for approximating qubit transformations (the Universal-NOT gate, the non-linear polarization rotation and the map $\varrho \longrightarrow \varrho^2$). We have analyzed the resulting operations and provided a tool to detect the quality of the input data.

Finally we note that in this Chapter we have considered an idealized situation when the input states of test particles can be prepared perfectly, i.e., the action of the initial-state preparator is totally known. Certainly, this is an approximation of a real situation, when test states are prepared with a finite precision. This additional source of uncertainty has to be taken into account in realistic estimation procedures of quantum channels.

Acknowledgments

This work was supported in part by the European Union projects QAP, HIP, and QUROPE, by the Slovak Academy of Sciences via the project CE-PI and the APVV project.

References

1. M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information*, (CUP, Cambridge, 2000).

2. J. Preskill, *Lecture Notes on Physics 229: Quantum Theory of Information and Computation* (1998),
available at `http://www.theory.caltech.edu/people/preskill/`.

3. G.M. D'Ariano and P. Lo Presti: "Quantum tomography for measuring experimentally the matrix elements of an arbitrary quantum operation" Phys.Rev.Lett. **86**, 4195 – 4198 (2001).
4. J.F. Poyatos, J.I. Cirac, and P. Zoller : "Complete characterization of a quantum process: The two-bit quantum gate, " Phys.Rev.Lett. **78**, 390 – 393 (1997).
5. V. Bužek: "Reconstruction of Liouvillian superoperators." Phys. Rev. A **58**, 1723 – 1727 (1998).
6. V. Bužek, G. Adam, and G. Drobný: "Reconstruction of Wigner functions on different observation levels." Ann. Phys. (N.Y.) **245**, 37 – 97 (1996)
7. V. Bužek, *Quantum tomography from incomplete data via* MaxEnt *principle*, in *Quantum Estimations: Theory and Experiment* (Springer-Verlag, Berlin, 2004), pp. 189 – 234.
8. M. Ziman, M. Plesch, V. Bužek, and P. Štelmachovič: "Process reconstruction: From unphysical to physical maps via maximum likelihood." Phys. Rev. A **72**, 022106 – 1-5 (2005).
9. H. Bechmann-Pasquinucci and N. Gisin: "Incoherent and coherent eavesdropping in the six-state protocol of quantum cryptography." Phys. Rev A **59**, 4238 – 4248 (1999).
10. V. Bužek, M. Hilery, and R.F. Werner: "Optimal manipulations with qubits: Universal-NOT gate." Phys.Rev.A **60**, R2626 - R2629 (1999).
11. V. Bužek, M. Hillery, and R.F. Werner: "Universal-NOT gate." J. Mod. Opt. **47**, 211 – 232 (2000).
12. P. Rungta, V. Bužek, C.M. Caves, M. Hillery, and G.J. Milburn: "Universal state inversion and concurrence in arbitrary dimensions." Phys. Rev. A **64**, 042315 – 1-13 (2001).
13. N. Gisin and S. Popescu: "Spin flips and quantum information for antiparallel spins." Phys. Rev. Lett. **83**, 432 – 435 (1999).
14. F. DeMartini, V. Bužek, F. Sciarrino, and C. Sias: "Experimental realization of the quantum universal NOT gate." Nature **419** No. 6909, 815 – 818 (2002).
15. C. Vinegoni, M. Wegmuller, B. Huttner, and N. Gisin: "Measurement of nonlinear polarization rotation in a highly birefringent optical fibre using a Faraday mirror." J. Opt. A **2**, 314 – 318 (2000).
16. V.Bužek and R.Derka, *Quantum observations* in *Coherence and Statistics of Photons and Atoms* ed. J.Peřina (John Wiley & Sons, New York, 2001) pp. 198 – 261.
17. Z. Hradil, J. Řeháček, J. Fiurášek, and M. Ježek, *Maximum likelihood methods in quantum mechanics*, in *Quantum Estimations: Theory and Experiment* (Springer-Verlag, Berlin, 2004), p. 63.

18. M.B. Ruskai, S. Szarek, and E. Werner: "An analysis of completely positive trace-preserving maps on $M - 2$." Lin. Alg. Appl. **347**, 159 – 187 (2002).

19. R.A. Fisher: "Theory of statistical estimation." Proc. Cambridge Phil. Soc. **22**, 700 – 725 (1925)

20. J. Fiurášek and Z. Hradil: "Maximum-likelihood estimation of quantum processes." Phys. Rev. A **63**, 020101(R) – 1-4 (2001).

21. M.F. Sacchi: "Maximum-likelihood reconstruction of completely positive maps." Phys. Rev. A **63**, 054104 – 1-4 (2001).

22. J. Fiurášek and Z. Hradil: "Comment on "Maximum likelihood reconstruction of CP maps", quant-ph/0009104." Los Alamos e-print arXiv `quant-ph/0101048`.

23. A.Peres: "Separability criterion for density matrices." Phys. Rev. Lett. **77**, 1413 – 1415 (1996).

24. P. Horodecki: "Separability criterion and inseparable mixed states with positive partial transposition." Phys. Lett. A **223**, 333 – 339 (1997);
M. Horodecki, P. Horodecki, and R. Horodecki: "Inseparable two spin-1/2 density matrices can be distilled to a singlet form." Phys. Rev. Lett. **78**, 574 –577 (1997).

25. A. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (Amsterdam, North Holland, 1982).

26. S. Massar and S. Popescu: "Optimal extraction of information from finite quantum ensembles." Phys. Rev. Lett. **74**, 1259 – 1263 (1995).

27. R. Derka, V. Bužek, and A. Ekert: "Universal algorithm for optimal estimation of quantum states from finite ensembles via realizable generalized measurement." Phys. Rev. Lett. **80**, 1571 – 1575 (1998), and references therein.

28. M. Ziman, M. Plesch, and V. Bužek: "Reconstruction of superoperators from incomplete measurements." Foundations of Physics **36**, 127 – 156 (2006).