THESIS

State estimations for 2-level quantum systems

László Ruppert

Supervisors:

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

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

Budapest University of Technology and Economics

2008

Contents

1 Introduction

State estimation is a fundamental problem in both quantum information theory and quantum control. In quantum control [2] its role is the same as in the classical control theory, i.e. to give an estimate of the unmeasured time-dependent state variable in order to be used in state feedback schemes. On the other hand, the measurement of a quantum mechanical system is probabilistic, so even the measurement of a measurable quantity asks for estimation methods, this is why state estimation is an important field in quantum information theory [16, 17], too.

To set up a quantum state estimation, or a quantum state tomography method, two ingredients has to be given: the measurement strategy used for getting information, and the estimator mapping the measurement data to the state space. If one uses a von Neumann measurement, the projective nature of the measurement forces the use of several copies of the same system being in the same state [20], what is in general rather difficult to implement in practice. However, in certain physical circumstances, for example in quantum optics, it is natural to have several copies of the quantum system in the same state.

The properties of different state estimations are investigated here in the simplest case of a quantum bit (qubit). This problem may be traced back to the seventies [7] and there is a well-known measurement scheme called the standard scheme [17] for estimating the state of a qubit. Although the state estimation problem is quite old, the interest in a thorough mathematical analysis of the quantum state estimation procedures has been flourishing recently [1, 6, 14, 19]. The methods using von Neumann type measurements for qubits in both pure and mixed states are now quite well developed. For example, an adaptive observable selection strategy based on a Bloch vector parametrization in spherical coordinates and on a Bayesian estimation method of qubits in mixed states is reported in the paper [5]. However, there are only a few papers [18, 20] that deal with the properties of the estimate when a finite number of measurements are only available. Here the quality of the estimate is usually characterized by the covariance matrix, the trace of which can be minimized if an optimal scheme is to be determined.

The methods using von Neumann measurements are useless, if one aims at the estimation of dynamically evolving states, as in quantum control. A possible way to circumvent the obstruction of the demolition property of von Neumann measurements is to use an indirect measurement scheme, where the 'unknown' quantum system is coupled with a 'measurement' (also called 'probe' or 'ancilla') system and the measurements are only applied on the measurement system [9, 11]. In the literature this method is often termed

weak measurement [4, 23]. Note, that most of the papers dealing with indirect or weak measurement schemes use a continuous-time approach [10].

The first aim of this paper is to generalize the standard scheme such a way that we are able to measure in non-symmetric way, i.e. we are not using the Pauli matrices as observables and not measuring the same number in each directions. This can be useful to decrease the variance of estimator if the qubit is nearly pure. A further aim is to construct a compromising estimator, that finds the trade-off between the effectiveness of the estimate and the number of qubits that are un-affected by the measurements. This will be achieved using indirect measurement scheme in the discrete time case. The simplest possible case is considered, where both the unknown and the measurement quantum systems are quantum bits.

The paper is organized as follows. Chapter 2 clarifies the notation used throughout the work and gives an introduction to the mathematics of quantum mechanics. At the end of this chapter the main results are placed in the context of the related literature. The generalization of standard scheme is discussed in Chapter 3. We will find the optimal ratios of number of measurements and optimal observables, then giving an asymptotically optimal algorithm. Thereafter, some approximately efficient numerical approaches are presented for the case of finite number of measurements. Chapter 4 presents a simple indirect measurement strategy and its properties. Then a comparison with other known strategies are given and some modified versions are presented. Finally, Chapter 5 concludes.

2 Basic notions

This chapter summarizes the background used in this thesis. Sections 2.1, 2.2 and 2.3 follows the presentation of [17].

2.1 State of a quantum system

The states of finite quantum systems are represented by $n \times n$ density matrices ($\rho(\theta)$), obeying the following 2 properties:

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

$$\rho(\theta) \geq 0 \tag{2}$$

In the simplest case we are talking about the quantum bit (shortly qubit), which is the key object of our examination. In this case ρ is a 2×2 matrix, and the so called Bloch parameterization gives a geometrically clear viewpoint of the state-space:

$$\rho(\theta) = \frac{1}{2}\left(I + \theta_1\sigma_1 + \theta_2\sigma_2 + \theta_3\sigma_3\right), \tag{3}$$

where

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

that is

$$\rho(\theta) = \frac{1}{2} \begin{bmatrix} 1+\theta_3 & \theta_1 - i\cdot\theta_2 \\ \theta_1 + i\cdot\theta_2 & 1-\theta_3 \end{bmatrix}.$$

The so called Bloch vector θ gives an equivalent representation of the quantum states. It is easy to see, that the second property of density matrices (2) transforms to

$$\theta_1^2 + \theta_2^2 + \theta_3^2 \leq 1,$$

in the Bloch-parametrization. It means, that the state space is the unit ball in $\mathbf{R}^3$. States for which the equation holds are called pure states, all the others are mixed states. In what follows, the states of a quantum bit will be represented either as a density matrix, or as a Bloch vector.

In Section 4 we will couple two separated qubits. Let us denote the Bloch representation of the unknown system and the probe (measurement device), where both are qubits, as

$$\rho_S = \frac{1}{2}(I + \sum_{i=1}^{3} \theta_i^S \sigma_i) \quad , \quad \rho_M = \frac{1}{2}(I + \sum_{i=1}^{3} \theta_i^M \sigma_i), \tag{4}$$

where θ_i^S are the Bloch parameters of the system qubit and θ_i^M are the Bloch parameters of the measuring qubit. The state of the composite system of two independent qubits is represented as a 4×4 density matrix:

$$\rho_{S+M} = \rho_S \otimes \rho_M. \tag{5}$$

The above 4-level state is a special product state that reflect the fact that the two qubits are separated. If the state of a 4-level system can be described in the form of convex combination of product states

$$\rho_{12} = \sum_i \alpha_i \ \rho_1^{(i)} \otimes \rho_2^{(i)}$$

then it is called separable state, else it is called an entangled state.

We can get the state of a subsystem of the composite system using the reduced density operator. Denote the reduced density operator for the first subsystem of the state ρ_{12} with ρ_1, then

$$\rho_1 = \mathrm{Tr}_2 \rho_{12}, \tag{6}$$

where Tr_2 is the partial trace over the second subsystem, i.e ρ_1 need to satisfy the equation

$$\mathrm{Tr}\left(\rho_{12}(A \otimes I)\right) = \mathrm{Tr}(\rho_1 A) \tag{7}$$

for all self-adjoint operators A on $\mathbf{C}^{2\otimes 2}$.

2.2 Dynamics

The Schrödinger picture is used here in discrete time that associates a unitary U to the time-evolution of the system such that

$$\rho(k+1) = U(k)\rho(k)U(k)^*. \tag{8}$$

We will only use dynamics for the coupled 4-level system, and only in one time-step:

$$\hat{\rho}_{S+M} = U_{S+M} \ \rho_{S+M} \ U_{S+M}^*. \tag{9}$$

The state of the composite system after the interaction is described by $\hat{\rho}_{S+M}$ in (9) where U_{S+M} is the overall system evolution unitary. Note that after the interaction the resulting state $\hat{\rho}_{S+M}$ will be entangled in the general case, even if the initial state ρ_{S+M} was separable (as in (5)).

When we are interested in the dynamical change of the system S, the first reduced density matrix should only be considered:

$$\hat{\rho}_S = \mathrm{Tr}_M \hat{\rho}_{S+M} \tag{10}$$

where Tr_M is defined in (6 - 7).

In order to have a simple parametrization of the interaction (coupling) between the unknown and measurement qubit, the Cartan decomposition [12, 21] of the discrete time evolution unitary U_{S+M} is used in the form

$$U_{S+M} = L_1 e^{ah} L_2 \tag{11}$$

where L_1 and L_2 are in $SU(2) \otimes SU(2)$, $h \in \mathbb{R}$ and $a \in \mathbf{a}$ with

$$\mathbf{a} = \mathrm{i}\ realspan\{\sigma_1^S \otimes \sigma_1^M, \sigma_2^S \otimes \sigma_2^M, \sigma_3^S \otimes \sigma_3^M\} \tag{12}$$

Because both L_1 and L_2 are in a product form, they describe the product of the local dynamical effects L_i^S and L_i^M $(i = 1, 2)$, and the interaction is parameterized by three real parameters a_1, a_2 and a_3.

Therefore, the dynamical equation of qubit S in (10) becomes

$$\hat{\rho}_S = L_1^S \mathrm{Tr}_M \left(e^{ah} \tilde{\rho}_S \otimes \tilde{\rho}_M e^{a^* h} \right) L_1^{S*} \tag{13}$$

where $L_1 = L_1^S \otimes L_1^M$, $L_2 = L_2^S \otimes L_2^M$ both time dependent, and $\tilde{\rho}_S = L_2^S \rho_S L_2^{S^*}$, $\tilde{\rho}_M = L_2^M \rho_M L_2^{M*}$. In order to simplify the forthcoming computations, we consider the case with no local dynamics, when $L_i^S = L_i^M = I$ $(i = 1, 2)$.

Instead of the unitary description (8) of the dynamics, we can use also the so-called T-representation of the linear mapping $\theta(k-1) \mapsto \theta(k)$ that corresponds to the original state transformation $\rho(k-1) \mapsto \rho(k)$ in (8) with a real 3×3 matrix T, such that

$$\theta(k+1) = T\theta(k), \tag{14}$$

so the T matrix describes the effect of dynamics on the Bloch vectors.

Example 1 *Let the unitary be the following:*

$$U_{S+M} = e^{-\mathrm{i}h(a_1 \sigma_1^S \otimes \sigma_1^M)} \tag{15}$$

i.e. the qubits are interacting only in the x direction. Computing the dynamics according to (13) we obtain that the system dynamics in T-representation is the following

$$\hat{\theta}_S = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(2a_1 h) & -\sin(2a_1 h)\theta_{M1} \\ 0 & \sin(2a_1 h)\theta_{M1} & \cos(2a_1 h) \end{bmatrix} \theta_S, \tag{16}$$

if there were no measurements performed.

2.3 Von Neumann measurements

In quantum mechanics measurements have a probabilistic nature. The subject of measurements are self-adjoint $n \times n$ matrices, so called observables. Let the spectral decomposition of an observable χ be the following:

$$\chi = \sum_i \lambda_i P_i. \tag{17}$$

Where λ_i are the different eigenvalues of χ, and P_i are the corresponding eigenprojections.

The possible outcomes of the measurement are the different λ_i eigenvalues and the corresponding probability is

$$\mathrm{Prob}(\lambda_i) = \mathrm{Tr}(\rho P_i). \tag{18}$$

A key element of quantum measurements is that it will change the actual state of the quantum system to

$$\rho_i = \frac{P_i \rho P_i}{\mathrm{Tr} P_i \rho P_i}, \tag{19}$$

if the outcome of measurement is λ_i.

Example 2 *Measurement of the Pauli operator σ_1:*

If one considers the measurement of the observable σ_1, then the possible outcomes are the different eigenvalues of the observable, i.e. ± 1. The probabilities of the different outcomes are

$$\begin{aligned} \mathrm{Prob}(+1) &= \mathrm{Tr}\rho E_{+1} = \tfrac{1}{2}(1+\theta_1) \\ \\ \mathrm{Prob}(-1) &= \mathrm{Tr}\rho E_{-1} = \tfrac{1}{2}(1-\theta_1) \end{aligned}$$

respectively, where the spectral decomposition of σ_1 is

$$\sigma_1 = E_{+1} - E_{-1} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix},$$

The state after measurement can be

$$\theta^{+1} = \begin{bmatrix} +1, & 0, & 0 \end{bmatrix}^T, \quad \theta^{-1} = \begin{bmatrix} -1, & 0, & 0 \end{bmatrix}^T.$$

depending on the actual outcome.

2.4 Literature review

The state estimation methods developed for qubits in the literature are briefly described in this section.

2.4.1 Standard method

We start with the most important method, the standard method [18], because we will make comparison with the standard method in all other cases.

Three different kinds of measurements are performed, they are represented by the three Pauli matrices $\sigma_1, \sigma_2, \sigma_3$. Since all the three are unitary their eigenvalues (which are the possible values of the measurement) are ± 1. If we measure σ_i on the state

$$\rho(\theta) = \frac{1}{2} \begin{bmatrix} 1+\theta_3 & \theta_1 - i\theta_2 \\ \theta_1 + i\theta_2 & 1-\theta_3 \end{bmatrix}$$

then the probability of outcome $+1$ is

$$\mathrm{Prob}(+1_i) = \frac{1+\theta_i}{2}.$$

similarly to Example 2. From the physical point of view, σ_1, σ_2 és σ_3 mean the spin measurement in directions x, y and z, respectively.

Because of the high symmetry and the independency of components we can easily make an estimation scheme for the state θ. Suppose, that m measurements are performed in each three directions. Then the relative frequency ν_i of the outcomes $+1$ is a sufficient statistic, so it is enough to use:

$$\nu_i := \frac{m_i}{m}, \quad \text{where } i = 1,2,3, \text{ and } m_i \text{ is the number of } (+1)\text{-s in the direction of } i.$$

The least squares estimator is a widely known and used method which minimizes the squared error. If the relative frequencies resulting from the measurements are ν_i, then the deviation from the real value of the state can be written in the form (because of the independent measurements, there are no cross terms):

$$\varphi_{LS}^m(\nu, \theta) = \sum_{i=1,2,3} \left(\nu_i - \frac{1+\theta_i}{2} \right)^2$$

We are going to minimize this expression. It is trivially minimal, when the expressions in parenthesis are zeros. This way, knowing relative frequencies ν_i, an estimate can be given for the Bloch vector θ:

$$\Phi^m(\nu) = \begin{bmatrix} 2\nu_1 - 1 \\ 2\nu_2 - 1 \\ 2\nu_3 - 1 \end{bmatrix} \tag{20}$$

Note that this estimation can provide physically meaningless result (the state will be outside of the Bloch ball), but the probability of having false estimate asymptotically vanishes as the number of measurement increases.

The covariance matrix ($V^{standard}$) of this estimator is the following:

$$V^{standard} = \frac{1}{m} \begin{bmatrix} 1-\theta_1^2 & 0 & 0 \\ 0 & 1-\theta_2^2 & 0 \\ 0 & 0 & 1-\theta_3^2 \end{bmatrix} \tag{21}$$

Because two matrices are usually not comparable, we will examine the traces of covariance matrices. If we have the opportunity to measure only N times, then $m = \frac{N}{3}$, and the trace of the covariance matrix will be the following:

$$\text{Tr } V_N^{standard}(\theta) = \frac{9 - 3\|\theta\|_2^2}{N}. \tag{22}$$

2.4.2 Minimal qubit tomography

Rehacek, Englert and Kaszlikowski has noticed in 2004, that in the standard method there is altogether 6 measurement directions, but in the 3 dimensional case it is enough to measure in only 4 directions to get all the possible states [20]. So they did not measure in the direction of the axis of the Bloch ball, but the projections were related to the vertices of a regular tetrahedron, and they applied the maximum likelihood principle to get an estimation for θ.

It is possible to calculate the covariance matrix of the estimator, but this is not comparable with the covariance matrix of the standard method in most of the cases, but if we compare the trace, then we get that the standard method is more efficient if $\theta \neq 0$ [18].

After all this conception is not useless, because they suggested an adaptive process, too, which rotates the measurement directions so that a possible outcome becomes on the opposite side of the assumed θ. This way they achieved that in that direction they never or only a few times get an outcome, so they obtained a better result for pure states than if they used standard method.

2.4.3 Indirect (weak) measurements

It is intuitively clear, that one must make a compromise between the information gained in a measurement and the disturbance or demolition caused by it. The general impossibility of determining the state of a single quantum system is proved in [3] whatever

measurement scheme is used. This indicates that the efficiency or precision provided by an indirect measurement scheme is necessarily smaller than that of a scheme that uses von Neumann measurements.

A related problem to the state estimation is to prepare the state of a given system in a specified way. Most papers apply some kind special dedicated measurements either to drive the system into a desired state or to compensate for the 'measurement back-action'. An application of weak measurements in bipartite state purification can be seen in [8], where the authors also use continuous time dynamics. Korotkov and Jordan [13] have shown that "it is possible to fully restore any unknown, pre-measured state, though with probability less than unity." A similar method will used in our approach, but we will improve a discrete-time model and make a state estimation from the measurements, too.

3 Optimization of the standard method

This chapter is devoted to a family of a methods that attempt to improve the standard method by varying the number of measurements in different directions and by changing the directions themselves.

3.1 Determination of optimal measurement directions and ratios

First we observe, that we measure in each direction the same number using the standard method independently of the state of the system. A possible way to improve the measurement scheme is that we optimize the number of measurements in certain directions such that the total variance (trace of the covariance matrix) will be minimal.

Assume we are allowed to measure on only N identically prepared copy of the system using the standard observable formed by the Pauli matrices and measure in the appropriate directions n_1, n_2 and n_3 times ($N = n_1 + n_2 + n_3$). Then similarly to the standard method the measurements in a certain direction only depend on the appropriate θ_i. But the measurements are still independent in each direction and the estimator will be the same as (20):

$$\Phi^{n_1,n_2,n_3}(\nu) = \begin{bmatrix} 2\nu_1 - 1 \\ \\ 2\nu_2 - 1 \\ \\ 2\nu_3 - 1 \end{bmatrix}. \tag{23}$$

The difference that here:

$\nu_i := \dfrac{m_i}{n_i}$, where $i = 1, 2, 3,$ and m_i is the number of $(+1)$-s in the direction of i.

Theorem 1 *The total variance is minimal if*

$$n_1 : n_2 : n_3 = \sqrt{1-\theta_1^2} : \sqrt{1-\theta_2^2} : \sqrt{1-\theta_3^2} \tag{24}$$

Proof: Similarly to (22) the covariance matrix will be:

$$V_{n_1,n_2,n_3}(\theta) = \begin{bmatrix} \frac{1-\theta_1^2}{n_1} & 0 & 0 \\ 0 & \frac{1-\theta_2^2}{n_2} & 0 \\ 0 & 0 & \frac{1-\theta_3^2}{n_3} \end{bmatrix}$$

Our aim is to minimize the trace of $V_{n_1,n_2,n_3}(\theta)$ for all possible n_1, n_2, n_3 (while for the standard measurement scheme $n_1 = n_2 = n_3 = n$). This is a simple optimization task:

$$\frac{1-\theta_1^2}{n_1} + \frac{1-\theta_2^2}{n_2} + \frac{1-\theta_3^2}{n_3} \to \texttt{min},$$

subject to: $n_1 + n_2 + n_3 = N$.

We can solve the above optimization problem using a Lagrange-function:

$$L(n_1, n_2, n_3, \lambda) = \frac{1-\theta_1^2}{n_1} + \frac{1-\theta_2^2}{n_2} + \frac{1-\theta_3^2}{n_3} + \lambda \cdot (n_1 + n_2 + n_3 - m)$$

From the necessary conditions for the minimum in each direction we obtain:

$$\lambda = \frac{1-\theta_1^2}{n_1^2} = \frac{1-\theta_2^2}{n_2^2} = \frac{1-\theta_3^2}{n_3^2}$$

□

Corollary 1 *The minimal trace of covariance matrix is:*

$$\mathrm{Tr}V_N^{minimal}(\theta) = \frac{1}{N}\left(\sqrt{1-\theta_1^2} + \sqrt{1-\theta_2^2} + \sqrt{1-\theta_3^2}\right)^2 \tag{25}$$

We can compare it with the trace for the standard method (22) by using the efficiency factor:

$$\eta(\theta) = \frac{\mathrm{Tr}V_{minimal}^n(\theta)}{\mathrm{Tr}V_{standard}^n(\theta)} = \frac{\left(\sqrt{1-\theta_1^2} + \sqrt{1-\theta_2^2} + \sqrt{1-\theta_3^2}\right)^2}{9 - 3\left(\theta_1^2 + \theta_2^2 + \theta_3^2\right)}$$

We can easily get that $0 \leq \eta(\theta) \leq 1$, $\eta(\theta) = 1$ if and only if $\theta_1 = \theta_2 = \theta_3$, when the minimal and standard schemes coincide, and the smaller is the value of η, the efficient is the estimation.

We examine 2 important examples:

Example 3 *If the state is in the x-axis direction, i.e. $\theta = (x, 0, 0)$ then the efficiency is*

$$\eta(x, 0, 0) = \frac{\left(\sqrt{1-x^2} + 2\right)^2}{9 - 3x^2} \tag{26}$$

The value of the efficiency in $x = 0$ is $\eta(0, 0, 0) = 1$, and as the state gets closer to the surface the efficiency factor decreases monotonously (i.e. $\eta'(x, 0, 0) < 0$, if $x > 0$, and $\eta'(x, 0, 0) > 0$, if $x < 0$). The value on the surface is $\eta(1, 0, 0) = \frac{2}{3}$. So as the true state approaches the surface, the minimal scheme becomes better and better (see Figure 1).

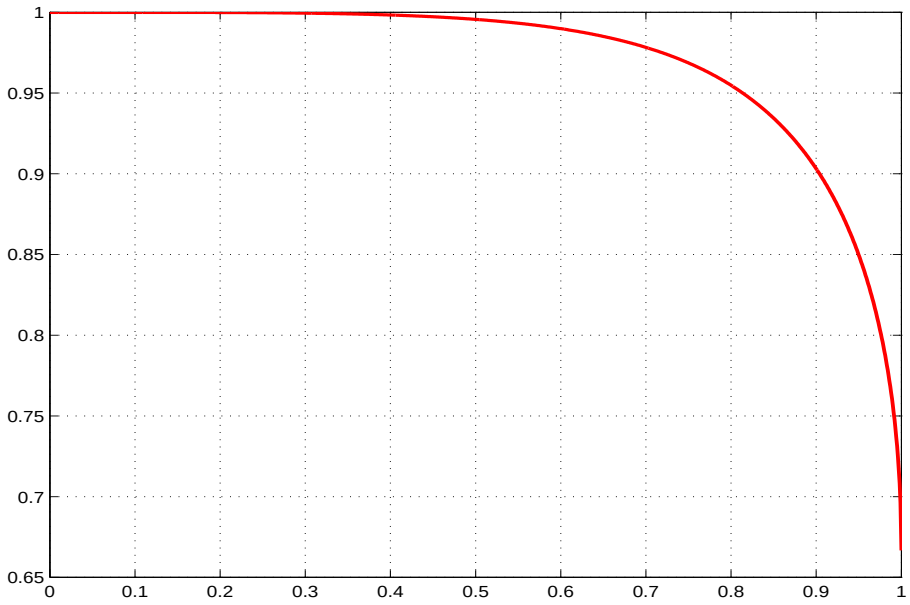

Figure 1: Efficiency along the axis $(x, 0, 0)$

Example 4 *Consider the case when the state is on the surface,* $\theta_1^2 + \theta_2^2 + \theta_3^2 = 1$

In this case we can observe a highly asymmetric pattern (see Figure 2): the measurement scheme is the most effective if the true state is near the direction of one axis.

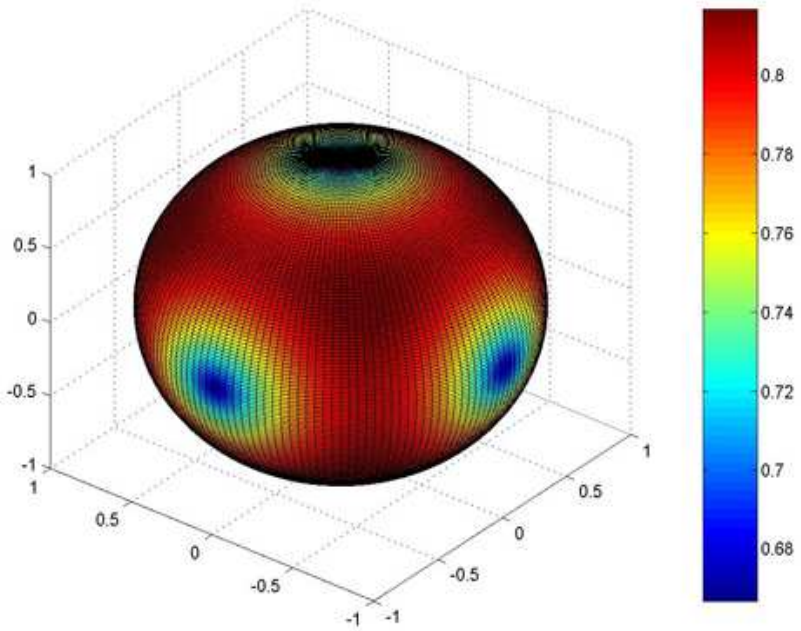

Figure 2: Efficiency on the surface

Theorem 2 *The minimal scheme is the most effective on the Bloch sphere ($\|\theta\|_2 = r$) in the axis directions.*

Proof: Because $\mathrm{Tr}V^n_{standard}(\theta) = (9 - 3r^2) \cdot \frac{1}{n}$, so the variance of standard method is constant on the whole sphere. Therefore it is enough to minimize (25):

$$\mathrm{Tr}V^n_{minimal}(\theta) = \frac{1}{n}\left(\sqrt{1-\theta_1^2} + \sqrt{1-\theta_2^2} + \sqrt{1-\theta_3^2}\right)^2,$$

subject to $\theta_1^2 + \theta_2^2 + \theta_3^2 = r^2$.

The above minimization problem can be solved by using spherical coordinates and using criteria for partial derivatives, but because of symmetry it is simpler with the Lagrange-method. Notice, that (25) is minimal if $\sqrt{1-\theta_1^2} + \sqrt{1-\theta_2^2} + \sqrt{1-\theta_3^2}$ is minimal, so the corresponding Lagrange function is

$$L(\theta_1, \theta_2, \theta_3, \lambda) = \sqrt{1-\theta_1^2} + \sqrt{1-\theta_2^2} + \sqrt{1-\theta_3^2} + \lambda \cdot \left(\theta_1^2 + \theta_2^2 + \theta_3^2 - r^2\right)$$

Its partial derivatives are equal to zero at the minima, so

$$2\lambda\theta_i - \frac{\theta_i}{\sqrt{1-\theta_i^2}} = 0, \quad i = 1, 2, 3$$

There are two different types of solutions to the above equations: $\theta_i = 0$, or $\theta_i^2 =$\{constant depending on λ\}. So we get 3 different solutions from the derivative criteria

$$\theta_1 = \theta_2 = 0 \text{ and } \theta_3^2 = r^2 \tag{27}$$

$$\theta_1 = 0 \text{ and } \theta_2^2 = \theta_3^2 = \frac{r^2}{2} \tag{28}$$

$$\theta_1^2 = \theta_2^2 = \theta_3^2 = \frac{r^2}{3} \tag{29}$$

If we examine these cases, we obtain that at (27) there are minimum places, at (28) there are saddles, and at (29) there are maximum places. □

Theorem 3 *With the optimal choice of the directions and the number of measurements the variance of estimator can decrease to*

$$\mathrm{Tr}V^n_{optimal}(\theta) = \frac{1}{n}\left(2 + \sqrt{1-r^2}\right)^2. \tag{30}$$

Proof: The statement is the consequence of Theorem 2 if we use such a coordinate system, that one of the axes is in the direction of θ. Then the state will be $[r, 0, 0]$ in the new coordinate system, so using (25) we get (30). □

Remark: The optimal case is equivalent to Example 3. The improvement from the above modification can be seen on Figure 1 and the efficiency grows up to $\frac{2}{3}$, if the state is on the surface.

3.2 Adaptive algorithm

In the realistic case we have an unknown θ, so the optimal measurement directions and measurement ratios are unknown, too. Still the following statement is true:

Theorem 4 *The optimal variance (30) from Theorem 3 is achievable asymptotically.*

Proof: We will give a measurement strategy for all N so that if N goes to infinity the limit of the total variance will be (30).

We need to make at least a draft, "orientation" estimate of the variance that can be followed by a refining step that applies the optimal measurement number ratios. According to Theorem 2, we can achieve the best result if one of the measurement directions points to the direction of the true Bloch vector. Therefore, if one rotates the measurement directions after the initialization step driven by the draft estimate $\widetilde{\theta}$, then the final estimation step can be performed by using the new directions and performing the measurements with the optimal ratio.

With the above strategy we can easily get an asymptotically optimal algorithm. Let be the length of initialization step N_{init}. Then a sufficient condition is $N_{init} \to \infty$ and $\frac{N_{init}}{N} \to 0$ (for example by choosing $N_{init} = \sqrt{N}$). Then $\widetilde{\theta}$ will converge to θ, and our estimation will be optimal with the number of measurements $N - N_{init}$. Since the variance is a continuous function of the measurement direction and ratios and $\frac{N_{init}}{N} \to 0$ so the variance goes to (30).

There is only one problem with this scheme. Namely it can happen that our estimation of θ may fall outside of the Bloch ball, so $\|\widetilde{\theta}\|_2 > 1$. In this case $r > 1$, so according to Theorem 1, the optimal ratio will contain non-real numbers which is obviously not possible. If our state is mixed ($\|\theta\|_2 < 1$) then this problem does not arise, because the probability that $\widetilde{\theta}$ fall outside of a circle with radius ε and center θ goes to zero according to the weak law of large numbers. So if ε is small enough then the probability of falling outside the Bloch ball ($\|\widetilde{\theta}\|_2 > 1$) goes to 0, too.

If the state is pure ($\|\theta\|_2 = 1$), then the probability of falling outside the Bloch ball ($\|\widetilde{\theta}\|_2 > 1$) does not go to zero. So we apply the following modification to the above algorithm: We know that the estimator $\widetilde{\theta}$ is inside of the circle with radius $\varepsilon(N)$ and center θ with $1 - \delta(N)$ probability. So if we use a translation by $\varepsilon(N)$ on the estimator $\widetilde{\theta}$ in the direction of the center of the Bloch ball, then this estimator ($\hat{\theta}$) will be inside of the Bloch ball at least with $1 - \delta(N)$ probability. If $\varepsilon(N)$ is constant then $\delta(N)$ goes to 0 according to the weak law of large numbers. It is then possible with a little change that $\varepsilon(N)$ goes to zero while $\delta(N)$ goes to 0, too. In this case we get an estimator ($\hat{\theta}$) that is

inside of the Bloch ball with asymptotically 1 probability, and the estimator $\hat{\theta}$ goes to θ (because $\varepsilon(N) \to 0$). So because of the continuity this estimation will be asymptotically optimal. □

Let us compare this result with the standard method (22). Unfortunately we can not compare the above result with the minimal qubit tomography because the authors did not present any analytical result for the variance, only numerical ones for special θ-s. The comparison can be seen in Figure 3.

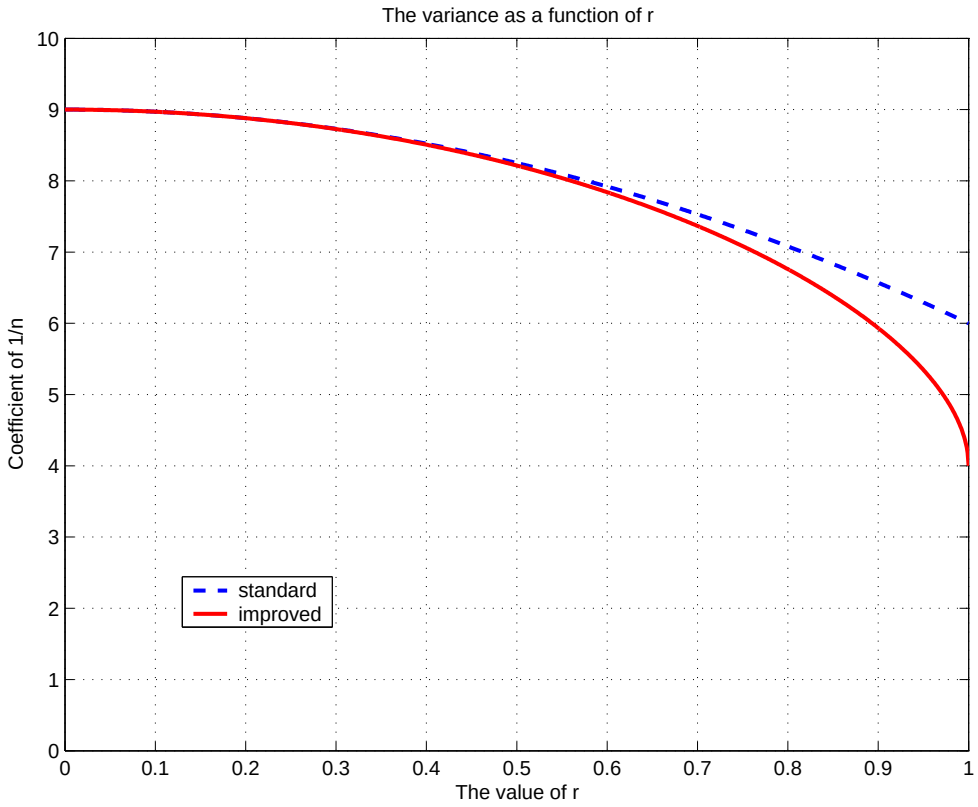

Figure 3: Comparison of different methods

Our estimation scheme is better than the standard one everywhere. Furthermore, when we are closer to the surface the estimation is more efficient. That is because the algorithm uses the asymmetry in the number of measurements and near the surface the asymmetry is greater.

3.3 Numerical approaches for finite number of measurements

An almost surely asymptotically optimal measurement scheme was introduced in Theorem 4, but in reality there is always a finite number of measurements available. So this

section contains the investigation of how the finite case changes the theoretical results.

In finite case we loose efficiency because the length of the initialization part is not negligible. Therefore, an important question is how to select the length N_{init}. Another crucial part of the algorithm described in Theorem 4 is that in finite case the draft estimation goes outside of the Bloch ball with positive probability, which needs to be handled. Here I give some possible algorithms to threat these problems.

In finite case the calculations of variance are really complicated so I investigate the cases by simulations in MATLAB [15] to approximate the real variance of different estimation schemes and compare it with the theoretical result (30). During the simulations the whole estimation process was performed 10000 times and the mean square error was computed. The simulations contain 3 different cases of the total number of measurements: $N = 1000,\ 3000$, and 5000. These numbers come from the bound of computational capacity.

In order to investigate the behavior of the algorithm on different real states, we perform simulations on 3 different states:

- $\theta = [0.95, 0, 0]$: a mixed state in the axis direction
- $\theta = [0.8, 0.5, 0.1]$: a mixed state in a general direction
- $\theta = [\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}]$: a pure state in a general direction

In all cases $\|\theta\|_2$ is chosen to be big because then the difference from the standard method will be significant. The results of the simulations are interpreted in tables with the same structure. In the first row the variance of standard method computed from (22) can be found. This is followed by the mean square errors obtained from the simulations. Finally the theoretical results from Theorem 4 are presented, computed from (30).

3.3.1 Using optimal ratios without rotation

The steps of the algorithm include an initialization giving a draft estimate using the standard scheme ($m = \frac{N_{init}}{3}$), that is followed by a final estimate.

In the initialization step a pre-defined function of the total allowable measurement number N is used: $N_{init} = N^{\alpha}$. If $0 < \alpha < 1$ then N_{init} fulfills the conditions discussed in Theorem 4. In our simulation experiments the values $\alpha = 0.75$ and $\alpha = 0.9$ are applied. The optimal ratio of the measurement numbers n_1, n_2 and n_3 (where $n_1 + n_2 + n_3 = N$) was calculated from Theorem 1. In the final estimation step we measure in each direction $n_i - m$ times, i.e. we take the number of measurements of the initialization phase also into account.

In addition, there is a trade-off between the length of the initialization step determining the variance of the draft estimate $\widetilde{\theta}$ and that of the final estimation step. If the initialization step is too short, then the draft estimation will be imprecise, so the optimal ratio will be far from $n_1,\ n_2,\ n_3$. On the other hand if the initialization step is too long then we can easily "over-measure" in some direction, i.e $n_i < m = \frac{N_{init}}{3}$, so we can not use the optimal ratios in the correcting step because we must measure $n_i - m$ times more in each direction, and in this case it is a negative number.

$[0.95, 0, 0]$	N=1000	N=3000	N=5000
standard	0.00629	0.00210	0.00126
$\alpha = 0.75$	**0.00553**	**0.00186**	**0.00109**
$\alpha = 0.9$	**0.00527**	**0.00179**	**0.00106**
optimal	0.00535	0.00178	0.00107

$[0.8, 0.5, 0.1]$	N=1000	N=3000	N=5000
standard	0.0063	0.0021	0.00126
$\alpha = 0.75$	**0.00636**	**0.00202**	**0.00124**
$\alpha = 0.9$	**0.00620**	**0.00202**	**0.00122**
optimal	0.00536	0.00179	0.00107

$[\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}]$	N=1000	N=3000	N=5000
standard	0.006	0.002	0.0012
$\alpha = 0.75$	**0.00616**	**0.00203**	**0.00121**
$\alpha = 0.9$	**0.00600**	**0.00201**	**0.00119**
optimal	0.004	0.00133	0.0008

Table 1: Simulation results using optimal ratios without rotation

The results of simulation can be seen in Table 1. Our first observation is that the efficiency does not depend well on the length on initialization step, i.e from α. But we get a bit better results for the $\alpha = 0.9$ case. According to Theorem 2, if $\theta = [0.95, 0, 0]$ then the variance is close to the optimal, but in the $\theta = [0.8, 0.5, 0.1]$ and $\theta = [\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}]$ cases the results are far from the optimal result because then the real state is not close to an axis. In the last case when $\theta_1 = \theta_2 = \theta_3$ the variance of our method will be equal to the standard case.

3.3.2 Using rotation with simplification

From Theorem 2 we know that if we want results to be close to the optimal variance (30), then we need to rotate the axis of measurements. In reality this can be done by rotating some mirrors (as in [20]) or using a Hamiltonian to rotate the state. The point is that this can add some extra time to the measurement process, but if N is big this will be irrelevant.

I will use the measurement scheme described in Theorem 4. The length of initialization step will be $N_{init} = N^{\alpha}$ with $\alpha = \frac{2}{3}$ and $\alpha = \frac{3}{4}$. Then I rotate a measurement direction in the direction of draft estimate $\widetilde{\theta}$. Then the optimal ratio (24) will change to

$$\widetilde{n}_1 : \widetilde{n}_2 : \widetilde{n}_3 = \sqrt{1 - \left(\widetilde{\theta}_1^2 + \widetilde{\theta}_2^2 + \widetilde{\theta}_3^2\right)} : 1 : 1, \tag{31}$$

because in the new coordinate system the state of system is approximately $(\|\widetilde{\theta}\|_2, 0, 0)$. If the draft estimation is outside of the Bloch ball($\|\widetilde{\theta}\|_2 > 1$), then in (31) will be a nonreal (i.e. a complex) number. Let us use notation $\varphi(\widetilde{\theta})$ for the non-trivial element of ratio. The simplest way is to handle the problem $\|\widetilde{\theta}\|_2 > 1$, that in this complex case we use a constant ratio, i.e. $\varphi(\widetilde{\theta}) = c$. In our simulations we use $c = 0.3$ and $c = 0.1$. Then we measure another $N - N_{init}$ times and get a new estimate $\widehat{\theta}$ for the state of the qubit. To avoid the loss of information from initialization step and get an unbiased estimate we will use the $\frac{N_{init}}{N} \cdot \widetilde{\theta} + \frac{N - N_{init}}{N} \cdot \widehat{\theta}$ estimator during simulations.

$[0.95, 0, 0]$	N=1000	N=3000	N=5000
standard	0.00629	0.00210	0.00126
α=$\frac{3}{4}$,c=0.1	**0.00593**	**0.00189**	**0.00112**
α=$\frac{2}{3}$,c=0.1	**0.00611**	**0.00205**	**0.00117**
α=$\frac{3}{4}$,c=0.3	**0.00552**	**0.00186**	**0.00111**
α=$\frac{2}{3}$,c=0.3	**0.00568**	**0.00188**	**0.00112**
optimal	0.00535	0.00178	0.00107

$[0.8, 0.5, 0.1]$	N=1000	N=3000	N=5000
standard	0.0063	0.0021	0.00126
α=$\frac{3}{4}$,c=0.1	**0.00588**	**0.00205**	**0.00125**
α=$\frac{2}{3}$,c=0.1	**0.00600**	**0.00199**	**0.00120**
α=$\frac{3}{4}$,c=0.3	**0.00557**	**0.00193**	**0.00116**
α=$\frac{2}{3}$,c=0.3	**0.00559**	**0.00186**	**0.00114**
optimal	0.00536	0.00179	0.00107

$[\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}]$	N=1000	N=3000	N=5000
standard	0.006	0.002	0.0012
α=$\frac{3}{4}$,c=0.1	**0.00487**	**0.00157**	**0.000935**
α=$\frac{2}{3}$,c=0.1	**0.00473**	**0.00155**	**0.000930**
α=$\frac{3}{4}$,c=0.3	**0.00494**	**0.00162**	**0.000987**
α=$\frac{2}{3}$,c=0.3	**0.00494**	**0.00162**	**0.000964**
optimal	0.004	0.00133	0.0008

Table 2: Simulation results using rotation with simplification

From the simulation results (see Table 2) we can find that this case is not sensitive to α, too. In case $\theta = [0.95, 0, 0]$ and $\theta = [0.8, 0.5, 0.1]$ the $c = 0.3$ is more efficient, but in case $\theta = [\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}]$, the $c = 0.1$ value will have better variance. This corresponds to the simple fact that the value of $\varphi(\theta)$ is closer to 0.3 in the first two cases, while in last

case $\varphi(\theta)$ is closer to 0.1. In all three cases the optimal variance is well approximated but these variances attached to different parameters in different cases.

3.3.3 Using rotation with numerical Bayes estimation

From the previous part it is clear that for different θ values different $\varphi(\theta)$ values are optimal, so to obtain a better estimate we need to make an estimation from $\widetilde{\theta}$ to θ even if $\|\widetilde{\theta}\| > 1$. We will use the numerical Bayes estimation to this purpose . We calculate the probabilities to some element of a Cartesian grid in such a way that the total number of the elements will be constant, so this calculation requires constant time even for huge N values, thus it will be negligible. We only take into account the elements of the grid that are inside of the Bloch ball so in this way we get an estimator $\widetilde{\theta}^*$ which is always inside of the Bloch ball. In simulation we use $\varphi(\widetilde{\theta}) = \nu \cdot \sqrt{1 - \left((\widetilde{\theta}_1^*)^2 + (\widetilde{\theta}_2^*)^2 + (\widetilde{\theta}_3^*)^2\right)}$, with $\nu = \frac{1}{2}$ and $\nu = \frac{1}{4}$. The other details of the measurement scheme is the same as in the previous scheme.

$[0.95, 0, 0]$	N=1000	N=3000	N=5000
standard	0.00629	0.00210	0.00126
α=$\frac{2}{3}$, ν=$\frac{1}{2}$	**0.00571**	**0.00190**	**0.00112**
α=$\frac{2}{3}$, ν=$\frac{1}{4}$	**0.00610**	**0.00204**	**0.00120**
α=$\frac{3}{4}$, ν=$\frac{1}{2}$	**0.00566**	**0.00187**	**0.00113**
α=$\frac{3}{4}$, ν=$\frac{1}{4}$	**0.00590**	**0.00194**	**0.00116**
optimal	0.00535	0.00178	0.00107

$[0.8, 0.5, 0.1]$	N=1000	N=3000	N=5000
standard	0.0063	0.0021	0.00126
α=$\frac{2}{3}$, ν=$\frac{1}{2}$	**0.00557**	**0.00189**	**0.00114**
α=$\frac{2}{3}$, ν=$\frac{1}{4}$	**0.00593**	**0.00204**	**0.00123**
α=$\frac{3}{4}$, ν=$\frac{1}{2}$	**0.00552**	**0.00196**	**0.00120**
α=$\frac{3}{4}$, ν=$\frac{1}{4}$	**0.00577**	**0.00210**	**0.00131**
optimal	0.00536	0.00179	0.00107

$[\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}]$	N=1000	N=3000	N=5000
standard	0.006	0.002	0.0012
α=$\frac{2}{3}$, ν=$\frac{1}{2}$	**0.00484**	**0.00158**	**0.000936**
α=$\frac{2}{3}$, ν=$\frac{1}{4}$	**0.00476**	**0.00155**	**0.000928**
α=$\frac{3}{4}$, ν=$\frac{1}{2}$	**0.00485**	**0.00158**	**0.000946**
α=$\frac{3}{4}$, ν=$\frac{1}{4}$	**0.00475**	**0.00159**	**0.000931**
optimal	0.004	0.00133	0.0008

Table 3: Simulation results using rotation with numerical Bayes estimation

The data of Table 3 show that the situation did not change too much. The length of the initialization does not affect well the efficiency, while for the $\theta = [0.95, 0, 0]$ and $\theta = [0.8, 0.5, 0.1]$ cases the bigger ν is more efficient, but in case of $\theta = [\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}]$

the smaller ν will lead us to better result. However the differences have been reduced compared to the previous algorithm.

3.3.4 Comparison of different algorithms

Instead of comparing the 10 different measurement schemes (see the rows of Table 1-3) let us choose the most efficient ones.

In case of no rotation the $\alpha = 0.9$ case is the most efficient. Similarly, if we regard the case with rotation using numerical Bayes estimation we can state that the $\nu = 0.5$ case has better variances, let us choose $\alpha = \frac{2}{3}$ (so the first data row of Table 3). Finally, in the case with rotation using simplification there are 2 nearly optimal but fundamentally different cases with $c = 0.3$ and $c = 0.1$. Let us choose in both cases $\alpha = \frac{2}{3}$ (so the second and fourth data row of Table 2). And finally we will plot the optimal variance (30).

To get comparable results for different θ and N values we do not plot the variances of each cases but the efficiencies, i.e we normalize the values with the variance of the standard method.

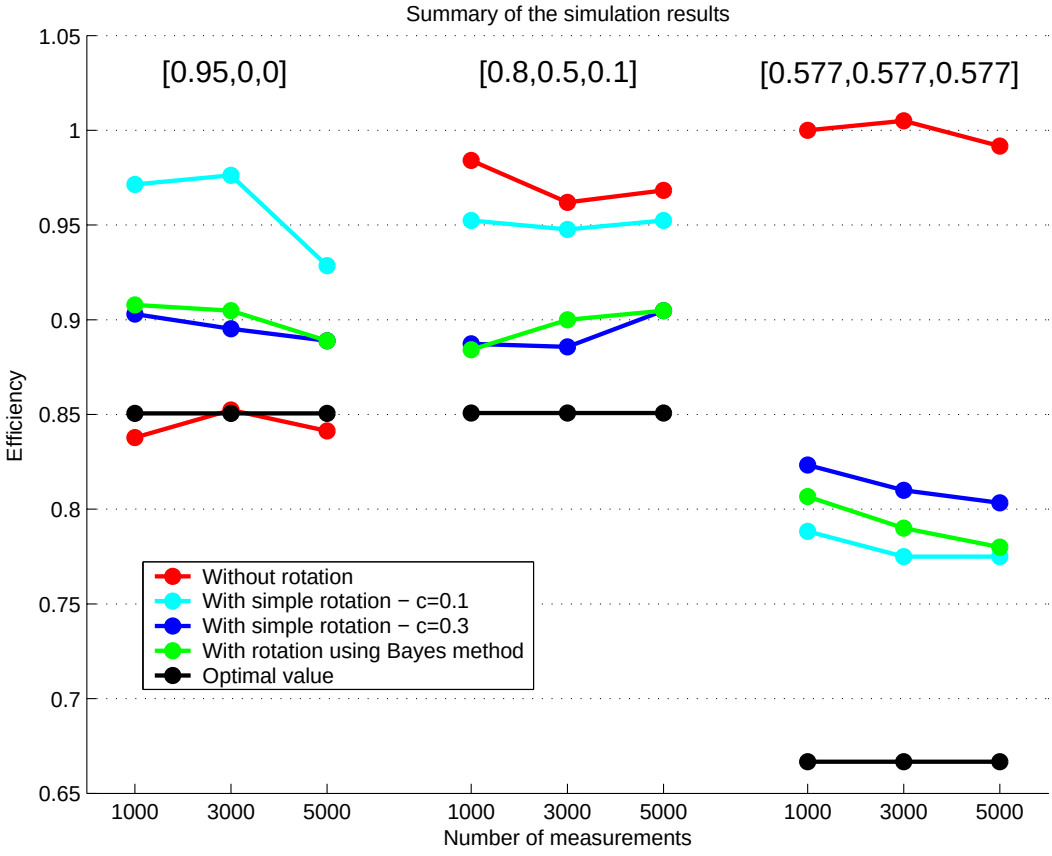

Figure 4: Comparison of efficiencies of different algorithms

The results can be seen in Figure 4. The efficiency of the method without rotation does not depend on the purity (from $\|\theta\|_2$), but on the distance from the axis. If the real state is in the direction of an axis ($\theta = [0.95, 0, 0]$), then the achieved variance is the smallest, in other cases it is bigger. If we use rotation then the key determining factor is the purity of the state. In the first two cases the value of $\|\theta\|_2$ is roughly the same just their direction are different, and indeed all three rotating measurement scheme has similar efficiency even by different values of θ. In the case of simplification there is a recognizable difference between the light and dark blue marks. The numerical Bayesian approach is a kind of compromising result, near the minimum of blue lines. Probably with a better choice of $\varphi(\widetilde{\theta})$ the variance can show further improvement but this is outside of the scope of this thesis.

The fact is visible too, that the estimation of pure state is more complicated than estimation of a mixed state since there is bigger gap between the simulation and theoretical results. Note that all of these methods are more efficient than the standard one. All of the methods begin with an initialization part so they can be combined to use of their advantage, for example if we get that $\widetilde{\theta}$ is close to an axis then we do not use rotation etc.

4 Indirect measurement scheme using coupled qubits

Indirect measurement means that the projective measurements are performed on the probe system (being in state θ_M) attached to the one we are interested in (θ_S). In the composite system (in state ρ_{S+M}) an indirect measurement corresponds to the observables of the form $I \otimes A_M$, where A_M is a self adjoint operator on the Hilbert space of system M. For the sake of simplicity, it is assumed, that A_M is a Pauli spin operator. The

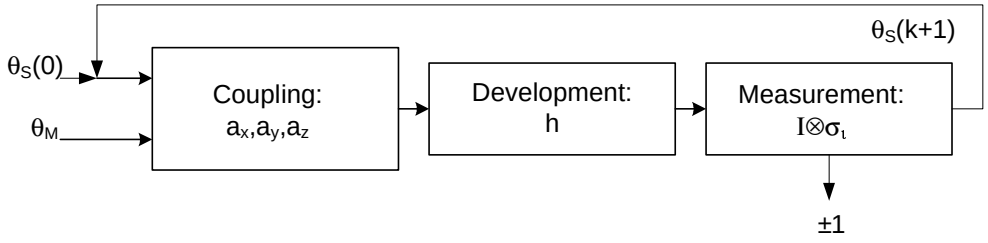

Figure 5: Signal flow diagram of indirect measurement

measurement strategy is shown in Fig. 5. At each time instant of the discrete time set, the measurement qubit is coupled to the unknown system S. They evolve according to the bipartite dynamics (13) for the sampling time h, and at the end of the sampling interval, a von Neumann measurement is performed on the measurement qubit. At the next time instant, the previous steps are repeated.

The above setting of the indirect measurement allows us to adjust various parameters of the measurement strategy. These can be and will be used to make an optimal compromise between the information gained from the measurement and the demolition caused by the measurement back-action.

The coupling parameters a_1, a_2, a_3 of the Cartan decomposition (11-12) determine how (in terms of strength and direction) the measurement system is coupled to the unknown one. The sampling time h amplifies this effect and appears as a multiplicative factor to the coupling parameters.

The state of the measurement qubit (θ_M) can be different at each time instant which allows us in the future to introduce a feedback to the measurement protocol.

It is important to note that one can make a 'no information - no demolition' situation by setting the coupling parameters to zero, and a 'maximal information - complete demolition' situation, too. Examples of such extreme cases will be given in the subsection 4.1.2 and 4.2.2.

4.1 A simple example for indirect measurement

A simple special case of an indirect measurement is investigated analytically here to brighten the effect of the protocol parameters. As we shall see later, this case can be used to selectively estimate one of the components of the unknown qubit's Bloch vector, similarly to the so called standard measurement scheme [18] for single qubits. A straightforward modification of the measurement setup leads to the estimators of the other two Bloch vector components.

4.1.1 Measurement setup

Suppose in the sequel that the qubits are interacting only in the y direction for time h (sampling time). Afterwards, an indirect measurement is performed, i.e. a von Neumann measurement of the observable $I \otimes \sigma_x$ on the composite system with state

$$e^{-\mathrm{i}h(a_2\sigma_2^S \otimes \sigma_2^M)} \cdot (\rho_S \otimes \rho_M) \cdot e^{-\mathrm{i}h(a_2\sigma_2^S \otimes \sigma_2^M)^*},$$

For the sake of simplicity, choose h and a_2 in such a way, that $2a_2h = 2\pi$. The above setting corresponds to the parameters

$$h = \frac{1}{10}, \quad a_2 = \frac{5}{2}\pi, \quad a_1 = a_3 = 0.$$

The probabilities of the different outcomes of $I \otimes \sigma_x$'s measurement are

$$\begin{aligned} \mathrm{Prob}(+1) &= \tfrac{1}{2}(1 + \theta_{S2}\theta_{M3}) \\ \mathrm{Prob}(-1) &= \tfrac{1}{2}(1 - \theta_{S2}\theta_{M3}). \end{aligned} \tag{32}$$

Now the probabilities depend on both the state of the unknown θ_S and that of the measurement qubit θ_M. The post-measurement states are

$$\theta_S^{+1} = \begin{bmatrix} \frac{\theta_{S3}\theta_{M2} + \theta_{S1}\theta_{M1}}{1+\theta_{S2}\theta_{M3}} \\ \frac{\theta_{S2} + \theta_{M3}}{1+\theta_{S2}\theta_{M3}} \\ \frac{\theta_{S3}\theta_{M1} - \theta_{S1}\theta_{M2}}{1+\theta_{S2}\theta_{M3}} \end{bmatrix}, \quad \theta_S^{-1} = \begin{bmatrix} \frac{\theta_{S3}\theta_{M2} - \theta_{S1}\theta_{M1}}{1-\theta_{S2}\theta_{M3}} \\ \frac{\theta_{S2} - \theta_{M3}}{1-\theta_{S2}\theta_{M3}} \\ \frac{-\theta_{S3}\theta_{M1} - \theta_{S1}\theta_{M2}}{1-\theta_{S2}\theta_{M3}} \end{bmatrix} \tag{33}$$

if $+1$ or -1 was the result, respectively. This measurement setup is useful for unknown state estimation since the probabilities and the new states depend on both θ_S and θ_M. This means, that we both gain information from the measurements and retrieve information in the new states after the measurement.

4.1.2 Properties of repeated indirect measurements

Let us concentrate on the estimate of the second unknown state co-ordinate, i.e. we want to describe the change of θ_{S2} (notation: $x = \theta_{S2}$) during the measurements. Let us further assume that θ_{M3} is constant (denoted by c) and $\theta_{M1} = \theta_{M2} = 0$, i.e. $\theta_M = [0, 0, c]^T$.

If $c = 1$ then we get the standard measurement scheme, where

$$\text{Prob}(\pm 1) = \frac{1}{2}(1 \pm x) \quad , \quad x^{\pm 1} = \pm 1$$

It is easy to see from (33) that this would be a totally invasive measurement, i.e. the information about the true state would be lost, thus we assume $|c| < 1$.

Propostion 1 *If we measure first +1, and thereafter -1 (or vice versa), then the state of θ_{S2} (=x) will not change.*

Proof: After the first measurement, x changes to $x^{+1} = \frac{x+c}{1+cx}$ then from x^{+1} it turns to be $x^{+1,-1} = \frac{x^{+1}-c}{1-cx^{+1}} = \frac{\frac{x+c}{1+cx}-c}{1-c\frac{x+c}{1+cx}} = \cdots = x$. The reverse goes similarly.

Corollary 2 *All of the possible cases of states can be ordered in a line such a way, that after each measurement we jump in the neighboring state on the left or right side.*

Propostion 2 *If we measure first +1, and thereafter -1 (or vice versa), then the probability of these outcomes doesn't depend on x.*

Proof: First from x it will be $x^{+1} = \frac{x+c}{1+cx}$ with probability: $P = \frac{1}{2}(1 + cx)$ then from x^{+1} it will be x with probability: $Q = \frac{1}{2}(1 - cx^{+1}) = \frac{1}{2}(1 - c\frac{x+c}{1+cx}) = \frac{1}{2}\frac{1-c^2}{1+cx}$. So the probability of this outcome is $P \cdot Q = \frac{1-c^2}{4}$. The reverse goes similarly.

Corollary 3 *If two outcome sequences contain the same number of $+1$ and -1 measurement outcomes, then their probabilities are the same.*

Corollary 4 *The probability that there are k times $+1$ and l times -1 outcomes ($k > l$) in the sequence can be computed as:*

$$\left(\frac{1-c^2}{4}\right)^l \cdot p_{k-l}$$

and the state after this sequence of outcomes will be x_{k-l}, where p_n is the probability of that from n measurements all outcomes are $+1$s, and x_n is the resulting state from n measurement when all outcomes are $+1$s. So we can represent the sequence of the measurement outcomes as a Markov-process.

Propostion 3 *p_k is a linear function of x.*

Proof: The proof goes by induction. Let $p_k := \frac{q_k}{2^k}$ and $x_k := \frac{y_k}{q_k}$, where (we will prove) q_k and y_k are simple polynomials of c and x.

If $k = 0$ then $q_0 = p_0 = 1$ and $y_0 = x$.

Next let us suppose that both q_k and y_k are simple polynomials, and p_k is linear in x. Then

$$p_{k+1} = p_k \cdot \frac{1}{2}(1 + cx_k) = \frac{q_k}{2^k} \cdot \frac{1}{2}\left(1 + c\frac{y_k}{q_k}\right) = \frac{1}{2^{k+1}} \cdot (q_k + cy_k).$$

On the other hand $p_{k+1} = \frac{q_{k+1}}{2^{k+1}}$, so

$$q_{k+1} = q_k + cy_k. \tag{34}$$

Furthermore,

$$x_{k+1} = \frac{x_k + c}{1 + cx_k} = \frac{\frac{y_k}{q_k} + c}{1 + c\frac{y_k}{q_k}} = \frac{y_k + cq_k}{q_k + cy_k} = \frac{y_k + cq_k}{q_{k+1}}.$$

On the other hand $x_{k+1} = \frac{y_{k+1}}{q_{k+1}}$, therefore

$$y_{k+1} = y_k + cq_k. \tag{35}$$

Finally we conclude that q_k and y_k are really simple polynomials, and from the recursion it can be seen that both q_k and y_k are linear in x, so p_k is linear in x, too.

The proof gives us a recursive calculation for x_k and p_k, so it is possible to build up a stochastic model based on the above 3 propositions, and develop a state estimation strategy.

4.2 Towards optimal quantum state estimation by indirect measurements

Let us suppose that we have N identical copies of the composite quantum system (the two coupled qubits, S and M). We shall use the following measurement strategy:

1. Perform 2 subsequent measurements (a *measurement pair*) on each copy with a pre-specified $c = \theta_{M3}$ and compute the maximum-likelihood (ML) estimate of x.

2. Retain the copies on which the measured outcomes were $+1$ and -1 (in any order) for further studies, because they are not affected by the measurements, i.e. their $\theta_{S2} = x$ is left unchanged (see Proposition 1).

Note that the above implies $n = 2$ for the results in sub-section 4.1.2. Now we investigate how the selection of c (the initial state of the measurement system) affects the variance of the estimate (we want it to be small), and the ratio of the un-affected system copies (we want this to be large).

Denote the number of the $(+1,+1)$ outcomes by N_+, and the probability that a measurement pair result in this outcome by $p_+ = p_2 = \frac{1+c^2+2cx}{4}$. Similarly, the number of the $(-1,-1)$ outcome is denoted by N_-, and its probability is by $p_- = \frac{1+c^2-2cx}{4}$. Then the number of the non-effective ($(+1,-1)$ or $(-1,+1)$) outcomes is $N_0 = N - N_+ - N_-$, and its probability is $p_0 = \frac{1-c^2}{2}$. Then the likelihood function of N measurement pairs is the following polynomial distribution:

$$P = \frac{N!}{N_+!\ N_-!\ N_0!}\ p_+^{N_+}\ p_-^{N_-}\ p_0^{N_0} \tag{36}$$

The maximum likelihood estimate of x is obtained by taking the logarithm of P in (36), and maximizing it with respect to x:

$$\hat{x}_{ML}(N_+, N_-, c) = \frac{1+c^2}{2c}\frac{N_+ - N_-}{N_+ + N_-} \tag{37}$$

This estimate is well-defined if at least one of N_+ or N_- is positive, that holds with probability one when number of measurements goes to infinity. On the other hand, this estimate is asymptotically unbiased.

4.2.1 The variance and the non-demolition probability

In the case of the investigated measurement setup (see section 4.1.1), the variance $\mathbf{V}_N$ of the Maximum Likelihood estimator (37) is as follows:

$$\begin{aligned}\mathbf{V}_N(c,x) &= \sum_{i=1}^{N} Var\left(\tfrac{1+c^2}{2c}\tfrac{N_+-N_-}{N_++N_-} \mid N_+ + N_- = i\right)\cdot \mathrm{Prob}(N_+ + N_- = i) = \\ &= \sum_{i=1}^{N}\left(\tfrac{1+c^2}{2c}\right)^2 \tfrac{1}{i^2}\, Var\left(N_+ - N_- \mid N_+ + N_- = i\right)\cdot \mathrm{Prob}(N_+ + N_- = i)\end{aligned} \tag{38}$$

where $Var(.)$ denotes the variance of a random variable.

Let be X_j a random variable that takes the value $+1$ if the outcome of the measurement pair is $(+1,+1)$, and -1 when the outcome is $(-1,-1)$. Then $X_j = 1$ with probability $\frac{p_+}{p_+ + p_-}$, and $X_j = -1$ with probability $\frac{p_-}{p_+ + p_-}$. These are the conditional properties of being $+1$,$+1$ and -1,-1, if we know that the two outcome is the same. Then

$$Var\left(N_+ - N_- \mid N_+ + N_- = i\right) = Var\left(\sum_{j=1}^{i} X_j\right) = i\cdot Var(X_1).$$

From simple calculation we obtain:

$$Var(X_1) = 1 - \left(\frac{p_+ - p_-}{p_+ + p_-}\right)^2 = 1 - \left(\frac{2cx}{1+c^2}\right)^2$$

Therefore, the variance of the Maximum Likelihood estimator is

$$\mathbf{V}_N(c,x) = \left(\frac{1+c^2}{2c}\right)^2 \left[1 - \left(\frac{2cx}{1+c^2}\right)^2\right] \sum_{i=1}^{N} \frac{1}{i} \cdot \mathrm{Prob}(N_+ + N_- = i),$$

$$\text{and} \quad \sum_{i=1}^{N} \frac{1}{i} \cdot \mathrm{Prob}(N_+ + N_- = i) = \mathbb{E}\left(\frac{1}{N_+ + N_-}\right) \sim \frac{1}{N(p_+ + p_-)},$$

where $\mathbb{E}$ denotes the mean value, and $\sim$ stands for asymptotic equality. Thus we obtain:

$$\lim_{N\to\infty} N\mathbf{V}_N(c,x) = \frac{(c+1/c)^2 - 4x^2}{2(1+c^2)} = W(c,x)$$

The other important aim would be to minimize the disturbed system instances, i.e. the cases when the outcomes are $(+1,+1)$, or $(-1,-1)$. The probability of having such outcomes is $p(c,x) = \frac{1}{2}(1+c^2)$. Note that $W(c,x)$ can be regarded as the asymptotic variance originating from a qubit, and $p(c,x)$ as the probability that the state of qubit will remain unchanged during the estimation process.

4.2.2 Optimal measurement strategy

If one wants to have a compromising strategy, then a possible way is to minimize the expression

$$\Psi(c,x) = \min_c \left[\, A \cdot W(c,x) + (1-A) \cdot p(c,x) \,\right], \tag{39}$$

where $A \in \mathbb{R}^+$ is a normalized parameter $(1 \geq A \geq 0)$ which determines our trade-off strategy. If $A \approx 1$, then the aim is accuracy, while in the case of $A \approx 0$ we aim at minimal demolition.

Figure 6 shows the substantial part of the loss function $\Psi(c,x)$ over the domain $(-1 \leq x \leq 1)$, $(0.2 \leq c \leq 1)$. Note that the function is symmetric to the $c = 0$ line, but it is indefinite at $c = 0$. It is seen that there is a definite optimal value $c \approx 0.6$ for the initial state of the measurement qubit in case $A = 0.1$ that is the same for every x. In the case of $A = 0.9$, however, the minimum is taken at $c = 1$, i.e. at the complete demolition situation.

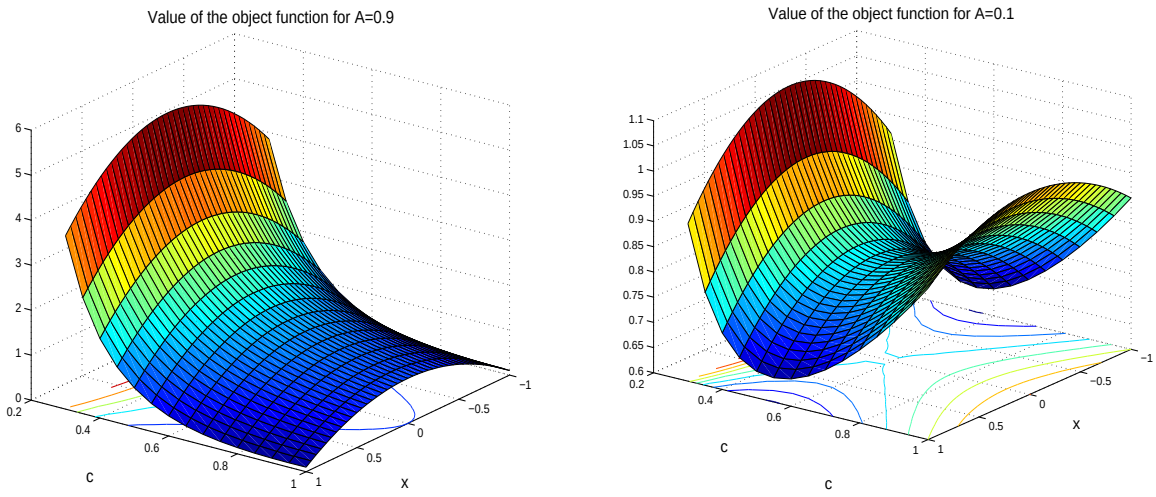

Figure 6: The optimal measurement qubit state for different A values: more information ($A = 0.9$, left) versus more non-demolished system ($A = 0.1$, right)

4.2.3 Comparison with the standard method

It is possible to compare the above estimation scheme with the so-called standard qubit tomography described in subsection 2.4.1, which uses information obtained from the von Neumann measurements of the three Pauli matrices.

For the sake of reasonability, it is assumed that the number of measurements m for the standard qubit tomography equals to the expected value of the changed qubits in indirect scheme, i.e. $m = N \cdot p(c, x)$. This way, the expected number of destroyed qubits will be the same in the two compared methods.

The variance of the standard method is $V_m^{stan} = \frac{1}{m} \cdot (1 - x^2)$ if the number of measurements is m. In this case $m = N \cdot \frac{1+c^2}{2}$, so $W^{stan}(c, x) = N \cdot V_m^{stan} = \frac{2(1-x^2)}{1+c^2}$. We can define efficiency η as the quotient of $W(c, x)$ and $W^{stan}(c, x)$, i.e.

$$\eta = \frac{W(c,x)}{W^{stan}(c,x)} = \frac{(c+1/c)^2 - 4x^2}{2(1+c^2)} \cdot \frac{1+c^2}{2(1-x^2)} = \frac{\frac{1}{4}(c+1/c)^2 - x^2}{1-x^2} \geq 1,$$

because $|c + 1/c| \geq 2$, equation holds if $c = \pm 1$.

The above result clearly shows, that the standard method is more accurate than the indirect one. Note, that if $c = \pm 1$, the two methods are the same, i.e. the standard qubit tomography is the special case of the indirect method (see Figure 7).

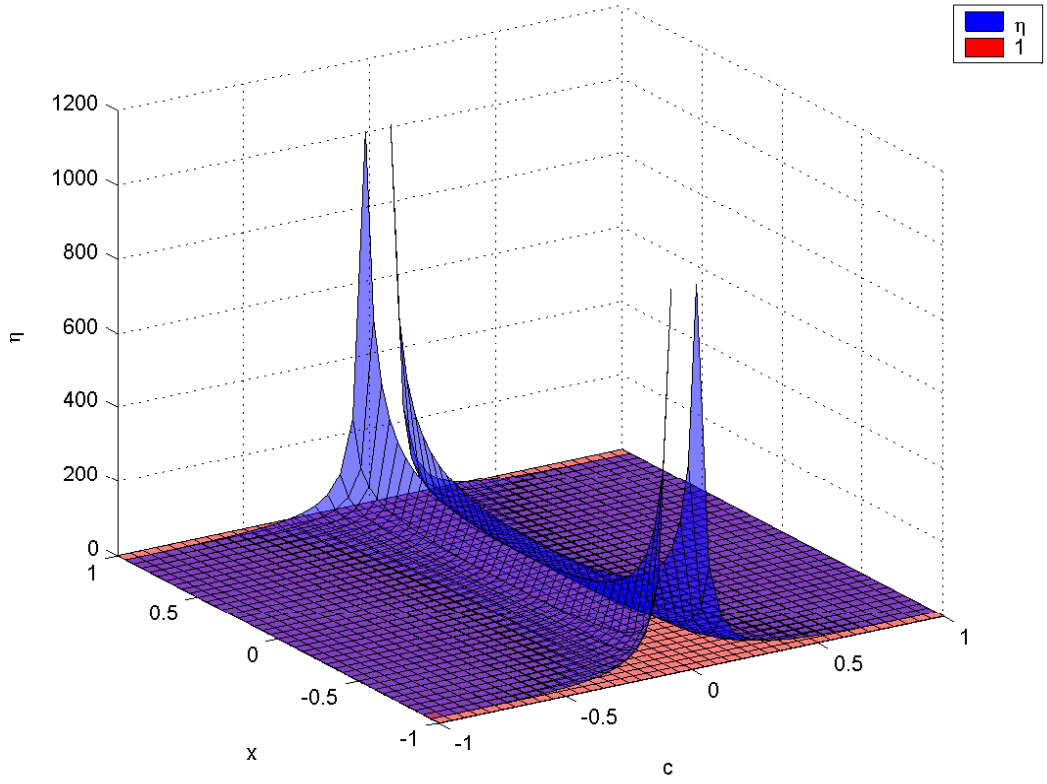

Figure 7: The efficiency η (blue) compared to constant one (red).

4.3 Modified methods

This section deals with possible modifications of the original method in order to improve its efficiency.

4.3.1 Recycling the unchanged qubits

The idea in the following is to use the qubits which remained unchanged during the indirect measurement to obtain more information about the state. So the measurement procedure is continued on the unchanged qubits until all of them is changed.

The variance of this modified method comes as the special case of the variance (38) of the method defined in section 4.1.1 with $i = N$. It is easy to see, that

$$\mathbf{V}_N^{recyc}(c,x) = Var\left(\frac{1+c^2}{2c}\frac{N_+ - N_-}{N_+ + N_-} \mid N_+ + N_- = N\right) = \frac{1}{N}\cdot\left[1-\left(\frac{2cx}{1+c^2}\right)^2\right].$$

The variance of standard method for N qubits is $\mathbf{V}_N^{stan}(c,x) = \frac{1}{N}\cdot(1-x^2)$. Computing

their quotient, it is

$$\eta^{recyc} = \frac{W^{recyc}(c,x)}{W^{stan}(c,x)} = \frac{\mathbf{V}_N^{recyc}(c,x)}{\mathbf{V}_N^{stan}(c,x)} = \frac{1-\left(\frac{2cx}{1+c^2}\right)^2}{1-x^2} = \frac{1-\left(\frac{x}{\frac{c+1/c}{2}}\right)^2}{1-x^2} \geq 1,$$

since $|c+1/c| \geq 2$. It means, that the indirect method cannot overcome the standard one even in the case when all the available qubits are measured.

However, it is expected that the modified estimation is more effective, than the original one. The fact, that $\eta = \eta^{recyc} \cdot \frac{1}{4}(c+1/c)^2$ supports these expectations, i.e. $\eta \geq \eta^{recyc}$.

4.3.2 Measuring four times

As another way of modification of the method described in subsection 4.1.2, it is possible to change n, i.e. the number of subsequent measurements performed on the coupled qubits. In what follows, n is assumed to be 4. This implies that the number of possible outcomes are $2^4 = 16$, but according to Proposition 1 and Proposition 2 the outcomes like $\{+---\}$, $\{-+--\}$, $\{--+-\}$ and $\{---+\}$ are indistinguishable and correspond to p_{-2} (see Figure 8).

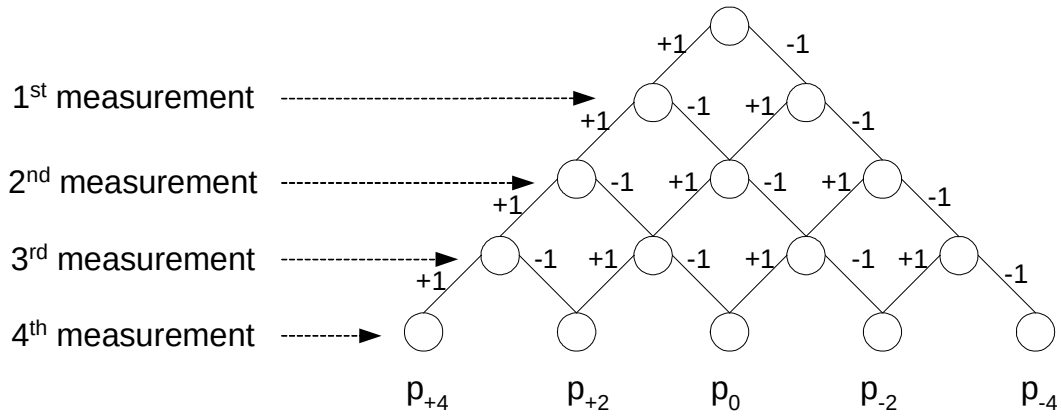

Figure 8: The possible outcome-combinations result in five different probabilities $p_{+4}, p_{+2}, p_0, p_{-2}, p_{-4}$.

The probabilities of such groups of outcomes can be determined using the result of Corollary 4 multiplied by the appropriate binomial coefficients:

$$p_{+4} = \tfrac{1}{16}\left(1+6c^2+c^4+4c\left(1+c^2\right)x\right)$$

$$p_{-4} = \tfrac{1}{16}\left(1+6c^2+c^4-4c\left(1+c^2\right)x\right)$$

$$\begin{aligned} p_{+2} &= \tfrac{1}{4}(1-c^2)(1+c^2+2cx) \\ p_{-2} &= \tfrac{1}{4}(1-c^2)(1+c^2-2cx) \\ p_0 &= \tfrac{3}{8}(1-c^2)^2 \end{aligned} \tag{40}$$

Using the probabilities (40) it is easy to construct estimators similar to (37):

$$\hat{x}_1(N_2, N_{-2}, c) = \frac{1+c^2}{2c}\frac{N_2 - N_{-2}}{N_2 + N_{-2}} \tag{41}$$

$$\hat{x}_2(N_4, N_{-4}, c) = \frac{1+6c^2+c^4}{4c(1+c^2)}\frac{N_4 - N_{-4}}{N_4 + N_{-4}} \tag{42}$$

Estimators (41) and (42) are independent, thus it is possible to use a linear combination of them as an estimation:

$$\hat{x}(N_2, N_{-2}, N_4, N_{-4}, c) = B \cdot \hat{x}_1(N_2, N_{-2}, c) + (1-B) \cdot \hat{x}_2(N_4, N_{-4}, c), \tag{43}$$

where $0 \leq B \leq 1$. In order to have an efficient estimator (43), it's variance

$$B^2 \cdot Var(\hat{x}_1) + (1-B)^2 \cdot Var(\hat{x}_2)$$

should be minimal. The covariance has been omitted because of the independency of estimators (41) and (42). Similarly to the derivation in subsection 4.2.1, the variances of the independent estimators are obtained as:

$$W^{\hat{x}_1}(c, x) = \frac{1 + c^4 + c^2(2 - 4x^2)}{2c^2 - 2c^6}$$

and

$$W^{\hat{x}_2}(c, x) = \frac{1}{2c^2} + \frac{2}{(1+c^2)^2} - \frac{8x^2}{1+6c^2+c^4},$$

using $W = \lim_{N\to\infty} N \cdot Var$. The minimal value of the variance is where its derivative is 0. It is reached with the following value of B:

$$B = \frac{(-1+c^2)\left(-(1+6c^2+c^4)^2 + 16c^2(1+c^2)^2 x^2\right)}{2\left(1 + c^2(10 + 24c^2 - 2c^4 - c^6 + 2(1+c^2)(-5-6c^2+3c^4)x^2)\right)}$$

The optimal value of the variance is then

$$W^{four}(c, x) = \frac{\left((1+c^2)^2 - 4c^2x^2\right)\left(1+6c^2+c^4-4c(1+c^2)x\right)\left(1+6c^2+c^4+4c(1+c^2)x\right)}{4c^2\left(1 + c^2\left(11 + 34c^2 + 22c^4 - 3c^6 - c^8 + 2(1+c^2)^2(-5-6c^2+3c^4)x^2\right)\right)}.$$

Since our other aim is to change as few qubits as possible, it is necessary to write up the ratio of changed qubits. This is simply the complementary probability of being unchanged, i.e.

$$p^{four}(c,x) = 1 - p_0 = 1 - \frac{3}{8}\left(1-c^2\right)^2$$

The efficiency is the following:

$$\eta^{four} = \frac{W^{four}(c,x)\cdot\left(1-\frac{3}{8}\left(1-c^2\right)^2\right)}{1-x^2}.$$

Figure 9 shows, that the relation $1 \leq \eta^{four} \leq \eta$ holds, i.e. the modification has slightly improved the performance of the estimator, but it is still worse than the standard one and the estimator formula is much more complicated. It can be seen in the case of both the

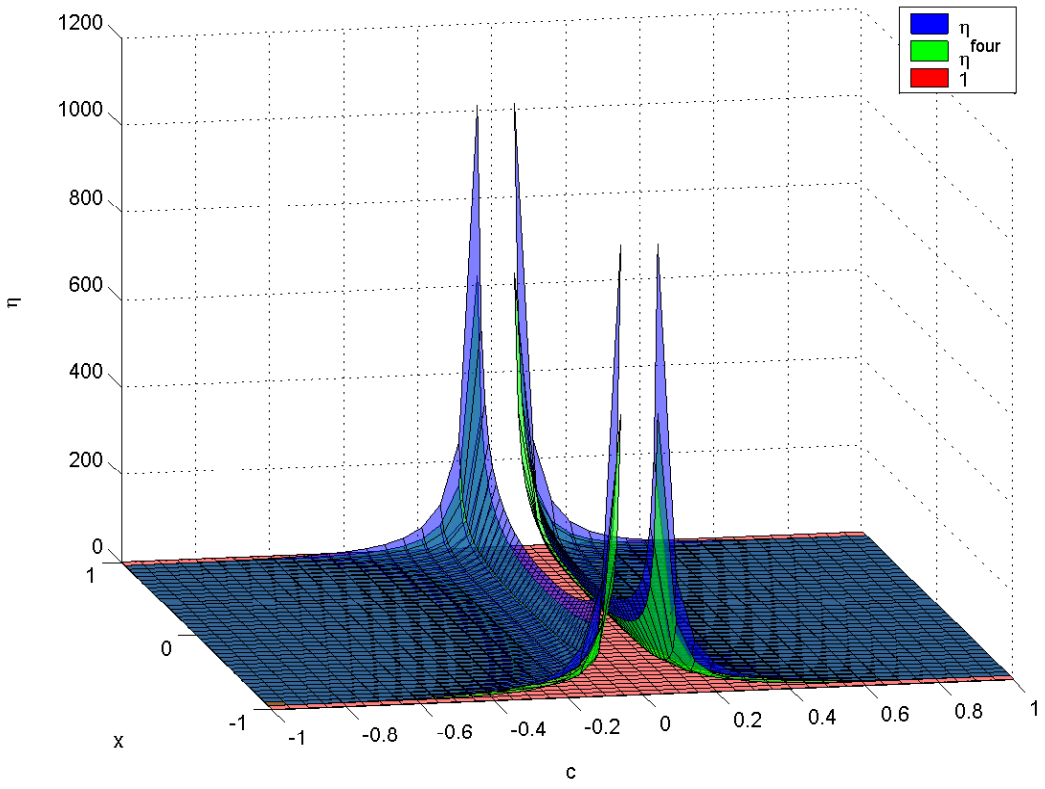

Figure 9: The quotients η (blue) and η^{four} (green) compared to constant one (red).

original and the modified estimator's η, that for a given measurement state with $c \neq \pm 1$, η is minimal, if $x = 0$. This means, that the estimators can achieve smaller variance (better performance) in the case of mixed states.

4.3.3 Undoing the measurement

In this case we use the method described in Section 4.2, perform the estimation, and then we re-measure the changed qubits until they get back to the initial state, i.e. until there is an equal number of measuring +1 and -1 outcomes.

Each path that returns in $2k$ steps to the origin has the probability $\left(\frac{1-c^2}{4}\right)^k$ (the number of + and - outcomes is both k). At the same time the number of pathes that return to the initial state first in the $2k$-th step is exactly $2 \cdot C_{k-1}$, where C_k is the k-th Catalan number. Let us use the notation $r = \frac{1-c^2}{4}$, so the probability that the state will ever return to the initial state, i.e we can undo the measurement, is

$$1 - p^{undo}(c,x) = \sum_{k=1}^{\infty} 2 \cdot C_{k-1}\ r^k = 2 \sum_{k=0}^{\infty} C_k\ r^{k+1} = 2r \sum_{k=0}^{\infty} C_k\ r^k.$$

The finally received sum is the generating function of Catalan numbers, and it has a well known closed formula:

$$\sum_{k=0}^{\infty} C_k\ r^k = \frac{1-\sqrt{1-4r}}{2r}$$

from that

$$1 - p^{undo}(c,x) = 2r \cdot \frac{1-\sqrt{1-4r}}{2r} = 1 - \sqrt{1-4r} = 1 - \sqrt{1-4\frac{1-c^2}{4}} = 1 - |c|$$

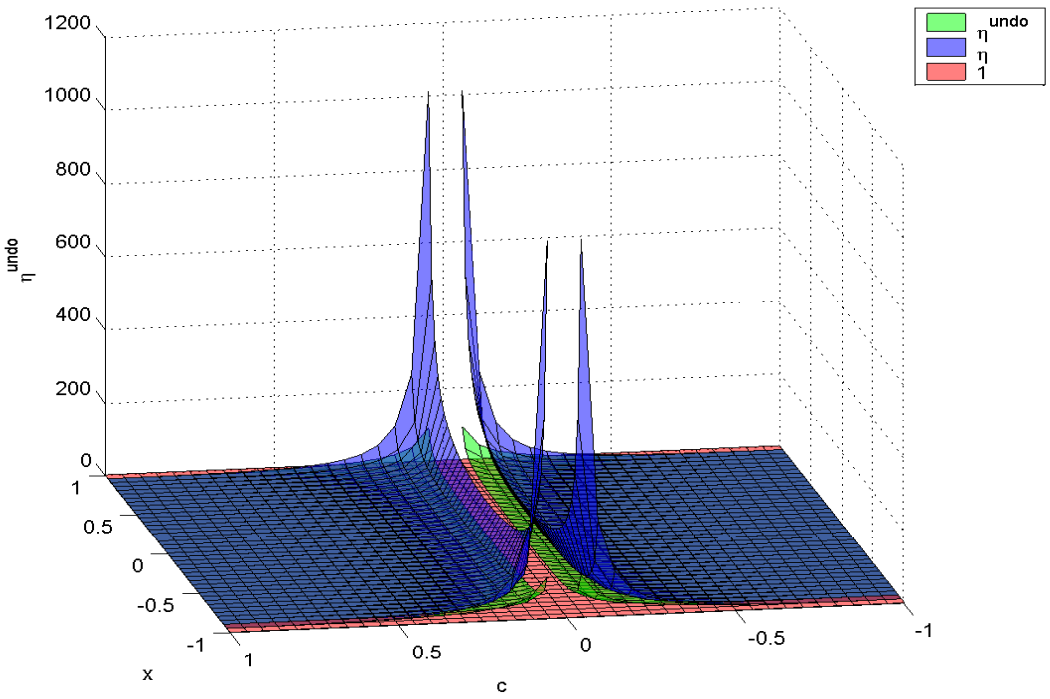

Figure 10: The efficiency η^{undo}

and we get that the probability that the qubit never returns $p^{undo}(c,x) = |c|$. So according the previous section ($W^{undo}(c,x) = W(c,x)$) the efficiency is the following (see also Figure 10):

$$\eta^{undo} = \frac{(c+1/c)^2 - 4x^2}{2(1+c^2)} \cdot \frac{|c|}{1-x^2}$$

The greatest improvement is at small c values, but the original efficiency (η) is the worst exactly here. Therefore, the situation is like in previous cases: we have achieved an improvement in the efficiency but it is still under the efficiency of the standard method, i.e. $1 < \eta^{undo} < \eta$ (see Figure 10). Note that this calculation gives back the trivial fact, that in the standard case ($c = 1$) we never return to the original state: $p^{undo} = 1$.

5 Summary

In this thesis several state estimation methods for a single qubit have been developed and examined. Wooters and Fields showed that the observable set consisting orthogonal observables results in the most information [24], so many people think that the standard scheme is optimal.

In Chapter 3 some kind of generalization of the standard scheme was proposed. The variance of the estimation with different number of measurements was investigated and it was shown that we can achieve better variance if we do not measure the same number in each direction but in a determined ratio which depends on the real state of the qubit. Moreover, the variance will be optimal if one of the observables is in the direction of the state of the quantum bit. Since the state of qubit is unknown, we have proposed a two-step algorithm to achieve the optimal variance asymptotically with probability 1. This way we can improve the variance if the state is not $[0, 0, 0]$, while the modification is more efficient when the state is more closer to the surface of Bloch ball. That is because the method is using asymmetry, and near the pure states the asymmetry increases. Some numerical methods for the finite number of measurement case have also been developed in order to help the application of the method in realistic circumstances.

Chapter 4 presents indirect qubit tomography methods in the discrete-time case [22]. The statistical properties of the estimate in terms of the variance of the ML estimator and the non-demolition probability have been analytically calculated in the case of subsequent measurements applied in the x direction while the qubits interact in the y direction and the initial Bloch-vector of the measurement qubit is $[0, 0, c]$. A way of finding an optimal compromising measurement strategy between the asymptotic variance and the non-demolition probability has been proposed. The efficiency of the results have been compared with a classical 'standard' state estimation procedure available in the literature. Although the classical one performs better by means of the variance, the indirect one gives a degree of freedom in the above mentioned trade-off problem. The estimation method has also been modified in a few ways to improve its precision. It has been shown that the modified measurement strategies may reach the efficiency of the standard method in the limit when the complete demolition situation is achieved. This is in good agreement with the results of D'Ariano and Yuen [3].

Acknowledgement

I would like to thank Professor Dénes Petz and Attila Magyar for their help and especially to Professor Katalin Hangos for her support and encouragement.

References

[1] E. Bagan, M.A. Ballester, R.D. Gill, A. Monras and R. Munoz-Tapia, Optimal full estimation of qubit mixed states, *Phys. Rev. A*, 73, 032301, 2006.

[2] V.P. Belavkin. Towards the theory of control in observable quantum systems. *Automatica and Remote Control*, 44, 178–188, 1983.

[3] G.M. D'Ariano and H.P. Yuen. Impossibility of measuring the wave function of a single quantum system. *Physical Review Letters*, 76, 2832–2835, 1996.

[4] L. Diósi. Weak measurement in bipartite quantum mechanics. `arXiv: quant-ph/0505075v1`, 2007.

[5] D.G. Fisher and M. Freyberger, Estimating mixed quantum states, *Physics Letters A*, 273, 293-302, 2000.

[6] M. Hayashi, K. Matsumoto, Asymptotic performance of optimal state estimation in quantum two level system, `arXiv, quant-ph/0411073v2`, 2006

[7] C. W. Helstrøm, *Quantum decision and estimation theory*, Academic Press, New York, 1976.

[8] C. Hill and J. Ralph. Weak measurement and rapid state reduction in bipartite quantum systems. `arXiv:quant-ph/0610156v2`, 2006.

[9] M. Hotta, T. Karasawa, and M. Ozawa, Ancilla-assisted enhancement of channel estimation for low-noise parameters, *Phys. Rev. A*, 72, 052334, 2005.

[10] K. Jacobs and D. A. Steck. A straightforward introduction to continuous quantum measurement. `arXiv:quant-ph/0611067v1`, 2006.

[11] A.N. Jordan, B. Trauzettel, and G. Burkard. Weak measurement of quantum dot spin qubits. `arXiv:0706.0180v1[cond-mat.mes-hall]`, 2007.

[12] N. Khaneja and S. J. Glaser. Cartan decomposition of su(2^n), constructive controllability of spin systems and universal quantum computing. arXiv: quant-ph/0010100v1, 2000.

[13] A.N. Korotkov and A.N. Jordan. Undoing a weak quantum measurement of a solid-state qubit. *Physical Review Letters (arXiv:quant-ph/0610156v2)*, 97, 166805, 2006.

[14] R. L. Kosut, I. Walmsley, H. Rabitz, Optimal experiment design for quantum state and process tomography and Hamiltonian parameter estimation, arXiv, quant-ph/0411093, 2004

[15] Mathworks Inc, MATLAB Software System, http://www.mathworks.com/, 2001

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

[17] D. Petz, *Quantum information theory and quantum statistics*, Springer, Berlin and Heidelberg, 2008

[18] D. Petz, K.M. Hangos, A. Magyar, Point estimation of states of finite quantum systems, *J. Phys. A: Math. Theor.*, 40, 7955-7969, 2007

[19] D. Petz, K.M. Hangos, L. Ruppert, Quantum state tomography with finite sample size, *Quantum Bio-Informatics, From Quantum Information to Bio-Informatics*, World Scientific, 247-257, 2008

[20] J. Rehacek, B.-G. Englert, D Kaszlikowski, Minimal qubit tomography, *Physical Review A*, 70:052321, 2004

[21] R. Romano and D. D'Alessandro. Incoherent control and entanglement for two-dimensional coupled systems. arXiv:quant-ph/0510020v1, 2005.

[22] L. Ruppert, A. Magyar, K. M. Hangos, D. Petz, Towards optimal quantum state estimation of a qubit by using indirect measurements. Technical report of the Systems and Control Laboratory SCL-002/2008, Budapest, MTA SZTAKI, 2008

[23] G.A. Smith, A. Silberfarb, I.H. Deutsch, and P.S. Jessen. Efficient quantum state estimation by continuous weak measurement and dynamical control. arXiv: quant-ph/0606115v1, 2006.

[24] W.K. Wootters and B.D. Fields, Optimal state determination by mutually unbiased measurements, *Annals of Physics*, 191, 363–381, 1989.