UNIVERSITY OF CALIFORNIA
RIVERSIDE

Compressed Sensing Quantum Process Tomography for Superconducting Quantum Gates

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Electrical Engineering

by

Andrey Rodionov

March 2015

Dissertation Committee:

Dr. Alexander N. Korotkov, Chairperson
Dr. Roger Lake
Dr. Nissim Amos

The Dissertation of Andrey Rodionov is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

Acknowledgments

I am grateful to my advisor Professor Alexander N. Korotkov without whose help I would not be here. His teaching and explanations, his intuitive view of the problems and material, his way of asking the substantial questions about the research topics to help me gain a better understanding of the material have proven invaluable to me during my research. I am extremely grateful for Professor Korotkov's patience and support in those times when I could not understand things that seemed relatively simple to him. Honestly Professor, I can't thank you enough for your overall kindness, I will never forget the ways in which you've helped me.

I have had the pleasure to work with a number of brilliant postdocs in our research group. In particular Drs. Andrzej Veitia, Eyob A. Sete and Justin Dressel, to whom I am infinitely grateful for providing help in understanding the physical concepts behind my work and also the computer software I was using throughout my work, Mathematica and Matlab, and teaching me some technical aspects of LaTeX.

Particular thanks to former postdoc Dr. Andrzej Veitia who gave me valuable ideas about calculating statistical characteristics in my work and suggested consideration of the square for the average state fidelity and also provided immense help in the preparation of the manuscript of my published paper.

I want to thank former graduate student Dr. Kyle Keane for helping me to understand the basics of QPT. Eric Mlinar, thank you for your feedback and motivation to constantly improve my work. Thank you to all of my coworkers for providing a friendly and welcoming environment in the office.

I am also grateful to Dr. Robert L. Kosut for explaining the ideas of compressed sensing quantum process tomography and for his consultations about the details of the Matlab scripts that I was writing.

I am grateful to current and former postdocs and graduate students in the research group of Professor John M. Martinis at UCSB: Dr. Matteo Mariantoni, Dr. Rami Barends, Dr. Daniel Sank, Jim Wenner, Julian Kelly for providing the experimental data that I was using in my theoretical work. All of the characteristics of the two-qubit CZ gate that are presented in the dissertation have been calculated from experimental data they shared.

Thank you Professors Roger Lake and Nissim Amos for being the members of my dissertation committee and attending my dissertation defense.

I am also grateful to the funding agencies, which provided funding for the research of this dissertation: it was funded through the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), through the Army Research Office Grant No. W911NF-10-1-0334. I also acknowledge support from the ARO MURI Grant No. W911NF-11-1-0268.

The text of this dissertation, in part or in full, is an adaptation of the material as it appears in the following journal [1]: Physical Review B, vol. **90**, article 144504, Oct. 2014. The co-author Alexander N. Korotkov, listed in the above publication, directed and supervised the research which forms the basis for this dissertation. The remaining co-authors listed provided scientific and technical expertise and support as collaborators.

Finally, I am grateful to the University of California, Riverside and to the Department of Electrical Engineering for providing me the opportunity to be a graduate student and do my research here.

I dedicate this dissertation to my parents Valeriy and Helena Rodionov for giving me the belief that through hard work and perserverance I can accomplish my goals.

ABSTRACT OF THE DISSERTATION

Compressed Sensing Quantum Process Tomography for Superconducting Quantum Gates

by

Andrey Rodionov

Doctor of Philosophy, Graduate Program in Electrical Engineering
University of California, Riverside, March 2015
Dr. Alexander N. Korotkov, Chairperson

An important challenge in quantum information science and quantum computing is the experimental realization of high-fidelity quantum operations on multi-qubit systems. Quantum process tomography (QPT) is a procedure devised to fully characterize a quantum operation. We first present the results of the estimation of the process matrix for superconducting multi-qubit quantum gates using the full data set employing various methods: linear inversion, maximum likelihood, and least-squares. To alleviate the problem of exponential resource scaling needed to characterize a multi-qubit system, we next investigate a compressed sensing (CS) method for QPT of two-qubit and three-qubit quantum gates. Using experimental data for two-qubit controlled-Z gates, taken with both Xmon and superconducting phase qubits, we obtain estimates for the process matrices with reasonably high fidelities compared to full QPT, despite using significantly reduced sets of initial states and measurement configurations. We show that the CS method still works when the amount of data is so small that the standard QPT would have an underdetermined system of equations. We also apply the CS method to the analysis of the three-qubit Toffoli gate with simulated noise, and similarly show that the method works well for a substantially reduced set of data. For the

CS calculations we use two different bases in which the process matrix is approximately sparse (the Pauli-error basis and the singular value decomposition basis), and show that the resulting estimates of the process matrices match with reasonably high fidelity. For both two-qubit and three-qubit gates, we characterize the quantum process by its process matrix and average state fidelity, as well as by the corresponding standard deviation defined via the variation of the state fidelity for different initial states. We calculate the standard deviation of the average state fidelity both analytically and numerically, using a Monte Carlo method. Overall, we show that CS QPT offers a significant reduction in the needed amount of experimental data for two-qubit and three-qubit quantum gates.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

"In theory, theory and practice are the same. In practice, they are not."

– Albert Einstein

Quantum information science is a relatively young interdisciplinary research field that combines ideas and methods of physics, mathematics, information theory and computer science [2]. Quantum information science is based on the control and use of quantum systems like photons, atoms, trapped ions or superconducting qubits to process and transmit information. It has been shown that certain quantum algorithms offer tremendous advantages over classical algorithms in a number of special computational tasks. For example, the efficient quantum algorithm of factorization of a large integer number n was invented by Peter Shor in 1994 [3, 4, 5, 6]. Shor's algorithm is performed in polynomial time in $\log n$ and is exponentially more efficient than any known classical algorithm. Specifically, it takes time $O((\log n)^3)$, which is much faster than the best known classical factoring algorithm for which the time scales as $O(e^{1.9(\log n)^{1/3}(\log\log n)^{2/3}})$. Shor's algorithm could be used to break public-key cryptography schemes such as the widely used RSA scheme [7]

which is based on the assumption that factoring large numbers is computationally infeasible on a classical computer. Another famous quantum algorithm, proposed by Lov Grover [8] in 1996, addresses the problem of database searching. The computational time of Grover's algorithm equals the square root of time of the fastest classical algorithm. Another well-known quantum protocol, formulated by Artur Ekert [9] in 1991, allows secure quantum key distribution. It uses entangled pairs of photons and allows detection of eavesdropping by testing Bell's inequality violation.

In classical information theory, information is encoded in strings of bits, which can either be in the state 0 or 1. The physical realization of a classical bit can be an arbitrary physical system with two states, for example, the p – n junction diode: switched-off (0) and switched-on (1). Analogously, in quantum computing, a unit of quantum information is called a "qubit" or "quantum bit", which is a two-level quantum mechanical system. Unlike its classical counterpart, the quantum bit can be in any superpositions of its basis states $|0\rangle$ and $|1\rangle$ (for example, the ground and excited states of an atom), which is the manifestation of the phenomenon of quantum superposition. Examples of physical realization of qubits can be the spin of the electron, or electron in a quantum dot, or polarized photon, etc. A quantum computer uses qubits to store and process information. A quantum logic gate is a device (a basic quantum circuit) that performs a fixed unitary operation on selected qubits [10].

The object of the research of this dissertation are quantum gates implemented with the superconducting qubits, which are currently considered one of the most promising platforms for quantum computing (see, e.g., [11]). Superconducting qubits are electrical resonators with a nonlinearity that is strong enough so that microwaves excite transitions only between the two lower states, and this is why they are often called "artificail atoms". Superconducting qubits include such

elements as Josephson junctions, capacitors and inductors. Nonlinearity in such nonlinear LC-resonators is due to the presence of Josephson junctions. We give a review of basic notions and concepts from the quantum information science and discuss in details various types of superconducting qubits in Chapter 2.

Experimental realizations of quantum gates usually differ from their ideal theoretical models due to small imperfections and experimental errors in the laboratory. Quantum tomography describes methods to identify non-ideal quantum states and quantum gates (or, equivalently, quantum channels, quantum processes, or quantum operations) in the laboratory. Quantum State Tomography [2] (QST) is a method by which a quantum state is measured, and Quantum Process Tomography (QPT) is a procedure by which an unknown quantum operation can be fully characterized experimentally [2, 12, 13]. The main idea of QST is the following: for a quantum state produced by a preparation device, a set of different measurements has to be performed on the ensemble of identical quantum states on the output of that device, in order to estimate all the parameters of the produced quantum state. These parameters form the density operator (equivalent name is "density matrix") of the state, which fully describes the state of a physical system. We describe the procedure of QST and present our results of QST for the phase qubits in Chapter 3.

The methods of quantum process tomography provide us with the opportunity to predict the evolution of a quantum state propagating through an imperfect quantum gate. In real experiments, when the interaction between the qubit system and the environment cannot be neglected, it is not possible to describe the evolution of a quantum state by a unitary operator. The experimenter attempts to apply an ideal quantum gate U, but what really occurs is a noisy quantum operation $\mathcal{E}$, which can be characterized by the methods of quantum process tomography. The main idea be-

hind QPT is the following: for a quantum channel (process), that takes an input state and outputs some transformed state, a combination of preparation of various linearly independent input states and measurements of the multiple copies of the corresponding output states is used to determine the quantum channel (process) parameters. The parameters that characterize the quantum channel form the process matrix χ, which contains all accessible information about the quantum process. Note that while the dimension of the unitary operator U, describing the ideal evolution of a N-qubit quantum system, is $2^N \times 2^N$, the dimension of a process matrix χ, which describes the non-unitary evolution in the presence of interaction with the environment, is $4^N \times 4^N$. The role of QPT in experimental characterization of quantum gates is twofold. First, it allows us to quantify the quality of the gate; that is, it tells us how close the actual and desired quantum operations are. Second, QPT may aid in diagnosing and correcting errors in the experimental operation [14, 15, 16, 17, 18]. The importance of QPT has led to extensive theoretical research on this subject (e.g., [19, 20, 21, 22, 23, 24]). Therefore the quantum process tomography is an essential tool in reliable quantum information processing.

Various methods exist for the estimation of the process matrix χ from experimental data. One of the methods of χ-matrix estimation from the set of measurement results is the so-called linear inversion procedure, in which the process matrix is expressed as a matrix multiplication of several other matrices. This method is described for the systems of one and two qubits in [2, 13, 25], and it is formulated slightly differently in [17]. Using the expressions for the process matrix from [17], we wrote a Matlab code that calculates the process χ-matrix for the evolution of quantum systems consisting of several qubits. We established that the maximum number of qubits in the system that allows the exact computation of the χ-matrix on an average personal computer equals six. The

details of the method of linear inversion for the reconstruction of the process matrix are given in Chapter 4.

As statistical and systematic errors are usually unavoidable in experiments, the density or process matrix obtained directly from measurement results (using the so-called "linear inversion" method) can often result in parameters that do not have a physical meaning, for example, they would correspond to negative probabilities, or probabilities greater than one, or violate the positive semidefinite or trace-preservation property, in other words the obtained estimator of either the density or the process matrix cannot be used for statistical predictions. One of the possible methods to avoid the problem of non-physical results is to use the maximum likelihood estimator [26, 27], which is a convex optimization problem [27, 28, 29]. Convex optimization problems are optimization problems, where the objective function and the constraints are convex. In convex optimization, any local minimum must be a global minimum, which makes it much easier to deal with a convex optimization problem in comparison to a non convex problem, as it may have multiple local optimal points. Convex optimization has a variety of applications in science, engineering, and finance. In the maximum likelihood approach to QST or QPT tasks, the likelihood of reproducing the experimental results is maximized, where the likelihood function is a function of the parameters of the measurement model of the quantum system, and such method guarantees the density or process matrix to be theoretically valid while giving a close fit to the measured experimental data. The specific mathematical form of the likelihood function can be chosen as a log-likelihood function, or as the least squares function, or some other function. In this dissertation, we provide results for the log-likelihood QST estimation of a two-qubit state of the system of two superconducting qubits in Chapter 3, and also provide results for the least squares QPT estimation of the process matrix for the systems of both two and three qubits in Chapter 5.

Although conceptually simple, QPT suffers from a fundamental drawback: the number of required experimental configurations scales exponentially with the number of qubits (e.g., [30]). Even for few-qubit systems, QPT involves collecting large amounts of tomographic data and heavy classical postprocessing. Specifically, an N-qubit quantum operation can be represented by a $4^N \times 4^N$ process matrix χ containing 16^N independent real parameters (or $16^N - 4^N$ parameters for a trace-preserving operation) which can be determined experimentally by QPT. To alleviate the problem of exponential scaling of QPT resources, alternative methods have been developed, e.g., randomized benchmarking [31, 32, 33] and Monte Carlo process certification [34, 35]. These protocols, however, find only the fidelity of an operation instead of its full process matrix. Both randomized benchmarking and Monte Carlo process certification have been demonstrated experimentally for superconducting qubit gates (see [36, 37, 38] and references therein). Although these protocols are efficient tools for the verification of quantum gates, their limitation lies in the fact that they do not provide any description of particular errors affecting a given process and therefore they cannot be used to improve the performance of the gates.

Recently, a new approach to QPT which incorporates ideas from signal processing theory has been proposed by Kosut, Shabani and coworkers [39, 40]. The basic idea is to combine standard QPT with compressed sensing (CS) theory [41, 42, 43, 44], which asserts that sparse signals may be efficiently recovered even when significantly undersampled. Compressed sensing quantum process tomography (CS QPT) enables one to recover the process matrix χ from far fewer experimental configurations than standard QPT. The method proposed in [39, 40] is hoped to provide an exponential speed-up over standard QPT. In particular, for a d-dimensional system the method is supposed to require only $O(s \log d)$ experimental probabilities to produce a good estimate of the

process matrix χ, if χ can be approximated by an s-sparse matrix in some known basis (for comparison, standard QPT requires at least d^4 probabilities, where $d = 2^N$ for N qubits). Note that there are bases in which the process matrix describing the target process (the desired unitary operation) is maximally sparse, i.e. containing only one non-zero element; for example, this is the case for the so-called singular-value-decomposition (SVD) basis [39] and the Pauli-error basis [18]. Therefore, if the actual process is close to the ideal (target) process, then it is plausible to expect that its process matrix is approximately sparse when written in such a basis [40].

The main results of this dissertation, published in Ref. [1] and presented in Chapter 5, are the characterization of quantum gates based on superconducting Xmon and phase qubits, by the method of compressed sensing quantum process tomography. Our research [1] was inspired by the previous work by A. Shabani *et al.*, presented in Ref. [40], where the CS QPT method was experimentally validated for a photonic two-qubit controlled-Z (CZ) gate. In that experiment, sufficiently accurate estimates for the process matrix were obtained via CS QPT using much fewer experimental configurations than the standard QPT: from just 18 and 32 configurations, the authors of [40] reported fidelities of 94% and 97% with process matrices calculated from an overcomplete full set of all 576 available configurations. In this dissertation in Chapter 5 we apply the method of Ref. [40] to several two-qubit CZ gates realized with superconducting qubits. Using the experimental results, we find [1] that CS QPT works reasonably well when the number of used experimental configurations is up to $\sim$7 times less than for standard QPT with the full number of experimental configurations of 144. Specifically, we first present our numerical results for the CS QPT of the superconducting two-qubit CZ gates, based on superconducting Xmon and phase qubits. We also compare numerical results obtained by applying the CS QPT method in two different operator bases,

the Pauli-error basis and the SVD basis. We also compare the performance of the CS QPT method with the least squares optimization, using partial data. We also study the CS QPT of a simulated three-qubit Toffoli gate with numerically added noise, and we find [1] that the reduction factor in the three-qubit case is $\sim$40, compared with standard QPT using the full number of experimental configurations of 1728. In our analysis in Chapter 5 we have primarily used two characteristics. The first characteristic is the comparison between the CS-obtained process matrix $\chi_{\rm CS}$ and the matrix $\chi_{\rm full}$ obtained from the full data set; this comparison is quantitatively represented by the fidelity $F(\chi_{\rm CS}, \chi_{\rm full})$. The second characteristic is how well the CS method estimates the process fidelity F_χ, i.e., how close $F(\chi_{\rm CS}, \chi_{\rm ideal})$ is to the full-data value $F(\chi_{\rm full}, \chi_{\rm ideal})$. Besides calculating the fidelities, we also calculate in Chapter 6 the standard deviation of the fidelity, defined via the variation of the state fidelity for different initial states. We show that this characteristic is also estimated reasonably well by using CS QPT. Chapter 7 is a conclusion. In Appendices we give the list of publications of the author of this dissertation (Appendix A), discuss the Pauli-error basis (Appendix B) and SVD basis (Appendix C).

Chapter 2

Overview of quantum information processing and superconducting qubits

2.1 Overview of quantum information processing

2.1.1 Qubit

According to quantum mechanics, the state of a physical system is represented by a state vector in a Hilbert space, which is a complex vector space with an inner product. In Dirac notation vectors in Hilbert space are denoted by $|v\rangle$, called a "ket-vector". The smallest nontrivial Hilbert space is two-dimensional.

The most basic unit of quantum information is qubit, which is a two-dimensional quantum system described by a state vector in a two-dimensional Hilbert space. Pure states for a qubit are superpositions of the basis states $|0\rangle$ and $|1\rangle$, which means that every state vector can be written as

a linear combination of these two basis states:

$$\big|\Psi\big\rangle = a\big|0\big\rangle + b\big|1\big\rangle, \tag{2.1}$$

where a and b are complex numbers, characterized by a relative phase and subject to the normalization condition $|a|^2 + |b|^2 = 1$. In contrast to the classical bit, which can be only in the states 0 or 1, a qubit can be in an arbitrary coherent superposition of the states $|0\rangle$ and $|1\rangle$, these states are traditionally called *computational basis*. If we measure a qubit in the computational basis, then according to the postulates of quantum mechanics the state of the qubit after measurement will be $|0\rangle$ with probability $|a|^2$, and $|1\rangle$ with probability $|b|^2$. This means if one prepares many copies of the same quantum system in the state $|\Psi\rangle$, a measurement of the state of the qubit will produce the outcome 0 with rate $|a|^2$, and the outcome 1 with rate $|b|^2$. In general, the dimension d of the Hilbert space for a system of N qubits is $d = 2^N$.

Numerous physical realizations of qubits have been proposed and demonstrated, in which the qubits are encoded in the polarization state of a photon [45, 46, 47, 48, 49, 50, 51], in the spin state of a trapped ion [52, 53, 54], in the the magnetic moment of a hydrogen atom in nuclear magnetic resonance [55, 56, 57], in the ground and first excited center-of-mass vibrational states of trapped neutral atoms [58, 59, 60], in the spin state of NV centers in diamond [61, 62, 63, 64], in the states which represent the presence or absence of excess Cooper pairs in the superconducting island (the charge qubit) [65], in the direction of current in a superconducting loop (the flux qubit) [66], or in the lowest two energy levels in the local minimum of the potential energy of the superconducting phase qubit [67, 68]. The superconducting qubits will be discussed in more details in Sections 2.2.2 — 2.2.5 of this dissertation.

In general, a quantum state is mixed. The above example Eq. (2.1) where the quantum state $|\Psi\rangle$ is represented by a state vector (wave function) in a Hilbert space over complex numbers is an example of a very special case of a pure state. A mixed quantum state corresponds to a probabilistic mixture (statistical ensemble) of pure states, and in order to describe it, the formalism of "density operator" needs to be used, which is discussed in the next Section 2.1.2.

When the quantum state, described by Eq. (2.1), interacts with some kind of environment, the state of a qubit may change in an uncontrollable way due to the following two main mechanisms of decoherence. First, if the energy of the state $|1\rangle$ is higher than the energy of the state $|0\rangle$, due to the dissipation and energy loss while interacting with the environment, the qubit's state $|1\rangle$ can decay to the $|0\rangle$ state at a rate typically described by an inverse decay time $1/T_1$. T_1 is usually called "longitudinal coherence time" or the "amplitude damping". The second mechanism of decoherence is dephasing, it can be thought of as noise in the relative phase between the two qubit states, when the relative phase between the two basis states $|0\rangle$ and $|1\rangle$ may change uncontrollably as

$$\big|\Psi\big\rangle = a\big|0\big\rangle + b\big|1\big\rangle \longrightarrow a\big|0\big\rangle + e^{i\theta}b\big|1\big\rangle, \tag{2.2}$$

where θ is the relative phase shift that changes the expectation values of measured quantities. This mechanism of decoherence is called "pure dephasing" and is characterized by the time T_ϕ (sometimes denoted in the literature as T_2'). Actually, the mechanism of "amplitude damping" also contributes to dephasing since it leads to the change and fluctuation of the eigenenergies of the quantum system and, consequently, to a random relative phase change between the two qubit states. Therefore, dephasing can occur via two different mechanisms, "amplitude damping" and "pure dephasing". In the presence of both types of decoherence mechanisms, relaxation time T_2 (which characterizes the total dephasing rate), T_1 and T_ϕ are related as $\dfrac{1}{T_2} = \dfrac{1}{2T_1} + \dfrac{1}{T_\phi}$, and $T_1 > T_2$.

Decoherence is the main agent destroying the quantum information and one of the key obstacles in implementing quantum computers devices. A key problem is the fact that quantum noise, and in particular, decoherence, are non-unitary operations that cause a pure state to become a mixed state.

2.1.2 Density matrix

In most of the situations, a quantum system appears to be in a mixed state, and therefore the state of a quantum system must be described by a *density matrix* (or equivalently *density operator*). For example, it is not possible to describe a quantum mechanical system that undergoes general quantum operations such as measurement, using exclusively states represented by wave functions (ket vectors). A density matrix $\hat{\rho}$ contains all the physically significant information about a quantum system. If a density matrix is given, we can calculate the ensemble average $\langle \hat{O} \rangle$ of any arbitrary operator $\hat{O}$ on the Hilbert space of the system.

The density matrix is a linear, Hermitian operator on the Hilbert space of wave functions. If a quantum system can be in one of the states $|\Psi_i\rangle$ with probability p_i, then the density operator $\hat{\rho}$ for such system is

$$\hat{\rho} \equiv \sum_i p_i \big|\Psi_i\rangle\langle\Psi_i\big|. \tag{2.3}$$

The density operator has the following properties:

• Hermicity: The density matrix is a hermitian, or self-adjoint, operator: $\hat{\rho} = \hat{\rho^{\dagger}}$;

• Positive semi-definiteness: All of the eigenvalues of the density operator are nonnegative, or, equivalently, $\langle\phi\big|\hat{\rho}\big|\phi\rangle \geq 0$ for any state $|\phi\rangle$;

• Trace condition: its trace must equal 1, $\mathrm{Tr}(\hat{\rho}) = \sum_i \rho_{ii} = 1$;

• Measurement: For a system in state $\hat{\rho}$, the expectation value of a measurement described by Hermitian operator $\hat{\mathbf{O}}$ is $\mathrm{Tr}(\hat{\rho}\hat{\mathbf{O}})$.

The elements of the N-qubit density operator form a $2^N \times 2^N$ matrix (or, equivalently, $d \times d$ matrix, where d is the dimension of Hilbert space). For example, in case of a single qubit ($N = 1$), a system is in a pure state described by a state vector (wavefunction) $|\Psi\rangle$ as in Eq. (2.1), and there is only one term in the sum Eq. (2.3) with $p_1 = 1$. Therefore, the density matrix can be written as

$$\hat{\rho} = \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix} = \rho_{00}|0\rangle\langle 0| + \rho_{01}|0\rangle\langle 1| + \rho_{10}|1\rangle\langle 0| + \rho_{11}|1\rangle\langle 1|, \tag{2.4}$$

or explicitly as

$$\hat{\rho} = |\Psi\rangle\langle\Psi| = |a|^2|0\rangle\langle 0| + ab^*|0\rangle\langle 1| + a^*b|1\rangle\langle 0| + |b|^2|1\rangle\langle 1| = \begin{pmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{pmatrix}. \tag{2.5}$$

The elements ρ_{00} and ρ_{11} on the main diagonal are real numbers (usually referred to as "populations" of the two energy levels, as they determine the probabilities to find the system in one of the two energy states), and the off-diagonal elements ρ_{01} and ρ_{10} are complex numbers (usually referred to as "coherences"). Therefore we need three real parameters to write the density operator of a one-qubit system, which can be chosen as ρ_{11}, $\mathrm{Re}(\rho_{10})$ and $\mathrm{Im}(\rho_{10})$:

$$\hat{\rho} = \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix} = \begin{pmatrix} 1 - \rho_{11} & \mathrm{Re}(\rho_{10}) - i\,\mathrm{Im}(\rho_{10}) \\ \mathrm{Re}(\rho_{10}) + i\,\mathrm{Im}(\rho_{10}) & \rho_{11} \end{pmatrix}. \tag{2.6}$$

In general, the density operator for the system of N qubits is a $2^N \times 2^N$ matrix that contains $(4^N - 1)$ independent real parameters.

As we will be discussing in details the Quantum State Tomography algorithms for the two-qubit systems in Chapter 3, specifically in Section 3.3, we want to introduce here the following

parametrization for a two-qubit density matrix:

$$\hat{\rho} = \begin{pmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ r_{21} & r_{22} & r_{23} & r_{24} \\ r_{31} & r_{32} & r_{33} & r_{34} \\ r_{41} & r_{42} & r_{43} & r_{44} \end{pmatrix}, \tag{2.7}$$

which is a 4×4 matrix containing 16 real parameters, 15 of which are independent. There are 4 real parameters r_{jj} on the main diagonal, subject to the normalization condition $r_{11}+r_{22}+r_{33}+r_{44} = 1$ (this leaves only three independent parameters out of four), and 12 off-diagonal elements which are complex numbers, subject to the hermicity condition $r_{ij} = r^*_{ji}$, which leaves us with 12 independent real parameters for the off-diagonal elements.

2.1.3 Bloch sphere and Pauli operators

There exists a nice way of graphical representation of a single qubit state as a point on a surface of the so-called *Bloch Sphere* (named after a physicist Felix Bloch, one of Heisenberg's students). Pure states are represented by points on the surface of the Bloch sphere of unit radius, while mixed states are inside of the sphere. Such representation can be derived as follows. The most general state of a qubit, described by Eq. (2.1), can be represented as

$$|\psi\rangle = e^{i\gamma}\Big(\cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle\Big), \tag{2.8}$$

because a and b are complex numbers in Eq. (2.1). The overall phase factor γ in Eq. (2.8) has no observable effect and can be set to zero. The variables $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi]$ are numbers that together define a point on the surface of a unit sphere, which is known as the Bloch sphere. This

fact becomes obvious if we introduce the parametrization

$$x = \sin\theta\cos\phi, \tag{2.9}$$

$$y = \sin\theta\sin\phi, \tag{2.10}$$

$$z = \cos\theta. \tag{2.11}$$

Note that the north and the south poles of the Bloch sphere correspond to the states $|0\rangle$ and $|1\rangle$ respectively. Points that are antipodal on the Bloch sphere represent orthogonal pure states of the qubit, and overlaps between states can be calculated from the relative angle between the two corresponding points on the Bloch sphere.

It is convenient to visualize rotations of a qubit on the Bloch sphere. Rotations on the Bloch sphere about $x-$, $y-$ and $z-$axes are described mathematically by the exponentiated Pauli matrices (Pauli operators) σ_x, σ_y and σ_z, where

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.12}$$

The Pauli operators play a very important role in quantum information processing. The set $\{\sigma_0, \sigma_x, \sigma_y, \sigma_z\}$ forms a traceless (except for σ_0), unitary, Hermitian and orthogonal basis (orthonormality can be achieved by scaling each element by $1/2$).

The rotation operators on the Bloch sphere about the axes x, y and z by angle θ are described by the following operators $R_x^{\ \theta}$, $R_y^{\ \theta}$ and $R_z^{\ \theta}$:

$$R_x^{\ \theta} \equiv e^{-i\theta\sigma_x/2} = \cos\Big(\frac{\theta}{2}\Big)I - i\sin\Big(\frac{\theta}{2}\Big)\sigma_x = \begin{pmatrix} \cos\dfrac{\theta}{2} & -i\sin\dfrac{\theta}{2} \\ -i\sin\dfrac{\theta}{2} & \cos\dfrac{\theta}{2} \end{pmatrix}; \tag{2.13}$$

$$R_y^\theta \equiv e^{-i\theta\sigma_y/2} = \cos\Big(\frac{\theta}{2}\Big) I - i \sin\Big(\frac{\theta}{2}\Big)\sigma_y = \begin{pmatrix} \cos\dfrac{\theta}{2} & -\sin\dfrac{\theta}{2} \\ \sin\dfrac{\theta}{2} & \cos\dfrac{\theta}{2} \end{pmatrix}; \tag{2.14}$$

$$R_z^\theta \equiv e^{-i\theta\sigma_z/2} = \cos\Big(\frac{\theta}{2}\Big) I - i \sin\Big(\frac{\theta}{2}\Big)\sigma_z = \begin{pmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{pmatrix}. \tag{2.15}$$

A single-qubit density matrix (not necessarily of a pure state) can be expressed in terms of Pauli matrices, which are defined by Eq. (2.12):

$$\rho = \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix} = \frac{1}{2}\left[r_0\sigma_0 + r_x\sigma_x + r_y\sigma_y + r_z\sigma_z\right], \qquad r_0 = 1, \tag{2.16}$$

or

$$\rho = \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 + r_z & r_x - ir_y \\ r_x + ir_y & 1 - r_z \end{pmatrix}, \tag{2.17}$$

where the 3-vector $\hat{r} = [r_x, r_y, r_z]$ lies in the unit Bloch's sphere. This vector is called the Bloch vector (or Bloch representation) of $\hat{\rho}$. For single qubit states, every point in the unit sphere is associated to a unique quantum state and the pure states lie on the boundary of the unit sphere.

Note that the Bloch sphere is closely analogous to the Poincaré sphere, which was developed by Henri Poincaré in the 19th century. Invented in 1891 by Henri Poincaré, the Poincaré sphere represents classical polarizations in an analogous way (up to the 90° rotation) that the Bloch sphere represents qubits.

The generalization of the Bloch vector to multi-qubit systems is not straightforward, because the pictorial representation of states associated to every point on the sphere becomes extremely complicated due to exotic geometry of multi-qubit state spaces, and not every state on the surface would be physical. On the other hand, the Pauli basis can be easily generalized for a sys-

tem of N qubits as the Kronecker (tensor) products of N single-qubit Pauli operators from the set $\{\sigma_0, \sigma_x, \sigma_y, \sigma_z\}$.

2.1.4 Definition of a quantum channel

Generally speaking, a quantum channel $\mathcal{E}$ describes the evolution of a state of a physical system characterized by the density matrix $\hat{\rho}^{\text{in}}$ to a new state $\hat{\rho}^{\text{fin}}$ under some process $\mathcal{E}$: $\hat{\rho}^{\text{fin}} = \mathcal{E}(\hat{\rho}^{\text{in}})$. The most prominent example is the unitary time evolution of a closed system, $\hat{\rho}^{\text{fin}} = U\hat{\rho^{\text{in}}}U^\dagger$ with $U = e^{-iHt/\hbar}$ and Hamiltonian H: The unitary evolution $\mathcal{E}(\hat{\rho})$ of a density operator $\hat{\rho}$, described by a unitary operator U, is

$$\mathcal{E}(\hat{\rho}) = U\hat{\rho}\,U^\dagger. \tag{2.18}$$

In general, the notion of quantum channels is much broader, for example, the transformation $\mathcal{E}$ does not have to be unitary. The formal definition of a quantum channel can be formulated in the following way. Let H_A and H_B be the state spaces (finite-dimensional Hilbert spaces $H_{A,B} \in \mathbb{C}^n$) of dimensions d_1 and d_2 of the sending and receiving ends, respectively, of a quantum channel. Let $\mathcal{L}(H_A)$ denote the set of bounded linear operators on H_A, in other words it is a matrix algebra of bounded operators on a Hilbert space H_A, and analogously for $\mathcal{L}(H_B)$. By definition, a quantum process (quantum channel) is a linear, completely positive, trace-preserving map $\mathcal{E}$:

$$\mathcal{E} : \mathcal{L}(H_A) \longrightarrow \mathcal{L}(H_B). \tag{2.19}$$

A map is called positive, if it maps positive operators to positive operators. It is called completely positive, if this is the case even when the map is only applied to a subsystem. That is, $\mathcal{E}$ is completely positive if

$$\mathcal{E} \otimes I_n : \mathcal{L}(H_A) \otimes \mathcal{L}(\mathbb{C}^n) \longrightarrow \mathcal{L}(H_B) \otimes \mathcal{L}(\mathbb{C}^n) \tag{2.20}$$

is positive for all $n \in \mathbb{N}$, where I_n denotes the identity map on $\mathcal{L}(\mathbb{C}^n)$. In other words, if an ancilla of an arbitrary finite dimension n is coupled to the system, then the induced map $\mathcal{E} \otimes I_n$ is also positive. Complete positivity allows one to use channels to describe operations local to a subsystem. A map is called trace-preserving if

$$\mathrm{Tr}[\mathcal{E}(\hat{\rho})] = \mathrm{Tr}(\hat{\rho}) \ \text{ for all } \hat{\rho} \in \mathcal{L}(H_A). \tag{2.21}$$

Sometimes in the literature the property of trace preservation is weakened so that $\mathcal{E}$ is only required to be trace-nonincreasing. A map is called linear if applying the map to a linear combination of input density matrices yields the same result as applying it to the input density matrices separately and taking the linear combination of the resulting states:

$$\mathcal{E}\big(\sum_i p_i \hat{\rho}_i^{\text{in}}\big) = \sum_i p_i \mathcal{E}(\hat{\rho}_i^{\text{in}}) \ \text{ for all } \rho_i^{\text{in}} \in \mathcal{L}(H_A). \tag{2.22}$$

Sometimes in the literature a condition for $\mathcal{E}$ to be unital is added to the definition of a quantum channel. A map $\mathcal{E} : \mathcal{L}(H_A) \longrightarrow \mathcal{L}(H_B)$ is called unital if it maps the unity of $\mathcal{L}(H_A)$ to the unity of $\mathcal{L}(H_B)$,

$$\mathcal{E}(\mathbb{I}_A) = \mathbb{I}_B. \tag{2.23}$$

Unitary channels are unital and describe the dynamics of a quantum system that is isolated from environmental interactions.

2.1.5 Distance between quantum states

In order to compare the two density matrices $\hat{\rho}$ and $\hat{\sigma}$ we use the standard measures of the distance between two quantum states from the quantum information theory: "fidelity" and "trace

distance". The general definition of the fidelity [69, 70, 71, 72]:

$$F(\hat{\rho},\hat{\sigma}) = \left\{ \mathrm{Tr}\left(\sqrt{\rho^{\frac{1}{2}}\sigma\rho^{\frac{1}{2}}} \right) \right\}^2. \tag{2.24}$$

The fidelity takes values between 0 and 1, i.e. $0 \leq F \leq 1$. If the states $\hat{\rho}$ and $\hat{\sigma}$ are equal, then $F(\rho,\sigma) = 1$, and if the states have orthogonal support (which means that they are completely different) then $F(\rho,\sigma) = 0$. Also, this fidelity is symmetric in $\hat{\rho}$ and $\hat{\sigma}$. Note that some authors call $\sqrt{F}$ fidelity [2].

Another distance measure between quantum states is the trace distance [2], which is defined as

$$D(\rho,\sigma) = \frac{1}{2}\mathrm{Tr}\Big(|\rho-\sigma|\Big) = \frac{1}{2}\mathrm{Tr}\Big[\sqrt{(\rho-\sigma)^\dagger(\rho-\sigma)}\Big]. \tag{2.25}$$

2.1.6 Quantum gates

A quantum gate (or a quantum logic gate) is a basic quantum circuit that operates on a small number of qubits. Quantum logic gates can be represented by unitary matrices. The most common quantum gates operate on spaces of one, two or three qubits.

Important examples of **one-qubit gates** are the following:

• X-gate (equivalent names are NOT-gate or bit-flip), which maps $\big|0\rangle$ to $\big|1\rangle$ and $\big|1\rangle$ to $\big|0\rangle$ and has the meaning of rotation of the Bloch sphere around the X-axis by π radians, it is represented by the Pauli σ_x matrix, defined in Eq. (2.12),

• Z-gate (also called as phase-flip gate), which leaves the basis state $\big|0\rangle$ unchanged and maps $\big|1\rangle$ to $-\big|1\rangle$ and has the meaning of a rotation around the Z-axis of the Bloch Sphere by π radians (azimuthal rotation of the Bloch vector by π about Z-axis), it is represented by the Pauli σ_z matrix, defined in Eq. (2.12),

• There are other gates, such as Hadamard gate, Y-gate, the phase shift gates and the swap gate. Obviously, there are infinitely many 2×2 unitary matrices, therefore infinitely many possible single qubit gates.

Controlled gates act on two or more qubits, where one or more qubits act as a control for some operation, and one of the qubits, called target qubit, changes its state depending on the state of the control qubit(s). Important examples of the two-qubit quantum gates are the following:

• CNOT gate. It operates on two qubits as follows: leaves the second qubit unchanged when the first qubit is in state $\big|0\rangle$, and it flips the second qubit if the first qubit is in state $\big|1\rangle$. The vectors $\big|00\rangle$, $\big|01\rangle$, $\big|10\rangle$, and $\big|11\rangle$ form an orthonormal basis for the set of pure states for a two-qubit system. Therefore, in this basis the CNOT gate has the following representation:

$$\begin{array}{ll} \big|00\big\rangle \longrightarrow \big|00\big\rangle & \\ \big|01\big\rangle \longrightarrow \big|01\big\rangle & \\ \big|10\big\rangle \longrightarrow \big|11\big\rangle & \\ \big|11\big\rangle \longrightarrow \big|10\big\rangle & \end{array} \qquad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \tag{2.26}$$

• Controlled-Z (CZ) gate is a controlled two-qubit gate which implements the flip the sign of $\big|11\big\rangle$ state and keeps the other three basis states unchanged, its matrix representation is

$$\begin{array}{ll} \big|00\big\rangle \longrightarrow & \big|00\big\rangle \\ \big|01\big\rangle \longrightarrow & \big|01\big\rangle \\ \big|10\big\rangle \longrightarrow & \big|10\big\rangle \\ \big|11\big\rangle \longrightarrow & -\big|11\big\rangle \end{array} \qquad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{2.27}$$

The Toffoli gate. The Toffoli gate, which is essentially a CCNOT-gate, is a three-qubit controlled gate, which applies the X-gate on the third target qubit if both of the first two control

qubits are in the states $|1\rangle$ and it does nothing for the other three possible states of the two control qubits.

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}. \tag{2.28}$$

2.2 Review of superconducting qubits

As discussed in Section 2.1.1, qubit or quantum bit is a fundamental unit of quantum information, which can store information not only in the states $|0\rangle$ or $|1\rangle$, but also as a combination of both basis states simultaneously, which is the manifestation of the phenomenon of quantum superposition. There are various physical realizations of qubits. As discussed in Section 2.1.1, numerous implementations have been reported in optics where the two photon polarizations form the effective two-level system, in trapped ions, in semiconductors, in liquid NMR systems, in NV centers in diamond and in other physical systems, but the object of research of this dissertation is the analysis of the quantum states and quantum processes involving superconducting qubits, based on Josephson effect. Superconducting qubits are electrical resonators with a non-linearity that is strong enough so that microwaves excite transitions only between the the two lower levels of the system. Superconducting qubits include such elements as Josephson junctions, capacitors and inductors,

and can be considered as nonlinear LC-circuits. Josephson junctions introduce the nonlinearity into the circuits, which leads to unequal distances between energy levels of the system and the possibility to distinguish the lower two levels of the system from the rest of the levels. We discuss superconducting qubits in more details in the next Sections 2.2.2, 2.2.3, 2.2.4 and 2.2.5.

Comparing superconducting qubits to other possible physical realizations of qubits, superconducting qubits have the largest physical size: their typical size is tens of micrometers (from $\sim 1\mu m$ to $\sim 100\mu m$). This circumstance makes these types of qubits relatively easy to fabricate using standard microfabrication techniques and integrate them into the electric circuits, consisting of capacitors, inductors, transmission lines and other linear electronic components. This makes it easier, compared with other qubit types, to couple superconducting qubits with the readout and control circuits and to each other. Strong qubit - qubit coupling between superconducting qubits also means faster gates compared with the other types of qubits. As a disadvantage of superconducting qubits, strong coupling to the environment and the larger size of the qubits mean shorter coherence times, which are currently $10 - 100$ microseconds. Another difficulty is that superconducting qubits are not true two-level systems, there are always more than two energy levels $\big|0\big\rangle$ and $\big|1\big\rangle$ in the potential well. The upper level $\big|2\big\rangle$ can become accidentally populated, which is possible for the qubits with small anharmonicity when the $\big|0\big\rangle \longrightarrow \big|1\big\rangle$ transition frequency differs very insignificantly (several percent) from the $\big|1\big\rangle \longrightarrow \big|2\big\rangle$ transition frequency. For example, for this reason a small nonlinearity places restrictions on the speed of quantum gates in transmon qubits.

2.2.1 General introduction to Josephson junction dynamics

The Josephson effect is the phenomenon of supercurrent (a current that flows indefinitely long without any applied voltage) across a device known as a Josephson junction. In superconduc-

tors at low temperatures electrons bind into Cooper pairs (bound states of two electrons with opposite momenta and spins) that condense into a state with zero-resistance and a well-defined phase. The idea of forming the Cooper pairs is the following: an electron moving through a lattice of ions creates vibrations (phonons), which can be absorbed by another electron. The interaction that arises as a result can be attractive provided the electron-phonon coupling is strong enough. Cooper pairs are bosons (the net spin of a pair of electrons is zero), therefore they may form a Bose-Einstein condensate that possesses the property of superconductivity. All the Cooper pairs at a given point in a superconductor can be described by a single wavefunction with a phase ϕ.

A Josephson junction is a quantum mechanical device which consists of two superconducting electrodes separated by a thin[1] layer of a nonsuperconducting material, such as insulating tunnel barrier, thin normal metal, etc. The devices are named after Brian Josephson, who predicted in 1962 [73] that pairs of superconducting electrons can quantum-mechanically tunnel through the nonsuperconducting barrier from one superconductor to another, and also predicted the mathematical relationships for the current and voltage across the the junction. Anderson and Rowell [74] observed the effect in 1964. The supercurrent I_J that flows through the junction as the result of the tunneling of Cooper pairs, and the voltage V across the barrier, are described by the following two classical equations:

$$I_J = I_0 \sin\delta, \tag{2.29}$$

$$V = \frac{\Phi_0}{2\pi}\frac{d\delta}{dt}, \tag{2.30}$$

where $\Phi_0 = \dfrac{h}{2e} = 2.07 \times 10^{-15}$Wb is the superconducting flux quantum, I_0 is the critical-current parameter of the junction, $\delta = \phi_L - \phi_R$ and V are the superconducting phase difference and voltage

[1] Note that in a Josephson junction, the nonsuperconducting barrier separating the two superconductors must be very thin. If the barrier is an insulator, it has to be on the order of 30 angstroms thick or less. If the barrier is another metal (nonsuperconducting), it can be as much as several microns thick.

across the junction. It can be shown from these two equations that the Josephson junction can be considered as a nonlinear inductance: differentiating Eq. (2.29) and replacing $\frac{d\delta}{dt}$ with V using Eq. (2.30), one arrives at the equation $V = L_J \frac{dI_J}{dt}$ with the nonlinear inductance

$$L_J = \frac{\Phi_0}{2\pi I_0 \cos\delta} \tag{2.31}$$

(nonlinearity arises from the term $\frac{1}{\cos\delta}$). Therefore, as we see, a Josephson superconducting qubit can be understood as a nonlinear resonator formed from the Josephson nonlinear inductance and its geometric junction capacitance. The nonlinearity of the junction plays an extremely important role because it leads to non-equidistant energy levels in such resonator, namely the frequency ω_{10} that drives transitions between the qubit states $0 \leftrightarrow 1$ is different from the frequency ω_{21} for transitions $1 \leftrightarrow 2$.

Using Eq. (2.30), it is easy to derive another fundamental relation between the phase difference δ across the Josephson junction and the magnetic flux in the loop Φ. In a Josepson junction closed with a superconducting loop, the voltage across the Josephson junction can arise only as a result of a change of the magnetic flux Φ in the loop,

$$V = \frac{d\Phi}{dt}. \tag{2.32}$$

The magnetic flux Φ is defined as $\Phi \equiv \int \mathbf{B_n} dA$, where $\mathbf{B}$ is the magnetic field, and A is the area of the contour. Substituting into Eq. (2.32) the expression for the voltage across the junction from Eq. (2.30), we arrive at the following fundamental relation

$$\delta = 2\pi \frac{\Phi}{\Phi_0}. \tag{2.33}$$

As we see from Eq. (2.33), the phase difference δ across the junction is linearly proportional to the magnetic flux Φ in the loop. This is an important relation, in particular we will use Eq. (2.33) while discussing the phase qubits in Section 2.2.5.

There are three main types of superconducting qubits: the charge, the flux and the phase qubits. Charge qubits (also called Cooper pair box) were the first superconducting qubits to be demonstrated in 1998 [65, 75]. A year later the flux qubits were demonstrated [66]. The phase qubit was first demonstrated in 2002 [76, 77, 78]. The main difference between the charge, flux, and phase qubits is the shape of their nonlinear potentials, which are correspondingly cosine, quatric and cubic. In subsequent years various design modifications to these original qubit types have been done, and several other qubits were introduced, such as quantronium [79], transmon [80, 81], fluxonium [82], which have improved performance by reducing the sensitivity to decoherence mechanisms that existed in earlier designs. We are giving a brief introduction into the main types of superconducting qubits in the next sections of this dissertation, where we first describe the charge and flux qubits in Section 2.2.2 and 2.2.3, then discuss the transmon qubit as a refined charge qubit with charge noise insensitivity in Section 2.2.4, and finally we give an overview of the phase qubits in Section 2.2.5.

2.2.2 Charge qubit

The "Cooper pair box" or charge qubit [65], effective circuit diagram of which is presented in Fig. 2.1, is a small superconducting island ("box"), which is basically a small superconducting electrode, connected via a Josephson junction with capacitance C_J and Josephson energy E_J to a grounded reservoir. (There also exists a modified version, called "split Cooper pair box", presented at Fig. 2.2, with two Josephson junctions, connected together to form a superconducting loop, which can be biased by an external magnetic flux Φ_{ext}.) The superconducting island can be biased by a voltage source V_g in series with gate capacitance C_g. Cooper pairs can tunnel onto the island and off the island, one at a time, therefore the relevant degree of freedom is the

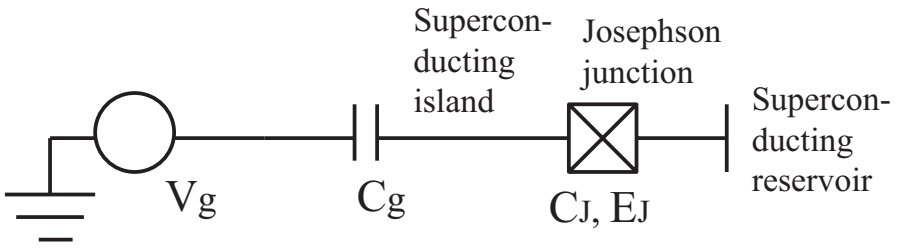

Figure 2.1: Effective circuit diagram of the Cooper-pair box. The small superconducting island is connected to a large superconducting reservoir by a Josephson junction with capacitance C_J and Josephson energy E_J. The island can be biased by a voltage source (gate voltage) V_g in series with a gate capacitance C_g.

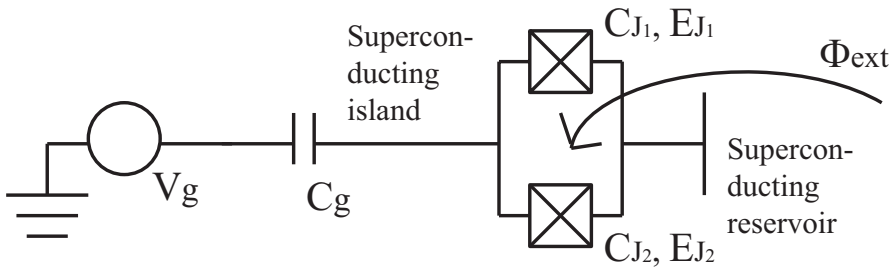

Figure 2.2: Effective circuit diagram of the split Cooper-pair box. The small superconducting island is connected to a large superconducting reservoir by two Josephson junctions with capacitances C_J and Josephson energies E_J (not necessarily equal). The island can be biased by a voltage source (gate voltage) V_g in series with a gate capacitance C_g and by an external magnetix flux Φ_{ext}.

number of Cooper pairs on the island. Such a qubit is described by its Josephson coupling energy $E_J = I_0\Phi_0/(2\pi)$ and by its Coulomb charging energy, corresponding to one Cooper pair on the junction, $E_C = (2e)^2/(2C_\Sigma)$, where $C_\Sigma = C_J + C_g$ is the total capacitance of the box (the sum of the capacitance of the junction C_J and the capacitance of the gate C_g), and e is the electron charge. Coherent oscillations in such Cooper-pair box have been observed in [83, 84].

To write the Hamiltonian of the charge qubit, we need to introduce the operator $\hat{n}$ of the number of excess Cooper pairs on the island. The total number of excess Cooper pairs n is related to the total charge q of the island as $q = -2en$, $n = 0$ corresponding to an electrically neutral island. Obviously, n can take both positive and integer values, corresponding to the excess

or deficit number of Cooper pairs on the island, and as Cooper pairs can tunnel in and out of the island by Josephson tunneling, this number n can fluctuate quantum-mechanically. Therefore the convenient basis for the charge states of the island are the eigenvectors of $\hat{n}$: $\hat{n}|n\rangle = n|n\rangle$. It is also convenient to introduce the operator $\hat{\delta}$, conjugated to $\hat{n}$, with the commutation relation $[\hat{\delta}, \hat{n}] = i$, where $\delta \in [0, 2\pi]$ defines the phase of the Cooper pair condensate on the island. Using the charge basis, we can write the Hamiltonian in the charge representation as the sum of its electrostatic and Josephson coupling parts:

$$\hat{H} = \sum_n \left[\frac{(2e)^2}{2C_\Sigma} \Big(\hat{n} - \frac{C_g V_g}{2e} \Big)^2 |n\rangle\langle n| - \frac{E_J}{2} \Big(|n\rangle\langle n+1| + |n+1\rangle\langle n| \Big) \right], \tag{2.34}$$

or, introducing the dimensionless gate charge $n_g = \dfrac{C_g V_g}{2e}$, which represents the charge on the gate capacitor in units of Cooper pairs and can be tuned by the voltage V_g at the gate electrode,

$$\hat{H} = \sum_n \left[E_C \Big(\hat{n} - n_g \Big)^2 |n\rangle\langle n| - \frac{E_J}{2} \Big(|n\rangle\langle n+1| + |n+1\rangle\langle n| \Big) \right]. \tag{2.35}$$

It can be shown that $\cos\delta = \dfrac{1}{2}\Big(|n\rangle\langle n+1| + |n+1\rangle\langle n| \Big)$, therefore the Hamiltonian can be rewritten as

$$\hat{H} = \sum_n \left[E_C \Big(\hat{n} - n_g \Big)^2 \right] - E_J \cos\delta. \tag{2.36}$$

For the "split Cooper pair box", presented at Fig. 2.2, the effective Josephson coupling E_J of the junctions in the charge qubit can be tuned with the external perpendicular magnetic flux Φ_{ext} through the loop [85] as

$$E_J = E_{J,\max} \cos\left(\pi \frac{\Phi_{\text{ext}}}{\Phi_0} \right), \tag{2.37}$$

then the Hamiltonian Eq. (2.36) takes the form

$$\hat{H} = \sum_n \left[E_C \Big(\hat{n} - n_g \Big)^2 \right] - E_{J,\max} \cos\left(\pi \frac{\Phi_{\text{ext}}}{\Phi_0} \right) \cos\delta. \tag{2.38}$$

The energy spectrum of Hamiltonian Eq. (2.36) or Eq. (2.38) depends on n_g and on the ratio E_J/E_C. The energy spectrum is discrete and periodic in n_g (the period equals one).

The eigenstates $|n\rangle$ and eigenenergies E_n of the charge qubit are given by

$$E_C\Big(\hat{n} - n_g\Big)^2|n\rangle = E_n|n\rangle, \;\; n \in \mathbb{N}, \tag{2.39}$$

and the number of Cooper pairs on the island is represented by the number operator, $\hat{n}$.

2.2.3 Flux qubit

In this Section the flux qubit [66] will be discussed very briefly (such type of qubits is never used in any calculations in this dDissertation, therefore we will discuss it briefly). The prototype of a flux qubit is a circuit called "RF-SQUID", which consists of a Josephson junction with capacitance C_J, the two sides of which are connected by a superconducting loop with inductance L (usually $L \leq L_J$ where L_J is defined in Eq. (2.31)). An external flux Φ_{ext} is imposed through the loop by an auxiliary coil. This external flux plays a role similar to the gate electrode in the charge qubit, allowing us to tune the effective Josephson coupling. The flux qubit parameters are such that $E_C \ll E_J$, minimizing the charge noise that is present in charge qubits. The flux qubit is operated with the external flux Φ_{ext} in the vicinity of the point $\Phi_{\text{ext}} = \Phi_0/2$. The potential of an RF-SQUID qubit has a form of a double well, this potential is symmetric for $\Phi_{\text{ext}} = \Phi_0/2$, and any change in Φ_{ext} tilts the potential. The two wells of the potential represent different "fluxoid" states of the RF-SQUID, corresponding to the counterclockwise and clockwise persistent currents circulating in the loop formed by L. Such coherent superposition of the two magnetic-flux states in different wells of the RF-SQUID was observed in [86]: one state corresponding to a few microamperes of current flowing clockwise, the other state corresponding to the same amount of current

flowing counterclockwise. In [87] and [88] a coherent time evolution between two quantum states of a superconducting flux qubit comprising three Josephson junctions in a loop was observed.

2.2.4 Transmon and Xmon

The drawback of the charge qubit, discussed in Section 2.2.2, is its sensitivity to the $1/f$ charge noise. A new type of superconducting charge qubit that has reduced sensitivity to the charge noise was developed at Yale University [80, 81] in 2007. Such a qubit is called "transmon", which stands for "transmission-line shunted plasma oscillation qubit". The transmon consists of two superconducting islands coupled through two Josephson junctions and isolated from the rest of the circuitry. The crucial distinction of the transmon from the "Cooper pair box" is a shunting connection of the two superconductors via a large capacitance C_B, accompanied by a similar increase in the gate capacitance C_g. The Hamiltonian of the transmon takes the same form as the Hamiltonian of the "Split Cooper pair box" Eq. (2.38), with the additional capacitance C_B in the expression for the total capacitance $C_\Sigma = C_J + C_B + C_g$, which enters into the formula for the Coulomb charging energy $E_C = (2e)^2/(2C_\Sigma)$. The presence of the additional capacitance term C_B allows us to make the charging energy E_C small compared to the Josephson energy E_J. The reduced sensitivity of the transmon to charge noise is made possible by significantly increasing the ratio of the Josephson energy to the charging energy E_J/E_C, which effectively flattens the energy bands (dependence of the eigenenergies E_n corresponding to Eq (2.39) on n_g becomes more and more flat as the ratio E_J/E_C increases), as explained in [80]. The typical values of the ratio E_J/E_C are about 20 — 100 for the transmon qubits (in other words $E_J \gg E_C$), while $E_J/E_C \sim 1$ for the charge qubits. Several experiments in which transmons were used are reported, for example, in [89, 90, 91].

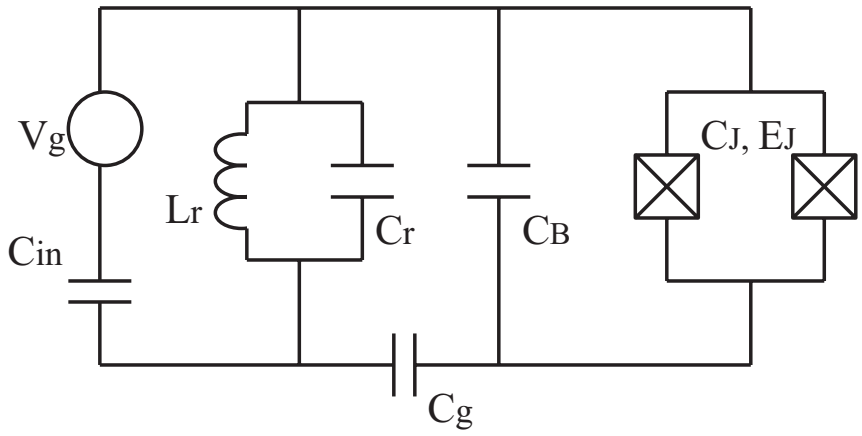

Figure 2.3: Effective curcuit diagram of the transmon qubit. The two Josephson junctions (with capacitance and Josephson energy C_J and E_J) are shunted by an additional large capacitance C_B, matched by a comparably large gate capacitance C_g. L_r and C_r model the coupling to the transmission line. V_g is the gate voltage.

The main idea of the method of measurement of the transmon qubit consists in coupling the transmon off resonance to a cavity, irradiating the cavity by microwave radiation, and then probing the transmitted or reflected photons. A cavity is realized in transmon experiments as a superconducting transmission line resonator, called coplanar stripline resonator. When a qubit is coupled off resonance to a cavity, the resonator's frequency depends on the qubit's quantum state, and photons populating the resonator acquire a qubit state dependent phase shift. The phase or amplitude of the microwave field outgoing from the resonator is measured, and the measurement contains information about the state of the qubit (more specifically, information about the qubit's projection along the z-axis of the Bloch sphere). The resonant frequency of the resonator can be controlled by changing the length of its middle stripe, and when the resonator is irradiated at the resonant frequency, a standing wave is formed inside of the resonator. The coupling of the qubit with the resonator is achieved by putting the qubit into the center of the transmission line.

This spectroscopic measurement scheme was first considered in [92], and this technique was first demonstrated for charge qubits [93, 94, 95] before the transmon was invented, but as shown in [80], this technique called dispersive measurement may be directly transferred to the transmon. In the dispersive limit, the detuning $\Delta \equiv \omega_q - \omega_r$ between the qubit and the resonator frequencies is large in comparison with the strength of the qubit-resonator coupling g, $\Delta \gg g$ ($\omega_q = \omega_1 - \omega_0$ is the transition frequency of the qubit). Theoretical analysis shows that the resonator frequency is shifted by the value $\chi \equiv \pm g^2/\Delta$, and the difference in resonator frequency for the two qubit states is 2χ, see Fig. 2.4. To perform a measurement of the qubit, a pulse of microwave radiation with the frequency of $\omega_{\mu w} = \omega_r$ or $\omega_r \pm g^2/\Delta$ is sent through the cavity. The choice of the frequency of the probe radiation $\omega_{\mu w}$ is determined by the ratio of the parameters Δ, g, and κ of the system (κ denotes the cavity decay rate which determines the width of the resonant peak at Fig. 2.4). As explained in [92], if $\dfrac{g^2}{\kappa\Delta} > 1$, i.e. the frequency shift of the resonator g^2/Δ is greater than the line width κ, then under irradiation of the cavity at one of the "pulled" frequencies $\omega_r \pm g^2/\Delta$, the transmission of the cavity will be close to unity for one state of the qubit and close to zero for the other state. Therefore, knowing the drive frequency and the intensity of the transmitted radiation, one can measure which state the qubit is in. If the opposite situation $\dfrac{g^2}{\kappa\Delta} < 1$ takes place, then it is preferable to irradiate the resonator at the bare cavity frequency $\omega_{\mu w} = \omega_r$. In this case the state of the qubit is encoded in the phase of the transmitted microwaves, and the homodyne detection technique can be used to measure it and to infer the state of the qubit. Such dispersive measurement with a microwave pulse has become a well-established technique of transmon readout, for example the measurement fidelity of 98% was reported recently in [96]. The dispersive measurement technique was generalized for the systems of two transmon qubits in [97], where a method to jointly and simultaneously read out

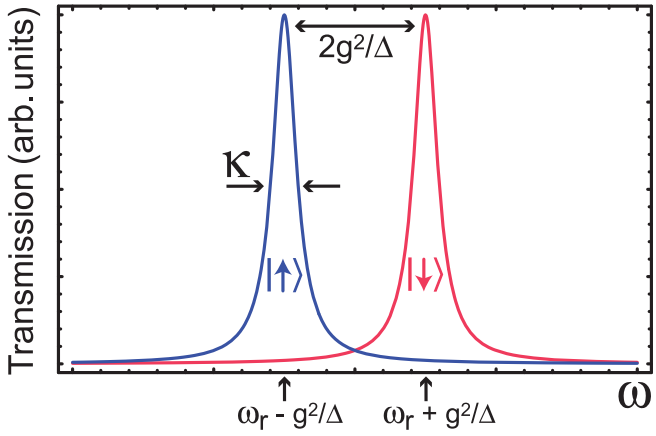

Figure 2.4: Transmission Spectrum of the cavity, which presents a peak of width κ at $\omega_r - g^2/\Delta$ or $\omega_r + g^2/\Delta$ depending on the state of the qubit, red curve for the excited state, blue curve for the ground state. κ denotes the cavity decay rate. To perform a measurement of the qubit, a pulse of microwave photons, at a probe frequency $\omega_{\mu w} = \omega_r$ or $\omega_r \pm g^2/\Delta$, is sent through the cavity. (*Adapted from Ref. [92].*)

the quantum state of two transmon qubits dispersively coupled to a microwave resonator has been presented.

In 2013 a new type of the transmon qubit, the Xmon, has been developed [98] in the research group of prof. J. Martinis (UCSB). Such Xmon qubit shows long coherence time exceeding $40\mu s$ (in the original paper [98] the reported measured energy relaxation time was $T_1 = 44\mu s$, the much longer time $T_1 \sim 60\mu s$ has been reported recently in [99]), allows for straightforward coupling to multiple elements, and has a low parasitic coupling. Such qubits are frequency tunable, which allows the high-fidelity implementation of fast two-qubit gates, such as CZ-gate. The device is shown in Fig. 2.5, it has a planar geometry consisting of the cross-shaped qubit capacitor, which connects at the bottom to the tunable Josephson junction, formed by the rectangular ring-shaped superconducting quantum interference device, and the four arms of the cross-shaped capacitor are connected to the following four elements: to the readout line on the top, to a quantum bus resonator

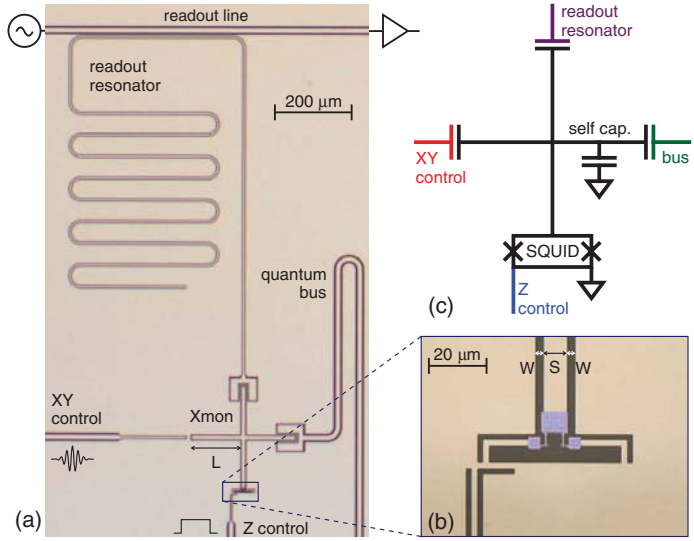

Figure 2.5: (a) Schematics of the Xmon qubit, formed by the Al superconducting film and the exposed sapphire substrate. The qubit is capacitively coupled to a readout resonator on the top, a quantum bus resonator (right), and an XY control line (left), and inductively coupled to a Z control line (bottom). (b) The inset shows the shadow evaporated Al junction layer (blue regions). (c) The electrical circuit of the qubit. (*Adapted from Ref. [98].*)

on the right, to an XY control line on the left for the qubit state excitations, and to a Z control line on the bottom for tuning the qubit frequency. The first three couplings are capacitive, while the last coupling to a Z control line is inductive. The typical sizes of the described qubit are the following: the arm length $L \sim 130 - 165\mu m$, and the width of the junction size S, W on the order of $10\mu m$ (see Table 1 in [98] for the discussion of the specific values of these parameters). Such Xmon qubit has the transition frequency between the ground and first excited states of about 6 GHz, and a ratio of Josephson to charging energy $E_J/E_C \sim 95$. The main experimental data used in this dissertation in Chapters 5 and 6, where we present our results for the compressed sensing method of QPT, were obtained using the two-qubit CZ gate realized with Xmon qubits [98].

2.2.5 Flux-biased phase qubit

Besides using the experimental data obtained with Xmon qubits in our calculations, we have also performed the analysis of the CS QPT of two CZ gates based on phase qubits in Chapters 5 and 6. Therefore we are giving a brief review of the type of superconducting qubits known as the "flux-biased phase qubit" [67, 68, 100] in this Section. A flux-biased phase qubit is a superconducting loop, formed by a single Josephson junction in a nonlinear LC-circuit, exposed to an external magnetic flux Φ_{ext}. Its schematics is presented in Fig. 2.6. When an external magnetic flux is applied to the loop, a current starts to circulate in the loop, which, in turn, produces a magnetic flux, which changes the total flux. One can derive the following expression for the potential energy of such phase qubit (see, for example, [100]):

$$U(\delta) = -\frac{I_0\Phi_0}{2\pi}\cos\delta + \frac{\Phi_0^2}{2L}\left[\phi - \frac{\delta}{2\pi}\right]^2 = -E_J\cos\delta + \frac{\left(\frac{\Phi_0}{2\pi}\delta - \Phi_{\text{ext}}\right)^2}{2L}, \tag{2.40}$$

where δ is a superconducting phase difference across the Josephson junction, $\Phi_0 = h/(2e)$ is the magnetic flux quantum, $\phi = 2\pi\Phi_{\text{ext}}/\Phi_0$ is the dimensionless external magnetic flux, $E_J = I_0\Phi_0/(2\pi)$ is the Josephson coupling energy, I_0 is the critical current, e is the electron charge, L is the inductance. Note that typically the ratio of the Josephson energy E_J to the charging energy E_C is much higher than that for the charge qubit or for the transmon: for phase qubits $E_J/E_C \sim 10^4$ [76, 101].

The derivation of Eq. (2.40) is pretty straightforward. First, we write down the expressions for the energy of the capacitance E_C, the Josephson junction E_{JJ}, and the inductance E_L. As is well known, the energy stored in the capacitor C is a function of the charge Q (or, equivalently, of voltage V across the capacitor), and the magnetic energy which is stored in the inductor L is a

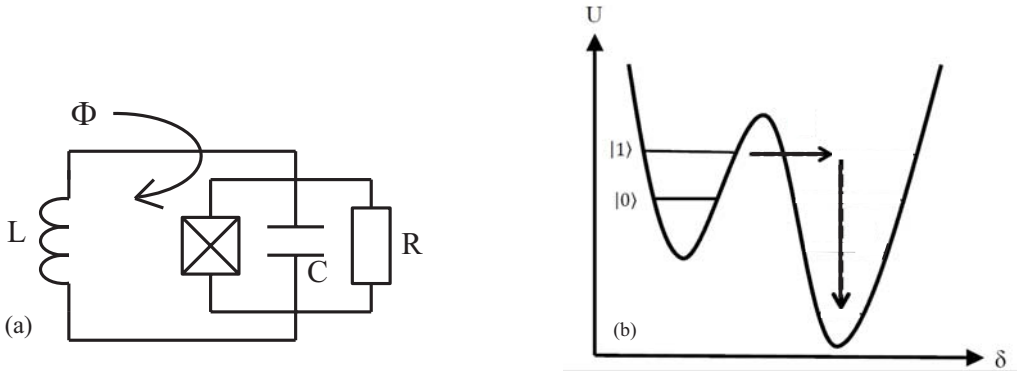

Figure 2.6: The circuit schematic of a flux-biased phase qubit (a), and the schematics of the plot of the potential energy U as a function of the phase difference δ across the Josephson junction with the measurement scheme (b). In Fig. (a): Φ is the external flux applied through the inductor loop, L is the inductance of the loop, C is the capacitance of the loop. In Fig. (b): When measuring the state of the phase qubit, the height of the potential barrier is lowered, the state $|1\rangle$ tunnels through the potential barrier, and relaxes into the deeper well. The superconducting phase difference δ changes its value.

function of the magnetic flux Φ:

$$E_C = \frac{CV^2}{2} = \frac{C}{2}\Big(\frac{\hbar^2}{4e^2}\Big)\Big(\frac{d\delta}{dt}\Big)^2, \tag{2.41}$$

$$E_{JJ} = -E_J \cos\delta, \tag{2.42}$$

$$E_L = \frac{1}{2L}\Phi^2 = \frac{1}{2L}\Big(\frac{\Phi_0}{2\pi}\delta - \Phi_{\text{ext}}\Big)^2. \tag{2.43}$$

While writing these relations, we have used some basic equations for Josephson junctions from Chapter 2.2.1. We have used Eq. (2.30) when writing here the equation Eq. (2.41), and we have also used the previously derived relation Eq. (2.33) between the phase difference δ across the junction and the magnetic flux Φ in the loop in Eq. (2.43) for E_L. Eq. (2.33) from Chapter 2.2.1 is valid without the external magnetic flux in the circuit, and when we turn on the external magnetic flux Φ_{ext}, we modify Eq. (2.33) to

$$\Phi + \Phi_{\text{ext}} = \frac{\Phi_0}{2\pi}\delta, \tag{2.44}$$

from which we get

$$\Phi^2 = \left(\frac{\Phi_0}{2\pi}\delta - \Phi_{\text{ext}}\right)^2, \tag{2.45}$$

which we used in Eq. (2.43). The kinetic part of the total energy of the phase qubit is E_C, and the potential part of the total energy of the phase qubit is $E_{JJ} + E_L$, therefore we arrive at the expression for the potential energy Eq. (2.40) by summing up Eq. (2.42) and Eq. (2.43).

The typical profile of the potential energy $U(\delta)$ for a flux-biased phase qubit as a function of the superconducting phase difference δ across the junction can be modeled by two wells with different numbers of quantized states in each well, as in Fig. 2.6. The "left" potential well is shallow, while the "right" potential well is much deeper. Typically, there are approximately 3 to 7 energy levels in the "left" shallow potential well, and at least 300 energy levels in the "right" deep potential well [78].

The most important feature of such potential is its anharmonicity, which means the energy levels in each potential well are not equidistant. The "shallow" left well can be approximated as a"cubic" potential (a quadratic parabola profile with a cubic anharmonicity), therefore the energy eigenstates are not equidistant. The lowest two energy eigenstates in the left shallow potential well form the effective two-level system and are used as the qubit basis states, the ground $|0\rangle$ and the first excited $|1\rangle$. Typically, the energy difference between these two lower levels is $E_{10}/\hbar \sim 6\,\text{GHz}$. Microwaves can be applied at the qubit frequency $\omega_{10} = E_{10}/\hbar$ to cause transitions and coherent oscillations between the qubit levels $|0\rangle$ and $|1\rangle$ without populating the upper levels. In addition, the flux bias ϕ_b applied to the qubit can be adjusted to tilt the potential, which allows changing the height of the barrier between the left and right potential wells. This possibility of controlling the height of the potential barrier allows the measurement of populations of the energy levels of the qubit.

The measurement scheme of the population of the upper qubit level of phase qubit was first implemented in 2004 [68, 101]. In this scheme, a low-frequency measurement pulse lowers the barrier between the shallow "left" potential well used for qubit states and a much deeper "right" well, which increases the probability of tunneling of a qubit from its upper energy level $|1\rangle$ from the "left" potential well into the "right" well. During the measurement pulse a qubit in the upper state $|1\rangle$ switches by tunneling to the right-hand well with probability close to one, while the state of a qubit in the lower state $|0\rangle$ does not change. When the tunneling of the state $|1\rangle$ into the deeper "right" well happens, it quickly relaxes into a state with a much lower energy, which is not sufficient for the return tunneling back into the shallow "left" well. If tunneling of $|1\rangle$ state happens, then the superconducting phase difference δ changes by π, which corresponds to the the magnetic flux variation by one magnetic flux quantum Φ_0, and the resulting magnetic field can be measured by the nearby SQUID (SQUID stands for "superconducting quantum interference device", which is a very sensitive magnetometer used to measure extremely weak magnetic fields, a voltage appears between the ports of the SQUID) [102]. This completes the measurement.

Chapter 3

Quantum State Tomography for superconducting qubits

3.1 The idea of the quantum state tomography

Quantum State Tomography (QST) is the procedure that allows us to completely determine an unknown quantum state from a set of experimental measurements. The need to estimate the quantum state appears when testing or claiming the preparation of specific quantum states. The density matrix $\hat{\rho}$ completely describes our knowledge of the state of a system, therefore the goal of Quantum State Tomography is measuring all the elements of the density matrix by several sets of multiple measurements. In this Chapter, we give a general overview of the idea of QST applied to the phase qubits in Section 3.1, provide mathematical details of the QST procedure in Section 3.2, and discuss in details the linear inversion method of QST in Section 3.3 and the maximum likelihood method of QST in Section 3.4. Matlab scripts implementing both these methods of QST in application to the phase qubit measurements have been written by the author of this thesis.

QST has been used to study a wide range of problems, such as determining the motional state of a trapped atom [103, 104] in 1996, determining the state of nuclear spins in NMR systems with a solution of chloroform molecules [105, 106] in 1998, determining the internal angular momentum state of an ensemble of cesium atoms [107] in 2001, also the technique of QST was described and implemented for quadrupole nuclei with nuclear spin 3/2 [108, 109] in 2004. In optics, the early ideas of quantum state tomography can be found in the work of G. S. Stokes [110] in 1852, where a minimal set of measurements to describe the polarization of light was developed. The first tomography of an entangled quantum state was performed on the polarization state of two photons generated by spontaneous parametric down-conversion [51] in 2001. The first realization of QST for a single phase qubit at UCSB was reported in 2006 [111], and for two phase qubits also in 2006 in [112].

For a general unknown quantum state the measurement of the density matrix cannot be done in a single attempt, it requires non-commuting measurements to be repeatedly done in different bases. Consider first a single phase qubit. The procedure for measuring a qubit by lowering the height of the potential barrier and reading the magnetic field by a SQUID, described above in Section 2.2.5 of Chapter 2, gives only a binary output of 0 or 1 for the occupation of the qubit upper state $|1\rangle$. By identically preparing the state and repeating this measurement many (thousands) times, an average occupation probability of the eigenstates $|0\rangle$ and $|1\rangle$ can be found. This gives information only about the diagonal elements of the density matrix, and leaves the off-diagonal elements of the density matrix unknown. In the language of the Bloch sphere, the population difference is called "z-component" of the Bloch vector. As these measurements do not reveal any phase information, we need additional sets of measurements along the $x-$ and $y-$ directions of the Bloch sphere, which

are achieved by the rotating the state to be measured by $\pm\pi/2$ around the $y-$ or $x-$axis before the measurement. This rotation is fulfilled experimentally by applying microwave pulses as explained in Section 2.2.5 of Chapter 2. From the sets of measurements of the Bloch vector in the $z-$, $x-$ and $y-$directions, the Bloch vector (and therefore all the elements of the density matrix) can be fully reconstructed. The procedure can be generalized for several qubits.

Before going into mathematical details of the procedure, let us make a couple of remarks.

As it was explained, we need basis rotations. There are two alternatives: either (1) measure identical qubits in different bases, or (2) keep the measurement basis fixed, but rotate the qubits. These two approaches are equivalent because of the property of the quantum-mechanical trace operation:

$$\mathrm{Tr}[\hat{\rho}UMU^{\dagger}] = \mathrm{Tr}[U^{\dagger}\hat{\rho}UM].$$

For experimental convenience, the qubit is always rotated by applying microwave pulses, while the measurement basis is kept fixed (it is the computational basis $|0\rangle$ and $|1\rangle$).

Also, let us emphasize that we rotate about the $y-$ axis to obtain the $x-$coordinate on the Bloch sphere and rotate about the $x-$axis to obtain the $y-$coordinate.

3.2 Mathematical details of QST for one and two qubits

In this Section, we first provide mathematical details of the measurement of all components of a single-qubit density matrix and then discuss the two-qubit QST.

The relations Eq. (2.17) from Section 2.1.3 of Chapter 2 can easily be inverted, and the coefficients of the Bloch vector $\hat{r} = [r_x, r_y, r_z]$ can be expressed in terms of the density matrix

elements as

$$r_x = \rho_{01} + \rho_{10} = 2\mathrm{Re}(\rho_{10}),$$
$$r_y = i(\rho_{01} - \rho_{10}) = 2\mathrm{Im}(\rho_{10}), \tag{3.1}$$
$$r_z = \rho_{00} - \rho_{11} = 1 - 2\rho_{11}.$$

As explained in Section 2.2.5 of Chapter 2, as the result of SQUID measurements, it is possible to measure the "z-component" of the Bloch vector, which is the coefficient in front of the σ_z-matrix in the decomposition Eq. (2.16) from Section 2.2.5 of Chapter 2. As this coefficient equals r_z, the diagonal elements ρ_{11} and ρ_{00} of the density matrix, introduced in Eq. (2.6) in Section 2.1.2 of Chapter 2, can be measured.

In order to measure the real and imaginary coefficients of the off-diagonal elements of the density matrix Eq. (2.6), $\rho_{10} = \mathrm{Re}(\rho_{10}) + i\ \mathrm{Im}(\rho_{10})$, $\pm\pi/2$ rotations about the $x-$ and $y-$axes need to be applied. Mathematically such rotations are described by the rotation operators introduced in Eqs. (2.13)-(2.15). For example, measurement in the same computational basis, after rotation by the angle $\theta = \pi/2$ about $x-$axis, yields the imaginary part of the element ρ_{10} of the density matrix (note that $\mathrm{Im}(\rho_{10}) = -\mathrm{Im}(\rho_{01})$), because after this rotation the r_y-coefficient of the Bloch vector appears to be in front of the σ_z-matrix in the Pauli decomposition of the density matrix:

$$R_x^{\theta}\hat{\rho}(R_x^{\theta})^{\dagger} = \begin{pmatrix} \cos(\theta/2) & -i\ \sin(\theta/2) \\ -i\ \sin(\theta/2) & \cos(\theta/2) \end{pmatrix} \hat{\rho} \begin{pmatrix} \cos(\theta/2) & i\ \sin(\theta/2) \\ i\ \sin(\theta/2) & \cos(\theta/2) \end{pmatrix} \Big|_{\theta=\pi/2} =$$
$$= (r_0\sigma_0 + r_x\sigma_x - r_z\sigma_y + r_y\sigma_z)/2,$$

because

$$R_x^{\pi/2}\ \sigma_x\ (R_x^{\pi/2})^{\dagger} = \sigma_x, \qquad R_x^{\pi/2}\ \sigma_y\ (R_x^{\pi/2})^{\dagger}\ = \sigma_z, \qquad R_x^{\pi/2}\ \sigma_z\ (R_x^{\pi/2})^{\dagger}\ = -\sigma_y.$$

Analogously, measurement in the same basis, after rotation by the angle $\theta = -\pi/2$ about $y-$axis, yields the coefficient r_x of the Bloch vector (now under this rotation the coefficient r_x

appears to be in front of the σ_z-matrix), and, therefore, $\mathrm{Re}(\rho_{10})$ element of the density matrix (2.6) is measured:

$$R_y^{\theta}\,\hat{\rho}(R_y^{\theta})^{\dagger} = \begin{pmatrix} \cos(\theta/2) & -\sin(\theta/2) \\ \sin(\theta/2) & \cos(\theta/2) \end{pmatrix} \hat{\rho} \begin{pmatrix} \cos(\theta/2) & \sin(\theta/2) \\ -\sin(\theta/2) & \cos(\theta/2) \end{pmatrix}\Bigg|_{\theta=-\pi/2} =$$

$$= (r_0\sigma_0 - r_z\sigma_x + r_y\sigma_y + r_x\sigma_z)/2,$$

because

$$R_y^{-\pi/2}\,\sigma_x\,(R_y^{-\pi/2})^{\dagger} = \sigma_z, \qquad R_y^{-\pi/2}\,\sigma_y\,(R_y^{-\pi/2})^{\dagger} = \sigma_y, \qquad R_y^{-\pi/2}\,\sigma_z\,(R_y^{-\pi/2})^{\dagger} = -\sigma_x.$$

For multiple qubits, the expansion of the density matrix in terms of Pauli matrices can be generalized to $\hat{\rho} = \frac{1}{2^N}\sum_{i=0}^{3}\sum_{j=0}^{3}...\sum_{k=0}^{3} r_{ij...k}\,\sigma_i \otimes \sigma_j \otimes ... \otimes \sigma_k$, and the idea of the measurement remains the same: measure some coefficients $r_{ij...k}$, apply basis rotations, and measure other coefficients $r_{ij...k}$.

As we see, for a single qubit it is tomographically sufficient to use $n_R = 3$ basis rotations in order to measure all independent elements of the density matrix Eq. (2.6). Such rotations are the identity rotation $\mathbb{I}$, and two rotations about $x-$ and $y-$axes, $R_x^{\pi/2}$, and $R_y^{-\pi/2}$. This technique with $n_R = 3$ basis rotations has been used in the first QST experiment with one and two phase qubits at UCSB in Prof. J. Martinis' research group in 2006 [111, 112]. In case of two or more qubits in the system, multi-qubit rotations are built as n_R^N Kronecker products of single-qubit rotations (N is the number of qubits in the system), so for the case of two qubits we have in total nine rotations, while for the three-qubit systems we have twenty seven rotations.

In case of the SQUID measurement scheme realized with two phase qubits, as described in Section 2.2.5 of Chapter 2, we get four possible outcomes for each of the 9 measurements of the state of a two-qubit system, corresponding to the following four joint probabilities:

• probability P_{00} that both qubits are in the ground states,

• probability P_{01} that the first qubit is in the ground state, second qubit is in the excited state,

• probability P_{10} that the first qubit is in the excited state, second qubit is in the ground state,

• probability P_{11} that both qubits are in the excited states.

Therefore, if $n_R = 3$ basis rotations per qubit are used in the two-qubit tomography protocol, then the actual experimental data consists of 36 distinct probabilities: we have 9 different measurements with 4 outcomes each. In general, for an N-qubit system, 3^N different measurement settings with 2^N distinct outcomes each are required for the density matrix measurement.

More recently, an improved measurement scheme has been used in two-qubit experiments at UCSB with $n_{\rm R} = 6$ single-qubit rotations (see for example [113, 114]), e.g., $\mathcal{R}_{\rm meas} = \{\mathbb{I}, R_y^{\pm\pi/2}, R_x^{\pm\pi/2}, R_x^{\pi}\}$. In this protocol, which researchers at UCSB have named "overconstrained tomography" or "octomography", we get $n_R^2 = 36$ different two-qubit measurements, each with 4 possible outcomes, therefore the actual experimental data consists of 144 distinct probabilities. Such measurement scheme has advantages from the symmetry viewpoint and can reduce some experimental imperfections, like taking into account the population of the upper noncomputational level $\big|2\big\rangle$.

3.3 Linear inversion method of QST for two phase qubits

In this Section we explain the linear inversion method for the "overconstrained tomography" protocol of QST for two phase qubits ($n_R = 6$, $N = 2$) for the experimental data obtained at UCSB. One of the first tasks of the author of this dissertation was to write a Matlab computer code that performs QST for the experimental data from UCSB by the linear inversion method. As

explained in Chapter 3.2, the experimental data consists of 144 measured probabilites, which we denote as $P_{ij}^{\rm rot}$. On the other hand, as was explained in Section 2.1.2 of Chapter 2, a two-qubit state is described by a 4×4 density matrix $\hat{\rho}$ introduced in Eq. (2.7), with 16 unknown parameters r_{11}, r_{12},...,r_{44}. One can derive theoretical (symbolic) expressions $P_{ij}(r_{11}, ..., r_{44})$ for all 144 measured probabilities $P_{ij}^{\rm rot}$ in terms of 16 parameters r_{11}, r_{12},...,r_{44} using $n_R^{\ N} = 36$ two-qubit basis rotation operators U_j ($j = 1, ..., 36$). These 36 operators U_j are calculated as the Kronecker products of the single-qubit basis rotation operators from the set $\mathcal{R}_{\rm meas} = \{\mathbb{I}, R_x^{\pm\pi/2}, R_y^{\pm\pi/2}, R_x^{\pi}\}$ (operators R_x^{θ} and R_y^{θ} were introduced in Eqs. (2.13)-(2.15)).

The unitary evolution of the density operator under basis rotations U_j is described by equation Eq. (2.18) from Section 2.1.4 of Chapter 2, therefore the probability of the measurement outcome described by a measurement operator $\hat{M}_i$ (index $i = 1, ..., 4$ corresponds to four possible measurement outcomes P_{00}, P_{01}, P_{10}, P_{11}) can be calculated as

$$P_{ij}(r_{11}, ..., r_{44}) = {\rm Tr}(\hat{M}_i \, U_j \hat{\rho} \, U_j^{+}). \tag{3.2}$$

Hence we get 144 linear equations for $P_{ij}(r_{11}, ..., r_{44})$, corresponding to $j = 1, .., 36$ two-qubit basis rotations and $i = 1, ..., 4$ measurement outcomes for each basis rotation, in terms of 16 unknown parameters r_{11}, r_{12},...,r_{44}. We can set each of these symbolic equations for $P_{ij}(r_{11}, ..., r_{44})$ equal to the experimental probabilities $P_{ij}^{\rm rot}$. It is convenient to arrange 144 probabilities $P_{ij}^{\rm rot}$ and 16 density matrix elements r_{mn} in columns, which allows us to write these 144 equations in matrix

form as

$$\underbrace{\begin{pmatrix} P_{00}^{I\otimes I} \\ P_{01}^{I\otimes I} \\ P_{10}^{I\otimes I} \\ P_{11}^{I\otimes I} \\ P_{00}^{I\otimes R_x^{\pi/2}} \\ P_{01}^{I\otimes R_x^{\pi/2}} \\ P_{10}^{I\otimes R_x^{\pi/2}} \\ P_{11}^{I\otimes R_x^{\pi/2}} \\ P_{00}^{I\otimes R_y^{-\pi/2}} \\ \dots \\ P_{11}^{R_x^{\pi}\otimes R_x^{\pi}} \end{pmatrix}}_{(144\,\mathrm{x}\,1)} = \underbrace{\begin{pmatrix} B_{1,1} & B_{1,2} & \dots & B_{1,16} \\ B_{2,1} & B_{2,2} & \dots & B_{2,16} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ B_{144,1} & B_{144,2} & \dots & B_{144,16} \end{pmatrix}}_{(144\,\mathrm{x}\,16)} \underbrace{\begin{pmatrix} r_{11} \\ r_{12} \\ r_{13} \\ r_{14} \\ r_{21} \\ \dots \\ r_{44} \end{pmatrix}}_{(16\,\mathrm{x}\,1)} \tag{3.3}$$

where $\mathbf{B}$ is a $144\,\mathrm{x}\,16$ transformation matrix, whose entries are calculated theoretically considering operators of basis rotations and projections to four possible measurement outcomes as explained above, and explicit two-qubit basis rotations are written in the superscript indices of expressions for probabilities P in the left column of Eq. (3.3). It is convenient to rewrite Eq. (3.3) in a compact matrix notation

$$\vec{P} = \mathbf{B}\vec{r}, \tag{3.4}$$

where $\vec{P}$ is a column-vector consisting of 144 measured probabilities, $\vec{r}$ is a column-vector consisting of 16 density matrix elements $r_{11}, ..., r_{44}$, which we need to determine, and $\mathbf{B}$ is a $144\,\mathrm{x}\,16$ transformation matrix which is calculated theoretically.

In order for Eq. (3.4) to have the exact unique solution, the transformation matrix $\mathbf{B}$ needs to have an inverse, which is possible only if the number of rows (number of measurements) in $\mathbf{B}$ is the same as the number of columns (the number of density matrix elements, d^2) in matrix $\mathbf{B}$. Usually in the case of full tomography the number of measurements significantly exceeds the number of density matrix elements, therefore Eq. (3.4) is the overdetermined system of equations for determining the unknown d^2 elements of the density matrix. For example, for the system of two qubits we have 144 equations for determining the unknown 16 elements $r_{11}, ..., r_{44}$ of the density matrix. In this situation, when the problem is overdetermined, the method of solution of overdetermined problem from linear regression theory [115] can be employed. Multiplying both sides of Eq. (3.4) by $\mathbf{B}^\dagger$, we arrive at

$$\mathbf{B}^\dagger \vec{P} = (\mathbf{B}^\dagger \mathbf{B})\vec{r}, \tag{3.5}$$

where $\mathbf{B}^\dagger \vec{P}$ is a d^2-dimensional column vector and $\mathbf{B}^\dagger \mathbf{B}$ is a square matrix of size $d^2 \times d^2$. Therefore, the equation Eq. (3.4) can be rewritten as

$$(\mathbf{B}^\dagger \mathbf{B})^{-1} \mathbf{B}^\dagger \vec{P} = \vec{r}. \tag{3.6}$$

Using the last equation Eq. (3.6), $\vec{r}$ can be found as the density matrix that minimizes the sum of the variances between the measured experimental data and the analytical expectation values for corresponding probabilities. Therefore the linear inversion method of QST consists of finding the least-square estimator of the density matrix from the overdetermined system of equations Eq. (3.4), which can be done, for example, using the "mldivide" operation in Matlab. The Matlab script, implementing this algorithm of linear inversion, was written by the author of this dissertation at the initial stage of his work. Example of the density matrix reconstruction for the experimental data for the CZ gate realized with the phase qubits is presented in Fig. 3.1.

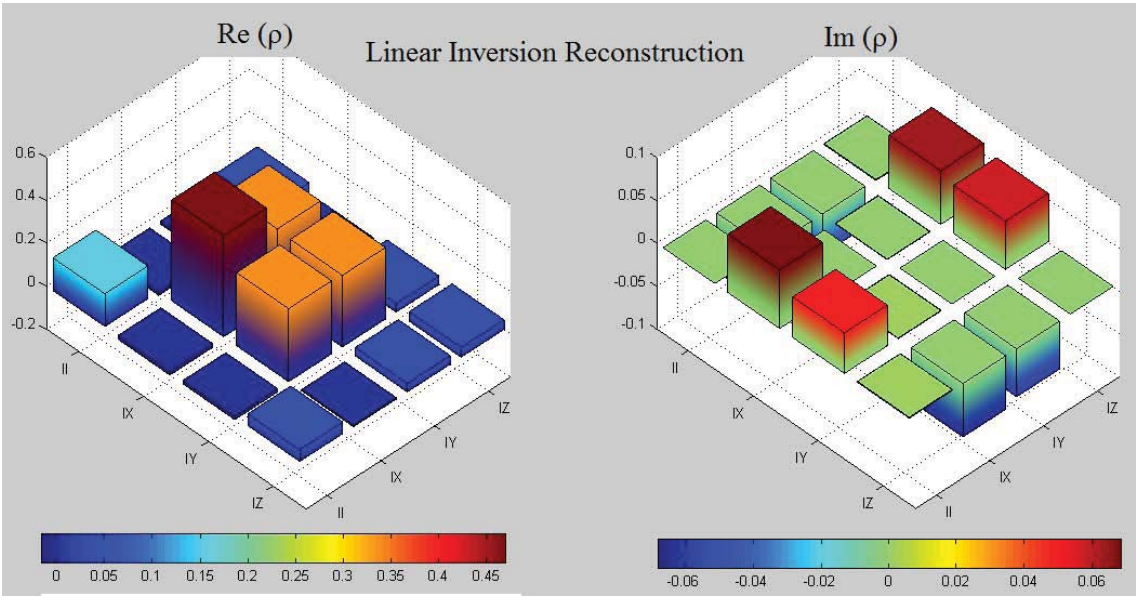

Figure 3.1: Results for the reconstruction of the density matrix by the linear inversion method of QST. Experimental data for the CZ gate based on the phase qubits have been used.

The linear inversion solution for the density matrix can produce results that do not satisfy some of the basic properties of a valid density matrix, because in this technique no positivity constraint or trace-normalization constraint is imposed. For instance, such solution could give negative probabilities or probabilities greater than one for some measurement outcomes. Mathematically it is equivalent to the violation of the condition that all eigenvalues of the physical density matrix must lie within the interval $[0, 1]$. To avoid this problem, the maximum likelihood estimation of the density matrix can be employed.

3.4 Maximum likelihood estimation of the density matrix

As we have just mentioned in the previous Section 3.3, methods of reconstruction of the density matrix based on the linear inversion may lead to certain physical artifacts such as the negative eigenvalues of the reconstructed density matrix. In order to avoid these physical artifacts, an

estimation method, based on statistical maximum-likelihood principle, has been proposed in [116]. In this approach, called maximum likelihood estimation, the positive semidefiniteness and trace normalization of the reconstructed density matrix are guaranteed, because these necessary conditions are incorporated into the parametrization of the density matrix. The first application of this method to the problem of density matrix reconstruction was done in optics by James *et. al.* in [51]. One of the tasks of the author of this thesis, in the beginning of his work, was the development of the computer Matlab scripts implementing the maximum-likelihood estimation of the density matrices for the experimental data for CZ gates based on phase qubits, obtained at Prof. J. Martinis' group at UCSB.

Maximum likelihood estimation is a widely used statistical technique for estimating some quantity based on a set of measurements. For any given value of the quantity, we calculate the "likelihood" that this value would produce the observed measurement results, consistent with the measured dataset, the error model, and a parametrized model of the system. Using a numerical search, the estimated quantity is then taken to be that for which the likelihood of the observed measurements is maximized. In case of the measurements on two phase qubits, using the "over-constrained tomography" protocol as it was explained in the previous Section 3.3, we have 144 measurement results (144 probabilities), and the quantity that we estimate are 16 real elements of the density matrix. We can get the theoretical expressions for each of 144 measured probabilities, as in Eq. (3.2) in the previous Section 3.3, and find a set of 16 parameters in the density matrix, which gives the best "likelihood" that we get these 144 measured probabilities.

As a reminder, for any density matrix to be physical, it must satisfy the three properties: be normalized, be hermitian, and have non-negative eigenvalues (due to the normalization condition

the eigenvalues cannot exceed one). The algorithm of construction of the density matrix satisfying all three conditions for the case of a quantum gate realized with two phase qubits and measured using "overconstrained tomography" protocol, is the following:

• Generate a formula for an explicitly "physical" density matrix, i.e., a matrix that has the three important properties of normalization, Hermicity, and positivity. This matrix is a function of 16 real variables denoted $t_1, t_2, ..., t_{16}$. We denote the density matrix as $\hat{\rho}(t_1, t_2, ..., t_{16})$.

• Introduce a "likelihood function" which quantifies how good the density matrix $\hat{\rho}(t_1, t_2, ..., t_{16})$ is in relation to the experimental data. This likelihood function is a function of the 16 real parameters t_n and of the 144 experimentally measured probabilities P_k ($k = 1, ..., 144$) (these probabilities were denoted as P_{ij}^{rot} in the previous Section 3.3). We will denote this function as $\hat{\mathcal{L}}(t_1, t_2, ..., t_{16}; P_1, P_2, ..., P_{144})$.

• Using standard numerical optimization techniques, find the optimum set of variables $t_1^{(opt)}, t_2^{(opt)}, ..., t_{16}^{(opt)}$ for which the function $\hat{\mathcal{L}}(t_1, t_2, ..., t_{16}; P_1, P_2, ..., P_{144})$ has its maximum value. The best estimate for the density matrix is then $\hat{\rho}(t_1^{(opt)}, t_2^{(opt)}, ..., t_{16}^{(opt)})$.

It is convenient to parametrize the density matrix as a function of 16 real parameters t_1, t_2, ..., t_{16} in the following way:

$$\hat{\rho}(t) = \frac{T^{+}(t)T(t)}{\mathrm{Tr}\{T^{+}(t)T(t)\}}. \tag{3.7}$$

This parametrization guarantees that the density matrix $\hat{\rho}(t)$ satisfies all three mathematical properties of a "physical" density matrix. Firstly, any matrix that can be written in the form $\hat{G} = \hat{T}^{+}\hat{T}$ is non-negative definite. To see this, note that mathematically the property of non-negative definiteness for any matrix $\hat{G}$ can be written as

$$\langle\psi|\hat{G}|\psi\rangle \geq 0 \quad \forall \ |\psi\rangle,$$

which is automatically true for a matrix $\hat{G}$ parametrized in the way $\hat{G} = \hat{T}^{+}\hat{T}$:

$$\langle\psi|\hat{T}^{+}\hat{T}|\psi\rangle = \langle\psi'|\psi'\rangle \geq 0, \text{ where } |\psi'\rangle = \hat{T}'|\psi\rangle.$$

Secondly, matrix $\hat{G} = \hat{T}^{+}\hat{T}$ is automatically Hermitian: $(\hat{T}^{+}\hat{T})^{+} = \hat{T}^{+}(\hat{T}^{+})^{+} = \hat{T}^{+}\hat{T}$. Finally, to insure normalization, we divide by the trace.

It is useful and convenient to choose a lower-triangular (tridiagonal) form for $\hat{T}$:

$$T(t) = \begin{pmatrix} t_1 & 0 & 0 & 0 \\ t_5 + it_6 & t_2 & 0 & 0 \\ t_{11} + it_{12} & t_7 + it_8 & t_3 & 0 \\ t_{15} + it_{16} & t_{13} + it_{14} & t_9 + it_{10} & t_4 \end{pmatrix} \tag{3.8}$$

with the total of 16 entries for the 4×4 two-qubit case, and where there are four real parameters on the main diagonal. This form of parametrization is based on the Cholesky decomposition of a Hermitian, positive-definite matrix into the product of a lower triangular matrix and its conjugate transpose in linear algebra. It has been shown by Andre-Louis Cholesky that every Hermitian positive definite matrix $\mathbf{A}$ can be decomposed into the product of a lower triangular matrix $\mathbf{L}$ with strictly positive diagonal entries and the Hermitian conjugated matrix $\mathbf{L}^{*}$: $\mathbf{A} = \mathbf{L}\mathbf{L}^{*}$. The Cholesky decomposition is unique: given a Hermitian, positive-definite matrix $\mathbf{A}$, there is only one lower triangular matrix $\mathbf{L}$ with strictly positive diagonal entries such that $\mathbf{A} = \mathbf{L}\mathbf{L}^{*}$. There is a standard Matlab command "chol" which calculates the Cholesky decomposition of a positive-definite matrix.

In order to write an expression for the likelihood function $\hat{\mathcal{L}}(t_1, t_2, ..., t_{16}; P_1, P_2, ..., P_{144})$, which we use in the maximum likelihood estimation algorithm, we need to establish the probability distribution that describes the measurement scheme. As explained in Chapter 3.2, there are

36 different two-qubit basis rotations U_j involved in the "overconstrained tomography" measurement protocol for a quantum gate based on two phase qubits, and measurement of the state of the qubit after each of those 36 rotations U_j can yield four different probabilities $P_{00}^{\rm rot}$, $P_{01}^{\rm rot}$, $P_{10}^{\rm rot}$ and $P_{11}^{\rm rot}$. Therefore the multinomial distribution with four outcomes ("tetranomial" distribution) needs to be employed for the description of the measurement scheme.

The tetranomial distribution can be viewed as a generalization of the well-known binomial distribution, which describes the measurements with two possible outcomes only. First, let us rewrite the well-known binomial distribution as

$$P(k_1, k_2; n, p_1, p_2) = \frac{n!}{k_1!\, k_2!} p_1^{k_1}\, p_2^{k_2}, \tag{3.9}$$

where it is presumed that we have observed k_1 outcomes of type $\#1$ with underlying probability p_1 and k_2 different outcomes of type $\#2$ with underlying probability p_2. Obviously, $p_1 + p_2 = 1$ and $k_1 + k_2 = n$, where n is the total number of trials. It is now easy to generalize this formula for the case of measurement with four possible outcomes. The tetranomial distribution gives the analytical expression for the probability that we have observed k_1 outcomes of type $\#1$ with underlying probability p_1, k_2 outcomes of type $\#2$ with underlying probability p_2, k_3 outcomes of type $\#3$ with underlying probability p_3 and k_4 outcomes of type $\#4$ with underlying probability p_4:

$$P(k_1, k_2, k_3, k_4; n, p_1, p_2, p_3, p_4) = \frac{n!}{k_1!\, k_2!\, k_3!\, k_4!} p_1^{k_1}\, p_2^{k_2}\, p_3^{k_3}\, p_4^{k_4}. \tag{3.10}$$

Obviously, $p_1 + p_2 + p_3 + p_4 = 1$ and $k_1 + k_2 + k_3 + k_4 = n$, where n is the number of repetitions of measurements for each of 36 basis rotations (for example, $n = 1500$ for the specific tomography experiments at UCSB which were used to obtain the experimental data with phase qubits, used for calculations in this dissertation). Let us call this probability $P(k_1, k_2, k_3, k_4; n, p_1, p_2, p_3, p_4)$ as the

likelihood L that the result of type j with the underlying probability p_j was observed k_j times ($j = 1, 2, 3, 4$). Note that the expression Eq. (3.10) as the product of four terms describes the probability distribution after only one two-qubit basis rotation U_j (operators U_j are built as the Kronecker products of the single-qubit basis rotation operators from the set $\mathcal{R}_{\rm meas} = \{\mathbb{I}, R_x^{\pm\pi/2}, R_y^{\pm\pi/2}, R_x^{\pi}\}$ as explained in Section 3.3). As there are 36 distinct two-qubit basis rotations U_j involved in the tomography experiment, each yielding the four possible outcomes, the full likelihood is the product of 144 terms.

For the convenience of numerical calculations, rather than finding the maximum value of L, it is easier to find the minimum of its logarithm taken with negative sign. Taking logarithm of L, we have

$$\ln L = \ln\Big(\frac{n!}{k_1!\,k_2!\,k_3!\,k_4!}\Big) + k_1 \ln p_1 + k_2 \ln p_2 + k_3 \ln p_3 + k_4 \ln p_4, \tag{3.11}$$

where the first term is independent of p_j, so it can be omitted. Dividing the other terms in Eq. (3.11) by the number of repetitions of experiment n for each of 36 basis rotations (where for the specific experiment $n = k_1 + k_2 + k_3 + k_4 = 1500$), we get the following functional (it is convenient from computational point of view to take the natural logarithm of the resulting functional):

$$-\ln L = -\Big\{\frac{k_1}{n}\ln p_1 + \frac{k_2}{n}\ln p_2 + \frac{k_3}{n}\ln p_3 + \frac{k_4}{n}\ln p_4\Big\}. \tag{3.12}$$

The terms $\dfrac{k_j}{n}$ in Eq. (3.12) have the meaning of experimentally measured probabilities, and p_j in Eq. (3.12) can be calculated theoretically from the elements of the density matrix parametrized through parameters t_1, t_2, ..., t_{16}. Therefore we need to minimize the functional $\mathcal{L}$

$$\mathcal{L} \equiv -\ln L = -\sum_{j=1}^{144}\Big\{p_j^{(\rm exp)} * \ln p_j^{(\rm th)}(t_1, t_2, ..., t_{16})\Big\} \tag{3.13}$$

over real parameters t_1, t_2,..., t_{16}, where $p_j^{(\rm exp)}$ are the 144 measured probabilities.

The author of this dissertation wrote the Matlab computer program that finds the minimum of the likelihood function Eq. (3.13) and calculates the density matrices by the maximum likelihood method for the experimental data for the CZ gates based on phase qubits. In our codes, we used the Matlab built-in function "*fminunc*" which requires an initial estimate for the values of t_1, t_2,..., t_{16}. For this, the tomographic estimate of the density matrix obtained by the linear inversion method, described in the previous Section 3.3, has been used. We checked that all eigenvalues of the density matrix obtained by the linear inversion method were positive, and in case some of them were negative, we assigned small positive values to those eigenvalues, then we used the Matlab built-in function "*chol*" to find the Cholesky decomposition of the density matrix obtained by the linear inversion method, which gave the initial guess for the real parameters t_1, t_2,..., t_{16}. The resulting density matrices, obtained by the maximum likelihood method, have been fully physical. Example of a density matrix, calculated by the maximum likelihood method for the experimental data for the CZ gate based on phase qubits, is presented in Fig. 3.2. Note that usually the difference between the density matrices calculated by the linear inversion method and by the maximum likelihood method is not very significant. For example, the state fidelity (defined in Eq. (2.24) of Chapter 2) between the two density matrices, presented in Figs. 3.1 and 3.2, equals $F = 0.998$, and the trace distance (defined in Eq. (2.25) of Chapter 2) between these density matrices equals $D = 0.008$. This means that the density matrices reconstructed by the two different methods are slightly different, but they are not too far from each other.

It is shown in statistics that binomial and multinomial distributions can be approximated by the Gaussian (normal) distribution for large n. Such transition is based on the Stirling formula $\ln n! \approx n \ln n - n + O(\log n)$, which allows one to replace factorials in the multinomial distribution

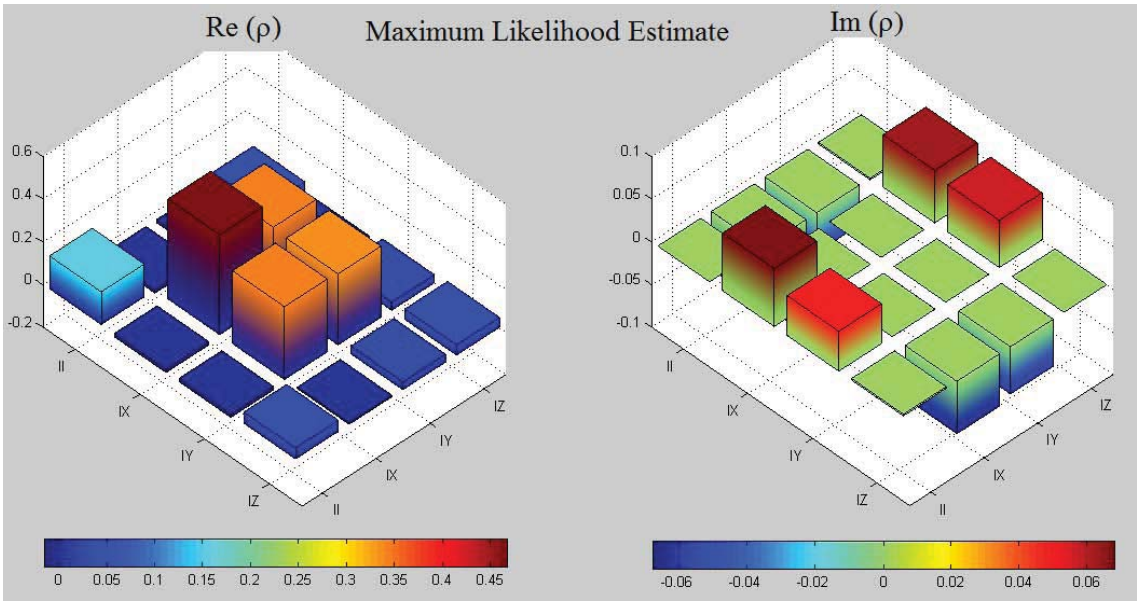

Figure 3.2: Results for the reconstruction of the density matrix by the maximum likelihood method of QST. Experimental data for the CZ gate based on the phase qubits have been used.

by $n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$. The rule of thumb that is mentioned in statistics about the validity of Gaussian approximation is that both products $n * p$ and $n * (1-p)$ should be greater than 5, which is usually the case in the tomography experiments at UC Santa Barbara, as $n = 1500$. Therefore the exact multinomial likelihood functional can be replaced by the approximate Gaussian one,

$$L(t_1, t_2, ..., t_{16}) = -\sum_{j=1}^{144} \frac{\left\{p_j^{(\exp)} - p_j^{(\mathrm{th})}(t_1, t_2, ..., t_{16})\right\}^2}{2p_j^{(\exp)}}. \tag{3.14}$$

This possibility of transition to Gaussian probability distribution is mentioned in a theoretical paper [117], and Gaussian distribution is always used in optics. In particular, Gaussian distribution Eq. (3.14) was used in the first optical implementation of the method [51].

Chapter 4

Standard Quantum Process Tomography of multi-qubit gates

4.1 Basics of "standard" quantum process tomography

Quantum Process Tomography is the procedure used to characterize a quantum gate, or quantum process, which takes an arbitrary given input state ρ^{in} and transforms it into the output state $\rho^{\text{fin}} = \mathcal{E}(\rho^{\text{in}})$. Usually, such task arises when we want to understand how closely an experimentally realized gate approximates the desired unitary operation. This Chapter is structured as follows. In this Section 4.1 we are giving an overall introduction into quantum process tomography, we also provide the details of the experimental setup of QPT for superconducting qubit gates in Section 4.2, and explain in Section 4.3 the so-called linear inversion method of QPT, which was programmed in Matlab by the author of this dissertation. We give the details of programming of the linear inversion method, discuss computational resources, such as time and memory, and then discuss in Section 4.4

alternative methods for the process matrix reconstruction, such as the maximum likelihood and the least squares method.

The theoretical fundamentals of QPT were formulated in [12, 13] in 1997. Quantum process tomography was first applied to an NMR system[1] in 2001 [118, 119, 120], then it was used to study optical quantum gates in 2003 and 2004 [121, 122], after that the first process tomography experiment in ion trap systems was done in 2006 [123, 124]. The first process tomography experiment for superconducting phase qubits was performed for a memory operation between a single phase qubit and a two-level state in 2008 [125], and the first demonstration of QPT in a system of two superconducting phase qubits was carried out in 2010 [126]. The method of quantum process tomography was also applied to superconducting qubits in [36, 114, 126, 127, 128].

The idea behind QPT is to reconstruct a quantum operation $\rho^{\text{in}} \mapsto \rho^{\text{fin}} = \mathcal{E}(\rho^{\text{in}})$ from experimental data. The quantum operation is a completely positive linear map $\mathcal{E}$ that transforms the density operator ρ^{in} into the output density operator ρ^{fin}. Then for an N-qubit system prepared in the state with density matrix ρ^{in}, the quantum operation $\mathcal{E}$ can be expressed in terms of the so-called Kraus operators as [2]

$$\mathcal{E}(\rho^{\text{in}}) = \sum_{j=1}^{K} A_j \rho^{\text{in}} A_j^{\dagger}, \tag{4.1}$$

where A_j are called "Kraus operators" of the quantum operation $\mathcal{E}$, and the number K of such operators does not exceed d^2, where $d = 2^N$ is the dimension of the system. The representation Eq. (4.1) is also known as the "operator-sum representation". Since the quantum operation $\mathcal{E}$ must be trace-preserving (or trace non-increasing), the Kraus operators A need to satisfy the condition

$$\sum_{j=1}^{K} A_j^{\dagger} A_j = I. \tag{4.2}$$

[1] the fidelity of a CNOT logic gate in an NMR apparatus was measured in Ref. [118].

To prove this trace-preservation condition Eq. (4.2), note that

$$1 = \mathrm{Tr}\,\mathcal{E}(\rho) = \mathrm{Tr}\Big(\sum_j A_j \rho^{\,\mathrm{in}} A_j^\dagger\Big) = \sum_j \mathrm{Tr}\Big(A_j \rho^{\,\mathrm{in}} A_j^\dagger\Big) = \mathrm{Tr}\Big(\sum_j A_j^\dagger A_j\, \rho^{\,\mathrm{in}}\Big), \tag{4.3}$$

which must be true for all input density matrices ρ^{in}, therefore the condition Eq. (4.2) needs to be satisfied. In case of a trace non-increasing quantum operation, the condition Eq. (4.2) becomes $\sum_{j=1}^{K} A_j^\dagger A_j \le I$.

The decomposition Eq. (4.1) of the quantum operation $\mathcal{E}$ into the Kraus operators is not unique as different sets of Kraus operators may describe the same quantum operation $\mathcal{E}$. Therefore, it is more convenient to introduce the following unique expansion of the quantum operation $\mathcal{E}$ over some fixed set of basis operators E_α:

$$\mathcal{E}(\rho^{\mathrm{in}}) = \sum_{\alpha,\beta=1}^{d^2} \chi_{\alpha\beta} E_\alpha \rho^{\mathrm{in}} E_\beta^\dagger, \tag{4.4}$$

where $d = 2^N$ is the dimension of the system, $\chi \in \mathbb{C}^{d^2\times d^2}$ is the process matrix and $\{E_\alpha \in \mathbb{C}^{d\times d}\}$ is a chosen basis of operators. We assume that this basis is orthogonal, $\langle E_\alpha | E_\beta\rangle \equiv \mathrm{Tr}(E_\alpha^\dagger E_\beta) = Q\,\delta_{\alpha\beta}$, where $Q = d$ for the Pauli basis and Pauli-error basis, and $Q = 1$ for the SVD basis (see Appendices B and C). Note that for a trace-preserving operation $\mathrm{Tr}(\chi) = 1$ if $Q = d$, while $\mathrm{Tr}(\chi) = d$ if $Q = 1$. We implicitly assume in this and the next Chapters of this dissertation the usual normalization $Q = d$, unless mentioned otherwise. The process matrix χ is positive semidefinite (which implies being Hermitian), and we also assume it to be trace-preserving,

$$\chi \ge 0 \quad \text{(positive semidefinite)}, \tag{4.5}$$

$$\sum_{\alpha,\beta=1}^{d^2} \chi_{\alpha\beta} E_\beta^\dagger E_\alpha = \mathbb{I}_d \quad \text{(trace-preserving)}. \tag{4.6}$$

These conditions ensure that $\rho^{\text{fin}} = \mathcal{E}(\rho^{\text{in}})$ is a legitimate density matrix for an arbitrary (legitimate) input state ρ^{in}. The condition Eq. (4.6) reduces the number of real independent parameters in χ from d^4 to $d^4 - d^2$. For example, for one qubit the 4×4 process matrix χ contains 12 independent elements, for two qubits the 16×16 process matrix contains 240 independent elements, and for three qubits the number of independent elements in the 64×64 process matrix becomes 4032. Hence the number of parameters needed to fully specify the quantum map $\mathcal{E}$ scales as $O(16^N)$ with the number of qubits N. Note that the set of allowed process matrices χ defined by Eqs. (4.5) and (4.6) is convex [29, 39]. The elements of the process matrix χ can be related to Kraus operators A as $\chi_{m,n} = \sum_{j=1}^{K} c_{j,m} c_{j,n}^*$, where $c_{j,m}$ are the coefficients of decomposition of Kraus operators A over the chosen basis of operators $\{E_\alpha\}$: $A_j = \sum_{m=1}^{d^2} c_{j,m} E_m$.

The essential idea of standard QPT is to exploit the linearity of the map Eq. (4.4) by preparing the qubits in different initial states, then applying the quantum gate $\mathcal{E}$, and then measuring a set of observables until the collected data allows us to obtain the process matrix χ through matrix inversion or other methods. More precisely, if the qubits are prepared in the state ρ_k^{in}, then the probability of finding them in the (measured) state $|\phi_i\rangle$ after applying the gate is given by

$$P_{ik} = \text{Tr}(\Pi_i \mathcal{E}(\rho_k^{\text{in}})) = \sum_{\alpha,\beta} \text{Tr}(\Pi_i E_\alpha \rho_k^{\text{in}} E_\beta^\dagger) \chi_{\alpha\beta}, \tag{4.7}$$

where $\Pi_i = |\phi_i\rangle \langle\phi_i|$. By preparing the qubits in one of the linearly independent input states $\{\rho_1^{\text{in}}, \ldots, \rho_{N_{\text{in}}}^{\text{in}}\}$ and performing a series of projective measurements $\{\Pi_1, \ldots, \Pi_{N_{\text{meas}}}\}$ on the output states, using the methods of quantum state tomography described in Chapter 3, one obtains a set of $m = N_{\text{in}} N_{\text{meas}}$ probabilities $\{P_{ik}\}$ which, using Eq. (4.7), may be written as

$$\vec{P}(\chi) = \Phi \vec{\chi}, \tag{4.8}$$

where $\vec{P}(\chi) \in \mathbb{C}^{m \times 1}$ and $\vec{\chi} \in \mathbb{C}^{d^4 \times 1}$ are vectorized forms of $\{P_{ik}\}$ and $\chi_{\alpha\beta}$, respectively. The

$m \times d^4$ transformation matrix Φ has entries given by $\Phi_{ik,\alpha\beta} = \mathrm{Tr}(\Pi_i E_\alpha \rho_k^{\mathrm{in}} E_\beta^\dagger)$. Note the similitude between the last equation Eq. (4.8) relating the elements of the unknown process matrix χ with the column-vector of the the measured probabilities in the QPT problem and the equation Eq. (3.4) from Section 3.3 of Chapter 3, which related the elements of the unknown elements of the density matrix with the column-vector of the measured probabilities in the QST problem.

4.2 Experimental details of QPT of multi-qubit superconducting gates

There are several different ways to perform standard QPT for an N-qubit quantum gate realized with superconducting qubits [113, 126, 129, 128, 130, 131, 132, 114]. The differences are the following. First, it can be performed using either $n_{\mathrm{in}} = 4$ initial states for each qubit [126, 129, 128, 130], e.g., $\{|0\rangle, |1\rangle, (|0\rangle + |1\rangle)/\sqrt{2}, (|0\rangle + i\,|1\rangle)/\sqrt{2}\}$, or using $n_{\mathrm{in}} = 6$ initial states per qubit [131, 132], e.g., $\{|0\rangle, |1\rangle, (|0\rangle \pm |1\rangle)/\sqrt{2}, (|0\rangle \pm i\,|1\rangle)/\sqrt{2}\}$, so that the total number of initial states is $N_{\mathrm{in}} = n_{\mathrm{in}}^N$. Multiqubit initial states are built as Kronecker product (also called tensor product or direct product) of single-qubit input states. (It is tomographically sufficient to use $n_{\mathrm{in}} = 4$, but the set of 6 initial states is more symmetric, so it can reduce the effect of experimental imperfections.) Second, the final measurement of the qubits can be realized in the computational basis after one out of $n_{\mathrm{R}} = 3$ rotations per qubit [126, 128], e.g., $\mathcal{R}_{\mathrm{meas}} = \{\mathbb{I}, R_y^{-\pi/2}, R_x^{\pi/2}\}$, or $n_{\mathrm{R}} = 4$ rotations [36, 129, 132], e.g., $\mathcal{R}_{\mathrm{meas}} = \{\mathbb{I}, R_y^{\pi}, R_y^{\pi/2}, R_x^{\pi/2}\}$, or $n_{\mathrm{R}} = 6$ rotations [113, 131, 114], e.g., $\mathcal{R}_{\mathrm{meas}} = \{\mathbb{I}, R_y^{\pi}, R_y^{\pm\pi/2}, R_x^{\pm\pi/2}\}$. This gives $N_{\mathrm{R}} = n_{\mathrm{R}}^N$ measurement "directions" in the Hilbert space. Again, rotations in multiqubit systems are built as the Kronecker product of single-qubit rotations. Third, it may be possible to measure the state of each qubit simultaneously [126, 113,

128], so that the probabilities for all 2^N outcomes are measured, or it may be technically possible to measure the probability for only one state (say, $|0...0\rangle$) or a weighed sum of the probabilities [131, 129, 130]. Therefore, the number of measured probabilities for each configuration is either $N_{\rm prob} = 2^N$ (with $2^N - 1$ independent probabilities, since their sum is equal 1) or $N_{\rm prob} = 1$. Note that if $N_{\rm prob} = 2^N$, then using $n_{\rm R} = 6$ rotations per qubit formally gives the same probabilities as for $n_{\rm R} = 3$, and in an experiment this formal symmetry can be used to improve the accuracy of the results. In contrast, in the case when $N_{\rm prob} = 1$, the use of $n_{\rm R} = 4$ or $n_{\rm R} = 6$ is natural for the complete tomography.

Thus, the number of measurement configurations (including input state and rotations) in standard QPT is $M_{\rm conf} = N_{\rm in} N_{\rm R} = n_{\rm in}^N n_{\rm R}^N$, while the total number of probabilities in the data set is $M = M_{\rm conf} N_{\rm prob}$. This number of probabilities can be as large as $M = 72^N$ for $n_{\rm in} = 6$, $n_{\rm R} = 6$, and $N_{\rm prob} = 2^N$ (with $72^N - 36^N$ independent probabilities). Since only $16^N - 4^N$ independent probabilities are necessary for the standard QPT, a natural choice for a shorter experiment is $n_{\rm in} = 4$, $n_{\rm R} = 3$, and $N_{\rm prob} = 2^N$; then the total number of probabilities is $M = 24^N$, with $24^N - 12^N$ independent probabilities. If $N_{\rm prob} = 1$ due to the limitations of the measurement technique, then the natural choices are $n_{\rm in} = 4$ and $n_{\rm R} = 4$, giving $M = 16^N$ or $n_{\rm in} = 4$ and $n_{\rm R} = 6$, giving $M = 24^N$.

In this and the following Chapters of this dissertation we focus on the case $n_{\rm in} = 4$, $n_{\rm R} = 3$, and $N_{\rm prob} = 2^N$. Based on this, for a two-qubit quantum gate there are $M_{\rm conf} = 12^N = 144$ measurement configurations and $M = 24^N = 576$ probabilities (432 of them independent).

4.3 Linear inversion method of quantum process tomography

4.3.1 The method of linear inversion for QPT

In principle, for tomographically complete sets of input states $\{\rho_1^{\text{in}}, \ldots, \rho_{N_{\text{in}}}^{\text{in}}\}$ and measurement operators $\{\Pi_1, \ldots, \Pi_{N_{\text{meas}}}\}$, one could invert Eq. (4.8) and thus uniquely find the process matrix χ by using the experimental set of probabilities $\vec{P}^{\text{exp}}$. This method of linear inversion in which the process matrix χ is represented in a block form as the product of several other auxiliary matrices, is described for the systems of one and two qubits in [2, 13, 25], and is formulated slightly differently in [17]. We wrote the Matlab computer code which implements the linear inversion method of [17], and we discuss this algorithm in this Section.

The algorithm of linear inversion, formulated in [17], describes a two-step procedure of calculation of the process matrix χ. First, it is convenient to compute some auxiliary matrix J of dimension $4^N \times 4^N$, called *Jamiolkowski operator*, and then in the second step of the algorithm convert J into the process matrix χ. As explained in Section 4.1, in the quantum process tomography procedure one needs to prepare n_{in}^N linearly independent initial states ρ^{in}, which are traditionally chosen as all possible Kronecker products of the single-qubit states, perform the evolution $\mathcal{E}$, and measure the resulting states ρ^{fin}, using the quantum state tomography procedure. Note that different number of basis rotations per qubit, $n_R = 3$ or $n_R = 4$, or $n_R = 6$, can be used in the algorithm. In order to calculate matrix J, in the first step of the linear inversion method of QPT, it is convenient to introduce $4^N \times 4^N$ matrices R and R_0 constructed from the density matrices ρ^{fin} and ρ^{in}, reshaped into the columns, so that the nth column of R is the nth density matrix ρ^{fin} reshaped into a column, and similar for R_0. Matrices R and R_0 can be related through the $4^N \times 4^N$ matrix $\mathcal{L}$: $R = \mathcal{L}R_0$.

Therefore the matrix $\mathcal{L}$ can be calculated as

$$\mathcal{L} = RR_0^{-1}. \tag{4.9}$$

Jamiolkowski matrix J consists of the reordered elements of $\mathcal{L}$. The reordering procedure is the following [17]: (1) each row of $\mathcal{L}$ is converted into a $2^N \times 2^N$ matrix by sequentially placing the rows of 2^N elements below each other and (2) these matrices are placed from left to right, with a new row of matrices starting after each set of 2^N steps. Therefore, the first step of the algorithm of the χ-matrix calculation by the method of linear inversion consists in calculation of the matrix $\mathcal{L}$ and its reordering, which yields the $4^N \times 4^N$ Jamiolkowski matrix J. In the second step of the algorithm, J needs to be transformed into the process matrix χ. As stated in [17], J matrix coincides with the process matrix χ in a special unphysical by-element basis $|i\rangle\langle j|$, the so-called elementary basis. Therefore, to obtain the process matrix χ for a physical operator basis E_n, one needs to make the following transformation of J. First, a matrix $\boldsymbol{E}$ of the dimension $4^N \times 4^N$ must be constructed from the elements of operator basis matrices E_n, which are reshaped into columns. Namely, the nth column of matrix $\boldsymbol{E}$ contains all elements of the $2^N \times 2^N$ matrix E_n. After calculating the matrix $\boldsymbol{E}$, some matrix multiplication of J, $\boldsymbol{E}$ and $\boldsymbol{E}^\dagger$ needs to be done for the calculation of χ-matrix. For the case of mutually orthogonal basis operators E_n (such as the Pauli basis operators), the final computational formula for the χ-matrix has the following block structure:

$$\chi = \boldsymbol{E}^\dagger J \boldsymbol{E}/4^N. \tag{4.10}$$

We have written the Matlab code which implements the above two-step linear inversion algorithm of the process matrix calculation. It is convenient to split the calculations into two separate scripts (files), because some matrices in the algorithm (namely, R_0 and $\boldsymbol{E}$) depend only on

the choice of the basis of operators for QPT (matrices $\{E_\alpha\}$ in Eq. (4.4)) and the initial states ρ^{in}, and therefore can be calculated just once in an auxiliary script and recorded into the data *mat*-files. Other matrices R, and therefore $\mathcal{L}$ and J, which depend on the density matrices ρ^{fin}, are calculated in the main Matlab script. The main Matlab script calculates the process matrix χ, using the n_{in}^N density matrices ρ^{fin}, each of the dimension $2^N \times 2^N$, as the input data. These n_{in}^N density matrices correspond to the output density matrices ρ^{fin} in Eq. (4.4) after the evolution of the known input basis states. These density matrices ρ^{fin} can either be experimentally measured (using QST methods) density matrices at the output of the quantum channel, or, as we do not have true experimental data for the systems of more than two qubits, can be randomly generated for an arbitrary number of qubits. According to the general QPT methodology, these density matrices correspond to measuring n_{in}^N states ρ^{fin} at the quantum channel's output, as the result of the evolution of the input states ρ^{in}, as in Eq. (4.4).

Before running the main Matlab script, an auxiliary script needs to be run first (just one time) in order to calculate and save into mat-files the inverse of matrix R_0 and the matrix $\boldsymbol{E}$, discussed above. These calculations of the inverse of the matrix R_0 and of the matrix $\boldsymbol{E}$ in a separate auxiliary script increase the speed of the main program by approximately a factor of two. In order to calculate the process matrix χ, the main Matlab script performs reshaping of the density matrices ρ^{fin} into columns, as described above, and builds the R matrix. Then the program reads from the *mat*-file the inverse matrix R_0^{-1}, which was calculated and recorded into the *mat*-file earlier by the auxiliary script. Now, when the main program has both matrices R and R_0^{-1}, it calculates the matrix $\mathcal{L}$ according to Eq. (4.9). In the next step, this matrix $\mathcal{L}$ is reshaped into the J-matrix according to the reshaping rule formulated above. After the calculation of J-matrix, the main program

reads already calculated $\boldsymbol{E}$ matrix from the *mat*-file, created beforehand by the auxiliary Matlab script, and calculates the process matrix χ by matrix multiplication of $\boldsymbol{E}^{\dagger}$, J and $\boldsymbol{E}$ according to the Eq. (4.10).

The author of this dissertation wrote the Matlab computer program, which performs the calculation of the process matrix χ by the linear inversion method. Example of the process matrix χ, calculated by the linear inversion method of QPT from the experimental data for the two-qubit CZ gate, realized with the phase qubits, is presented in Fig. 4.1.

Below follows the discussion of the computer resources, such as memory RAM and computational time, required for the linear inversion algorithm of QPT.

4.3.2 Computer memory and time requirements for the linear inversion method of QPT

The biggest numerical matrices that one needs to store in memory RAM of the computer are the complex matrices R and R_0, which enter into Eq. (4.9) for calculation of $\mathcal{L}$, and matrices $\mathcal{L}$, J, and χ that are calculated using R and R_0. All these matrices are complex matrices of the dimension $4^N \times 4^N$. It is possible to use two types of memory allocation for the elements of these arrays: the default "double-precision floating point numbers format" for each number in the arrays, which requires 64 bits, or equivalently 8 bytes, for each value stored, or we can change the format to "single-precision floating point numbers format", which requires 32 bits, or equivalently 4 bytes, for each value. The amount of computer memory, necessary to store the matrices R and R_0, is shown in Table 4.1 for various numbers of qubits in the system, N. As we see, the required memory scales as 16^N.

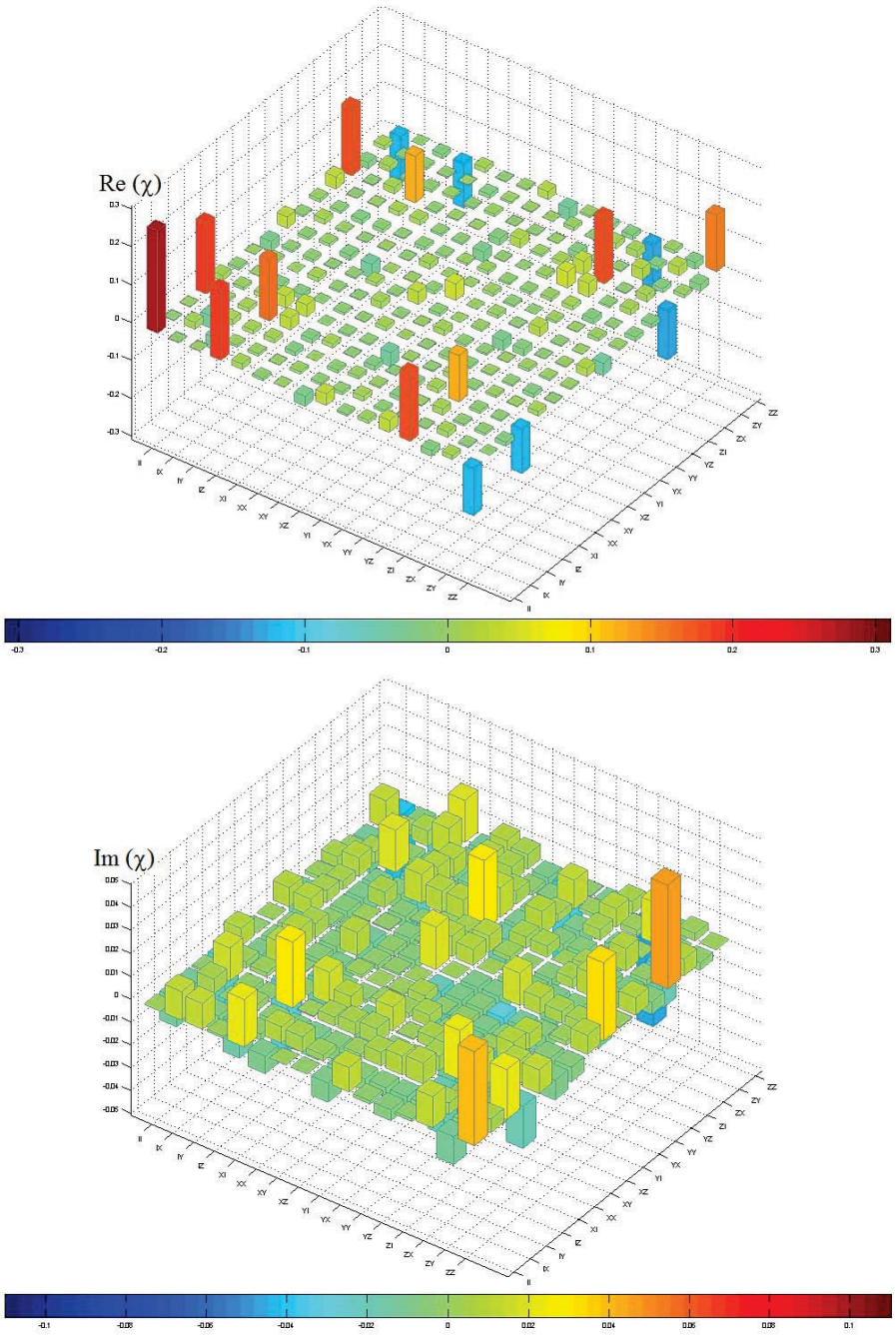

Figure 4.1: The process matrix χ [$\mathrm{Re}(\chi)$ on the top panel, $\mathrm{Im}(\chi)$ on the lower panel] calculated by the linear inversion method of QPT. The experimental data for the two-qubit CZ gate realized with the phase qubits have been used. The process fidelity with the ideal χ-matrix is $F \equiv \mathrm{Tr}(\chi\chi_{\mathrm{ideal}}) = 0.63$. The modified Pauli basis $\{E_\alpha\}$ has been used. Note that the scales on the vertical axes for $\mathrm{Re}(\chi)$ and $\mathrm{Im}(\chi)$ are different.

Strictly speaking, this consideration turns out to be not the upper memory limitation in our Matlab program. The program is written in the way that in order to calculate Jamiolkowski J-matrix, the *cell array* of dimension $4^N \times 4^N$ is introduced and calculated in the program, and later this cell array is converted into a numerical J-matrix. It is known that cell arrays require more memory than numerical arrays, because some information is recorded into *headers*, which require additional memory, but it turns out that such additional memory does not play a significant role. For a 64-bit Windows system, each cell array header consumes 112 bytes, and there are 4^N elements of the cell-array, therefore the additional memory is only $(4^N \times 112)$ bytes, which corresponds to only 448 kB for a 6-qubit system (a very small addition to 256 MB).

Table 4.1: The amount of computer RAM memory, required to store the arrays R and R_0 for various number of qubits in the system, N. Two types of memory allocation are compared.

N	Dimension of R and R_0	Memory "double-precision"	Memory "single-precision"
2	16×16	4 kB	2 kB
3	64×64	64 kB	32 kB
4	256×256	1 MB	512 kB
5	1024×1024	16 MB	8 MB
6	4096×4096	256 MB	128 MB
7	16384×16384	4GB	2 GB

Although it may seem that the dimensions of matrices and RAM memory needed in the $7-$qubit case are within the size of RAM-memory of an average computer (a computer with 6GB of RAM memory was used by the author of dissertation for these calculations), and it was indeed possible to calculate matrices R and R_0 for the $7-$ qubit case (it takes about 4 hours of calculation time on an average computer to calculate each matrix), but Matlab runs out of memory when it tries

to calculate the inverse of matrix R_0 in Eq. (4.9), either by using the Matlab command "inv", or by using the Matlab right-divide operator "mrdivide". Therefore six is the maximum number of qubits in a quantum system, for which the process matrix χ can be calculated by the linear inversion method of full QPT, described in this Section, on an average computer.

The computer time requirements are not demanding for this algorithm. It takes less than a second of computational time to run the QPT algorithm (both auxiliary and main scripts) for the systems of two and three qubits on an average PC, and the computational time increases with the increasing number of qubits. It takes about 1000 seconds to run the "auxiliary" script and about 260 seconds to run the "main" script for the case of 6 qubits on an average PC.

Let us provide the rough estimate of time resources it takes in an experiment to measure the necessary full set of experimental probabilities for different numbers of qubits N (from $N = 2$ to $N = 6$) in the system. We compare two protocols of measurements. First we consider the measurement protocol with the minimum number of basis rotations per qubit $n_R = 3$ and $n_{\text{in}} = 4$ different initial states per qubit in QST experiments. In this protocol an experimentalist needs to perform 12^N different experiments and measure 24^N different probabilities. If a bigger number of basis rotations per qubit $n_R = 6$ is used in the so-called "overconstrained tomography" QST experiment, keeping the same number of initial states $n_{\text{in}} = 4$ for every qubit, then the experimentalist needs to perform 24^N different experiments and measure 48^N different probabilities. If it takes about a millisecond ($t = 1\ ms$) for measuring a single experimental outcome, and if one needs to perform 1000 repetitions of experiment to achieve reasonable statistics for each probability (there were actually 1500 repetitions in the experiments with two phase qubits, performed at UCSB, we assume here 1000 repetitions for the simplification of our estimate), then we arrive at the follow-

ing crude estimate for the time, required to measure the full set of probabilities for QPT, which is presented in Table 4.2.

Table 4.2: Time required to measure the full set of probabilities in QPT, for different numbers of basis rotations per qubit, $n_R = 3$ and $n_R = 6$, and for various numbers of qubits in the system, N. The number of initial states for each qubit is $n_{\text{in}} = 4$.

N	$n_R = 3, \;\; n_{\text{in}} = 4$	$n_R = 6, \;\; n_{\text{in}} = 4$
2	2.4 minutes	9.6 minutes
3	29 minutes	3.8 hours
4	5.76 hours	3.8 days
5	2.9 days	3 months
6	34.6 days	6 years

Here is an example of how this rough estimate can be obtained for $N = 6$ qubits, if the number of basis rotations per qubit $n_R = 3$ and the number of different initial states per qubit is $n_{\text{in}} = 4$: 12^N different experiments $\times 1000$ repetitions$\times 0.001$sec $= 12^N$sec$= 829.4$ hours$= 34.6$ days. As we see, the "overconstrained" tomography with $n_R = 6$ becomes impractical for more than four qubits.

4.4 Maximum likelihood and least-squares methods for QPT

We have just discussed in Section 4.3 the method of linear inversion in QPT that allows one to invert Eq. (4.8) and thus find the process matrix χ from the experimental set of probabilities $\vec{P}^{\text{exp}}$. In this method, however, the conditions for the process matrix to be physically valid, such as the positive semidefinite condition Eq. (4.5) and the trace preservation condition Eq. (4.6), were

not imposed on the process matrix χ. In practice, because of experimental uncertainties present in $\vec{P}^{\text{exp}}$, the process matrix thus obtained is usually non-physical, that is, inconsistent with the conditions Eqs. (4.5) and (4.6). In standard QPT this problem is remedied by finding the physical process matrix [satisfying Eqs. (4.5) and (4.6)] that minimizes (in some sense) the difference between the probabilities $\vec{P}(\chi)$ and the experimental probabilities $\vec{P}^{\text{exp}}$.

Two popular methods used to estimate a physical process matrix χ compatible with the experimental data are the maximum likelihood (ML) method [133, 123, 134] (see also [51, 135]) and the least-squares (LS) method [122, 36, 136]. The ML method minimizes the cost function [133]

$$\mathcal{C}_{ML} = -\sum\nolimits_j P_j^{\text{exp}} \ln P_j(\chi), \tag{4.11}$$

where the index j labels the measured probabilities, while the LS method (often also called maximum likelihood) minimizes the difference between $\vec{P}(\chi)$ and $\vec{P}^{\text{exp}}$ in the ℓ_2-norm sense [**?**], so the minimized cost function is

$$\mathcal{C}_{LS} = ||\vec{P}(\chi) - \vec{P}^{\text{exp}}||_{\ell_2}^2 = \sum\nolimits_j [P_j^{\text{exp}} - P_j(\chi)]^2. \tag{4.12}$$

In both methods the conditions Eq. (4.5) and Eq. (4.6) should be satisfied to ensure that χ corresponds to a physical process. This can be done in a number of ways, for example using the Cholesky decomposition, the Lagrange multipliers, or just stating the conditions Eqs. (4.5) and (4.6) as a constraint (if an appropriate software package is used). The ML method Eq. (4.11) is natural when the inaccuracy of $\vec{P}^{\text{exp}}$ is dominated by the statistical error due to a limited number of experimental runs. However, this method does not work well if a target probability P_j is near zero, but P_j^{exp} is not near zero due to experimental imperfections (e.g., "dark counts"); this is because the cost function

Eq. (4.11) is very sensitive to changes in $P_j^{\rm exp}$ when $P_j(\chi) \approx 0$. Therefore the LS method Eq. (4.12) is a better choice when the inaccuracy of $\vec{P}^{\rm exp}$ is not due to a limited number of experimental runs.

Note that other cost functions can also be used for minimization in the procedure. For example, by replacing $\ln P_j(\chi)$ in Eq. (4.11) with $\ln[P_j(\chi)/P_j^{\rm exp}]$ (this obviously does not affect optimization), then expanding the logarithm to second order, and using condition $\sum_j P_j(\chi) = \sum_j P_j^{\rm exp}$ (which cancels the first-order term), we obtain [51] $\mathcal{C}_{ML} \approx {\rm const} + \sum_j [P_j^{\rm exp} - P_j(\chi)]^2/2P_j^{\rm exp}$. This leads to another natural cost function

$$\mathcal{C} = \sum_j \frac{[P_j(\chi) - P_j^{\rm exp}]^2}{P_j^{\rm exp} + a}, \tag{4.13}$$

where we phenomenologically introduced an additional parameter a, so that for $a \gg 1$ the minimization reduces to the LS method, while for $a \ll 1$ it is close to the ML method (the parameter a characterizes the relative importance of non-statistical and statistical errors). One more natural cost function is similar to Eq. (4.13), but with $P_j^{\rm exp}$ in the denominator replaced by $P_j^{\rm exp}(1 - P_j^{\rm exp})$ (see [133]), which corresponds to the binomial distribution variance.

Both problems Eq. (4.11) and Eq. (4.12) are convex optimization problems [29, 39, 137], therefore it is convenient to use the special Matlab software for solving these convex optimization problems, such as YALMIP [138] or CVX [139]. Unfortunately, the ability to solve the log-convex problems described by Eq. (4.11) has not been implemented in any of these Matlab packages YALMIP or CVX by its developers so far, but efficient numerical algorithms capable of solving the least-squares convex optimization problems of the type of Eq. (4.12) are implemented in both these packages. Therefore, it is much easier to solve the least-squares convex optimization problem of the type of Eq. (4.12) from the computational point of view. In this dissertation we use the LS method Eq. (4.12) for the standard QPT. In particular, we find the process matrix $\chi_{\rm full}$ for the full

data set $\vec{P}^{\text{exp}}_{\text{full}}$ by minimizing $||\vec{P}(\chi_{\text{full}}) - \vec{P}^{\text{exp}}_{\text{full}}||_{\ell_2}$, subject to the conditions Eqs. (4.5) and (4.6). Such minimization problem is computationally tractable as it is a convex optimization problem, and the author of this dissertation has written the Matlab computer program that calculates the process matrix χ by the least-squares method. As an example, two process matrices χ_{full} reconstructed by the LS method from the two sets of experimental data for two-qubit CZ gates are shown in Figs. 4.2 and 4.3. The process matrix presented in Fig. 4.2 is calculated from the experimental data for the CZ gate based on phase qubits, while the experimental data obtained with Xmon qubit are used in the calculation of the the process matrix presented in Fig. 4.3.

In the next Chapter 5 we benchmark the results of the compressed sensing QPT process matrix reconstruction from partial data against the process matrix χ_{full}, obtained by the LS method using the full data set.

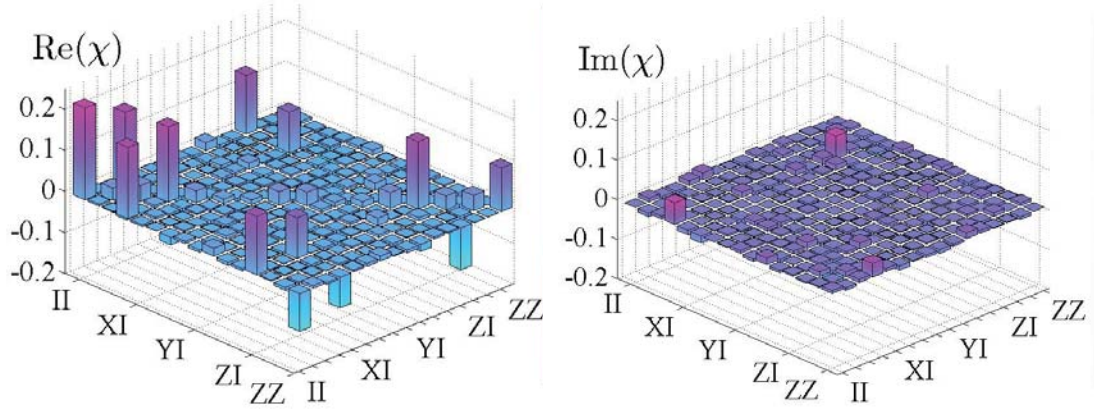

Figure 4.2: The process matrix χ calculated by the least squares method of QPT. The experimental data for the two-qubit CZ gate realized with the phase qubits have been used. The process fidelity with the ideal χ-matrix is $F \equiv \mathrm{Tr}(\chi\,\chi_{\mathrm{ideal}}) = 0.51$. The modified Pauli basis $\{E_\alpha\}$ has been used.

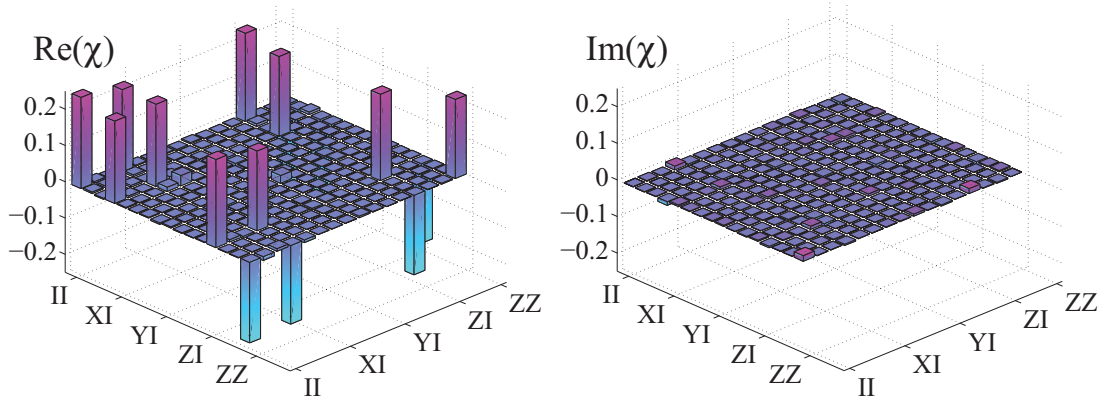

Figure 4.3: The process matrix χ calculated by the least squares method of QPT. The experimental data for the two-qubit CZ gate realized with the Xmon qubits have been used. The process fidelity with the ideal χ-matrix is $F \equiv \mathrm{Tr}(\chi\,\chi_{\mathrm{ideal}}) = 0.91$. The modified Pauli basis $\{E_\alpha\}$ has been used.

Chapter 5

Compressed Sensing Quantum Process Tomography of two and three qubit gates

5.1 Introduction to Compressed Sensing Quantum Process Tomography

As we have seen in Chapter 4, the standard QPT suffers from a fundamental drawback in that the number of required experimental configurations and therefore the number of measured probabilities scale exponentially with the number of qubits, this makes the standard QPT impractical for systems of several qubits. As we have discussed in Chapter 4, for an N-qubit system, the number of measurement configurations in standard QPT is $M_{\mathrm{conf}} = n_{\mathrm{in}}^N n_{\mathrm{R}}^N$, while the total number of probabilities in the data set is $M = M_{\mathrm{conf}} N_{\mathrm{prob}}$ (where n_{in} is the number of input single-qubit

states and n_{R} is the number of basis rotations per qubit, and the number of measured probabilities per configuration N_{prob} is usually 2^N). These numbers M_{conf} and M become very large even for few-qubits systems, implying the amount of information one needs to collect during the experiment and the computer resources required for post-processing the data are becoming unmanageable. The method of compressed sensing QPT (CS QPT) proposed in [39, 40] is hoped to alleviate this problem. Compressed sensing techniques are based on the idea that the ideal operation can be described by a maximally sparse process matrix in a special basis which includes the intended unitary process. Therefore, a nearly precise implementation of the intended process will be described by an approximately sparse matrix in this basis. The concepts of compressed sensing techniques as used in signals processing [41, 42, 43, 44] are then applied to efficiently characterize the implemented process. As it has been shown in [40], if the process matrix χ is known to be s-compressible[1] in some known basis, then for a d-dimensional quantum system the CS QPT method is supposed to require only $O(s \log d)$ experimental probabilities to produce a good estimate of the process matrix χ. The CS QPT method was experimentally validated in Ref. [40] for a photonic two-qubit controlled-Z (CZ) gate.

The CS idea also inspired another (quite different) algorithm for quantum state tomography (QST) [140, 141], which can be generalized to QPT [141, 142]. This matrix-completion method of CS QST estimates the density matrices of nearly pure (low rank r) d-dimensional quantum states from expectation values of only $O(rd \operatorname{poly} \log d)$ observables, instead of d^2 observables required for standard QST. It is important to mention that this method does not require any assumption about the quantum state of a system, except that it must be a low-rank state (in particular, we do not need to know the state approximately). The CS QST method has been used to reconstruct the quantum

[1] Roughly speaking, a matrix is s-compressible if it can be well approximated by an s-sparse matrix.

states of a 4-qubit photonic system [143] and cesium atomic spins [144]. In Ref. [141] it has been shown that using the Jamiołkowski process-state isomorphism [145] the formalism of CS QST can also be applied to the QPT, requiring $O(rd^2 \operatorname{poly} \log d)$ measured probabilities (where r is the rank of the Jamiołkowski state) to produce a good estimate of the process matrix χ. Therefore there is crudely a square-root speedup compared with standard QPT. Note that this algorithm requires exponentially more resources than the CS QPT method of Ref. [40], but it does not require us to know a particular basis in which the matrix χ is sparse. The performance of these two methods has been compared in the recent paper [142] for a simulated quantum system with dimension $d = 5$; the reported result is that the method of Ref. [141] works better for coherent errors, while the method of Ref. [40] is better for incoherent errors.

The method of compressed sensing is now an already well-developed mathematical field with numerous applications in signal processing, including medical magnetic resonance tomography [146], photography [147], face recognition [148], holography [149], seismic imaging [150], etc. Compressed sensing typically involves forming a convex optimization involving the ℓ_2-norm of the measurement error and the ℓ_1-norm (we define norms in the subsequent Section 5.2) of the estimation variable, the latter being a convex heuristic for sparsity. The estimate is obtained by solving a convex optimization problem where under suitable measurement conditions the underdetermined measurements plus the sparsity heuristic yields a very accurate solution. Perfect recovery of a sparse signal is achieved with no measurement noise, and the errors grow gracefully with noise and near-sparsity of the signal.

Compressed sensing quantum process tomography (CS QPT) enables one to recover the process matrix χ from far fewer experimental configurations than the standard QPT. In this Chapter

we apply the method of CS QPT, introduced in Ref. [40] and experimentally validated there for a photonic two-qubit controlled-Z (CZ) gate, to several two-qubit CZ gates realized with superconducting Xmon and phase qubits, as well as to the simulated data for the three-qubit Toffoli gate with numerically added noise.

This Chapter is structured as follows. We first formulate the problem of CS QPT mathematically in Section 5.2, then discuss the set of measurement configurations used to collect QPT data for the systems with Xmon and phase qubits, and also briefly discuss our way of computing the process matrix χ via compressed sensing in Section 5.3. In the following sections we present our numerical results for the CS QPT of a superconducting two-qubit CZ gate. We explain the importance of choosing the proper value for the so-called noise parameter ε in Section 5.4, then we present the results of the reconstruction of the process matrix χ by the CS QPT method from the reduced (partial) data set in Sections 5.5 and 5.6 for various values of the noise parameter ε. We also compare numerical results obtained by applying the CS QPT method in two different operator bases, where the process matrix is expected to be almost sparse (the Pauli-error basis and the SVD basis), in Section 5.7. We also compare the performance of the CS QPT method with the least optimization, using partial data, in Section 5.8. Finally, we present the results of our study of the CS QPT of a simulated three-qubit Toffoli gate with the numerically added noise in Section 5.9.

5.2 Fundamentals of Compressed Sensing Quantum Process Tomography

If the number of available experimental probabilities in the quantum process tomography problem is smaller than the number of independent parameters in the process matrix (i.e. $m <$

$d^4 - d^2$), then the set of linear equations Eq. (4.8) from Chapter 4 for the process matrix χ becomes underdetermined. Actually, the LS method may still formally work in this case for some range of m, but, as discussed in Sections 5.8 and 5.9, it is less effective. As a natural alternative, the methodology of compressed sensing is applicable to an underdetermined set of equations where the unknown signal is known to be sparse with an unknown sparsity pattern. By using the ideas of compressed sensing [41, 43, 42, 44], the method of CS QPT requires a significantly smaller set of experimental data to produce a reasonably accurate estimate of the process matrix.

Let us formulate the problem mathematically as follows: we wish to find the physical process matrix $\vec{\chi}_0$ satisfying the equation

$$\vec{P}^{\text{exp}} = \Phi\vec{\chi}_0 + \vec{z}, \tag{5.1}$$

where the vector $\vec{P}^{\text{exp}} \in \mathbb{C}^m$ (with $m < d^4 - d^2$) and the matrix $\Phi \in \mathbb{C}^{m\times d^4}$ are given, while $\vec{z} \in \mathbb{C}^m$ is an unknown noise vector, whose elements are assumed to be bounded (in the root-mean-square sense) by a known limit ε, $||\vec{z}||_{\ell_2}/\sqrt{m} \leq \varepsilon$. While this problem seems to be ill-posed since the available information is both noisy and incomplete, in Ref. [41] it was shown that if the vector χ_0 is sufficiently sparse and the matrix Φ satisfies the restricted isometry property (RIP), χ_0 can be accurately estimated from Eq. (5.1). Note that the CS techniques of Ref. [41] were developed in the context of signal processing; to adapt [39] these techniques to QPT we also need to include the positivity and trace-preservation conditions, Eqs. (4.5) and (4.6) from Section 4.1 of Chapter 4.

The idea of CS QPT [40] is to minimize the ℓ_1-norm of $\vec{\chi}$ in a basis where χ is assumed to be approximately sparse. The definitions of the ℓ_1 and ℓ_2 norms are the following: for a vector $x \in C^m$, the ℓ_2 norms is defined as $||x||_{\ell_2} = \sqrt{x^\dagger x} = \sqrt{\sum_{i=1}^m |x_i|^2}$, and the ℓ_1 norm as $||x||_{\ell_1} = \sum_{i=1}^m |x_i|$. Mathematically, the method of compressed sensing is solving the following convex

optimization problem:

$$\text{minimize } ||\vec{\chi}||_{\ell_1}\,, \tag{5.2}$$

$$\text{subject to } ||\vec{P}(\chi) - \vec{P}^{\exp}||_{\ell_2}/\sqrt{m} \leq \varepsilon \tag{5.3}$$

$$\text{and conditions (4.5), (4.6).}$$

As shown in Refs. [42, 40], a faithful reconstruction recovery of an approximately s-sparse process matrix χ_0 via this optimization is guaranteed (see below) if (i) the matrix Φ satisfies the RIP condition,

$$1 - \delta_s \leq \frac{||\Phi\vec{\chi}_1 - \Phi\vec{\chi}_2||^2_{\ell_2}}{||\vec{\chi}_1 - \vec{\chi}_2||^2_{\ell_2}} \leq 1 + \delta_s, \tag{5.4}$$

for all s-sparse vectors (process matrices) $\vec{\chi}_1$ and $\vec{\chi}_2$, (ii) the isometry constant δ_s is sufficiently small, $\delta_s < \sqrt{2} - 1$, and (iii) the number of data points is sufficiently large,

$$m \geq C_0 s \log(d^4/s) = O(sN), \tag{5.5}$$

where C_0 is a constant. Quantitatively, if $\chi_{\rm CS}$ is the solution of the optimization problem [Eqs. (5.2) and (5.3)], then the estimation error $||\chi_{\rm CS} - \chi_0||_{\ell_2}$ is bounded as

$$\frac{||\chi_{\rm CS} - \chi_0||_{\ell_2}}{\sqrt{m}} \leq \frac{C_1||\chi_0(s) - \chi_0||_{\ell_1}}{\sqrt{ms}} + C_2\,\varepsilon, \tag{5.6}$$

where $\chi_0(s)$ is the best s-sparse approximation of χ_0, while C_1 and C_2 are constants of the order $O(\delta_s)$. Note that in the noiseless case ($\varepsilon = 0$) the recovery is exact if the process matrix χ_0 is s-sparse. Also note that while the required number of data points m and the recovery accuracy depend on the sparsity s, the method itself [Eqs. (5.2) and (5.3)] does not depend on s, and therefore s need not be known.

The inequality Eq. (5.5) and the first term in the inequality Eq. (5.6) indicate that the CS QPT method is supposed to work well only if the actual process matrix χ_0 is sufficiently sparse.

Therefore it is important to use an operator basis $\{E_\alpha\}$ [see Eq. (4.4) in Section 4.1 of Chapter 4], in which the ideal (desired) process matrix χ_{ideal} is maximally sparse, i.e., it contains only one nonzero element. Then it is plausible to expect the actual process matrix χ_0 to be approximately sparse [40]. In this paper we will use two bases in which the ideal process matrix is maximally sparse. These are the so-called Pauli-error basis [18] and the SVD basis of the ideal unitary operation [39]. In the Pauli-error basis $\{E_\alpha\}$, the first element E_1 coincides with the desired unitary U, while other elements are related via the N-qubit Pauli matrices $\mathcal{P}$, so that $E_\alpha = U\mathcal{P}_\alpha$. The N-qubit Pauli matrices $\mathcal{P}$ are calculated as the Kronecker product (also called tensor product or direct product) of the single-qubit Pauli matrices σ_0, σ_x, σ_y, σ_z, introduced in Section 2.1.3 of Chapter 2. In the SVD basis $E_1 = U/\sqrt{d}$, and other elements are obtained via a numerical SVD procedure. More details about the Pauli-error and SVD bases are discussed in Appendices B and C.

As mentioned previously, the method of CS QPT involves the RIP condition Eq. (5.4) for the transformation matrix Φ. In Ref. [40] it was shown that if the transformation matrix Φ in Eq. (4.8) is constructed from randomly selected input states ρ_k^{in} and random measurements Π_i, then Φ obeys the RIP condition with high probability. Notice that once a basis $\{E_\alpha\}$ and a tomographically complete (or overcomplete) set $\{\rho_k^{\text{in}}, \Pi_i\}$ have been chosen, the matrix Φ_{full} corresponding to the full data set is fully defined, since it does not depend on the experimental outcomes. Therefore the mentioned above result of Ref. [40] tells us that if we build a matrix Φ_m by randomly selecting m rows from Φ_{full}, then Φ_m is very likely to satisfy the RIP condition. Hence the submatrix $\Phi_m \in \mathbb{C}^{m\times d^4}$, together with the corresponding set of experimental outcomes $\vec{P}^{\text{exp}} \in \mathbb{C}^m$ can be used to produce an estimate of the process matrix via the ℓ_1-minimization procedure Eq. (5.2) and (5.3).

5.3 CS QPT for two-qubit controlled-Z gate

In the following sections, we present results for the experimental CZ gate realized with superconducting Xmon and phase qubits, which were introduced in the sections 2.2.4 and 2.2.5 of Chapter 2. Before proceeding to presenting our results, we first introduce in this section the set of measurement configurations used to collect QPT data and the experimental device used, and also briefly explain our way of computing the χ-matrix via compressed sensing.

As was explained in Section 4.2 in Chapter 4 of this dissertation, for an N-qubit system the number of measurement configurations in standard QPT is $M_{\rm conf} = n_{\rm in}^N n_{\rm R}^N$, while the total number of probabilities in the data set is $M = M_{\rm conf} N_{\rm prob}$. We denote the number of input single-qubit states as $n_{\rm in}$, the number of basis rotations per qubit as $n_{\rm R}$, and the number of measured probabilities for each configuration as $N_{\rm prob}$. In this dissertation we focus on the case $n_{\rm in} = 4$, $n_{\rm R} = 3$, and $N_{\rm prob} = 2^N$. Then for a two-qubit quantum gate there are $M_{\rm conf} = 12^N = 144$ measurement configurations and $M = 24^N = 576$ probabilities (432 of them independent). For a three-qubit gate there are $M_{\rm conf} = 1728$ configurations and $M = 13824$ probabilities (12096 of them independent).

The main experimental data used in this dissertation are for the two-qubit CZ gate realized with Xmon qubits [98]. The data were obtained with $n_{\rm in} = 6$, $n_{\rm R} = 6$, and $N_{\rm prob} = 2^N$. However, since the main emphasis of this dissertation is the analysis of QPT with a reduced data set, we started by reducing the data set to $n_{\rm in} = 4$ and $n_{\rm R} = 3$ by using only the corresponding probabilities and removing other data. We will refer to these data as "full data" (with $M_{\rm conf} = 144$ and $M = 24^N = 576$). For testing the CS method we randomly choose $m_{\rm conf} \leq M_{\rm conf}$ configurations, with corresponding $m = 4m_{\rm conf}$ experimental probabilities ($3m_{\rm conf}$ of them independent). Since the

process matrix χ is characterized by $16^N - 4^N = 240$ independent parameters, the power of the CS method is most evident when $m_{\rm conf} < 80$, so that the system of equations (4.8) is underdetermined. [For a three-qubit gate the system of equations becomes underdetermined for $m_{\rm conf} < (16^N - 4^N)/(2^N - 1) = 576$.]

The main experimental data used for the analysis in this dissertation were taken on an Xmon device, similar to the one used in Ref. [11], but there were some differences in its design[2]. For the device used here the qubits were coupled via a bus, and the entangling gate between qubits A and B was implemented with three multiqubit operations: 1) swap state from qubit B to bus, 2) CZ gate between qubit A and bus, 3) swap back to qubit B. The swap was done with the resonant Strauch gate [151], by detuning the frequency of qubit A with a square pulse. Generating a square pulse is experimentally challenging, moreover this gate has a single optimum in pulse amplitude and time. We also note that the qubit frequency control was not optimized for imperfections in the control wiring, as described in Ref. [152] (see also Fig. S4 in Supplementary Information of [11]). The combination of device, non-optimal control, and multiple operations, leads to the experimental process fidelity $F_\chi = 0.91$ of the CZ gate used for the analysis here to be significantly less than the randomized benchmarking fidelity $F_{RB} = 0.994$ reported in [11]. Moreover, QPT necessarily includes state preparation and measurement (SPAM) errors [33], while randomized benchmarking does not suffer from these errors. This is why we intentionally used the data for a not-well-optimized CZ gate so that the gate error dominates over the SPAM errors. (Note that we use correction for the imperfect measurement fidelity [113]; however, it does not remove the measurement errors

[2]The data were collected during an experiment performed by postdoc Rami Barends and graduate student Julian Kelly at the University of California, Santa Barbara. The author of this dissertation used in his theoretical work the experimental data provided by Rami Barends.

completely.) It should also be mentioned that in the ideal case $1 - F_\chi = (1 - F_{\rm RB}) \times (1 + 2^{-N})$, so the QPT fidelity is supposed to be slightly less than the randomized benchmarking fidelity.

Besides performing the CS QPT analysis of the experimental data, obtained with an Xmon device, we also used in our calculations two other sets of experimental data, obtained with phase qubits. In fact, experiments with the phase qubits at UCSB were done much earlier than the experiments with Xmon qubits, and there were several technical problems in initial attempts to implement the CZ gate using phase qubits, such as low anharmonicity and low T_1 decoherence time, which resulted in a low quality of both CZ gates implemented with phase qubits. In the phase-qubit experiments, the data were obtained with $n_{\rm in} = 4$, $n_{\rm R} = 6$, and $N_{\rm prob} = 2^N$. In analogy to the Xmon data, we reduced both data sets, obtained with phase qubits, to $n_{\rm in} = 4$ and $n_{\rm R} = 3$ by using only the corresponding probabilities and removing other data. These "full data" have $M_{\rm conf} = 144$ and $M = 24^N = 576$, the same $M_{\rm conf}$ and M as the "full data" for Xmon qubits.

Now we briefly explain the overall procedure of our calculations using CS QPT method, the details of calculations and our results will be presented in the following sections of this Chapter.

We first calculate the process matrix $\chi_{\rm full}$ for the full data set by using the least-squares method described in Eq. (4.12) in Section 4.4 of Chapter 4. For that we use three different operator bases $\{E_\alpha\}$: the Pauli basis, the Pauli-error basis, and the SVD basis. The pre-computed transformation matrix Φ in Eq. (4.8) depends on the choice of the basis, thus giving a basis-dependent result for $\chi_{\rm full}$. We then check that the results essentially coincide by converting $\chi_{\rm full}$ between the bases and calculating the fidelity between the corresponding matrices (the infidelity is found to be less than 10^{-7}).

As a natural extension of the fidelity between quantum states to quantum channels, introduced in Eq. (2.24) in Chapter 2, the fidelity between two process matrices χ_1 and χ_2 is defined as the square of the Uhlmann fidelity [153, 69],

$$F(\chi_1, \chi_2) = \left(\mathrm{Tr} \sqrt{\chi_1^{1/2} \chi_2 \chi_1^{1/2}} \right)^2, \tag{5.7}$$

so that it reduces to $F(\chi_1, \chi_2) = \mathrm{Tr}(\chi_1 \chi_2)$ [154] when at least one of the process matrices corresponds to a unitary operation. Since $0 \leq F \leq 1$, we refer to $1 - F$ as the infidelity.

The LS method using the full data set produces the process matrix χ_{full}, which has the process fidelity $F(\chi_{\mathrm{full}}, \chi_{\mathrm{ideal}}) = 0.907$ relative to the ideal CZ operation in the experiments with Xmon qubits. Note that our full data set is actually a subset of an even larger data set (as explained above), and the χ matrix calculated from the initial set corresponds to the process fidelity of 0.928 for the Xmon qubits; the difference gives a crude estimate of the overall accuracy of the procedure. The values of the same process fidelities for the two sets of data obtained with phase qubits were significantly lower, only $F(\chi_{\mathrm{full}}, \chi_{\mathrm{ideal}}) = 0.62$ and $F(\chi_{\mathrm{full}}, \chi_{\mathrm{ideal}}) = 0.51$.

After calculating χ_{full} for the full data set, we can calculate its fidelity compared to the process matrix χ_{ideal} of the desired ideal unitary operation, $F_\chi = F_{\mathrm{full}} = F(\chi_{\mathrm{full}}, \chi_{\mathrm{ideal}})$. This is the main number used to characterize the quality of the quantum operation.

Then we calculate the compressed-sensing process matrix χ_{CS} by solving the ℓ_1 - minimization problem described by Eqs. (5.2) and (5.3), using the reduced data set. It is obtained from the full data set by randomly selecting m_{conf} configurations out of the full number M_{conf} configurations. We use the fidelity $F(\chi_{\mathrm{CS}}, \chi_{\mathrm{full}})$ to quantify how well the process matrix χ_{CS} approximates the matrix χ_{full} obtained from full tomographic data. Additionally, we calculate the process fidelity

$F(\chi_{\rm CS}, \chi_{\rm ideal})$ between $\chi_{\rm CS}$ and the ideal operation, to see how closely it estimates the process fidelity $F_{\rm full}$, obtained using the full data set.

Since both the least-squares and the ℓ_1-norm minimization are convex optimization problems [39, 137], they can be efficiently solved numerically. We used two ways for MATLAB-based numerical calculations: (1) using the package CVX [139], which calls the convex solver SeDuMi [155]; or (2) using the package YALMIP [138], which calls the convex solver SDPT3 [156]. In general, we have found that for our particular realization of computation, CVX with the solver SeDuMi works better than the combination YALMIP-SDPT3 (more details are below in the subsequent Chapters).

The CS method calculations were mainly done in the Pauli-error basis, using the CVX-SeDuMi combination for the ℓ_1-norm minimization. This is what is implicitly assumed in this and the next Chapters, unless specified otherwise. Note that the CS-method optimization is very different from the LS method. Therefore, even for the full data set we would expect the process matrix $\chi_{\rm CS}$ to be different from $\chi_{\rm full}$. Moreover, $\chi_{\rm CS}$ depends on the noise parameter ε [see Eq. (5.3)], which to some extent is arbitrary. To clarify the role of the parameter ε, we will first discuss the CS method applied to the full data set, with varying ε, and then discuss the CS QPT for a reduced data set, using either near-optimal or non-optimal values of ε.

5.4 Full data set, varying noise parameter ε

We start with calculating the process matrix χ_{CS} by solving the ℓ_1-minimization problem, Eqs. (5.2) and (5.3), using the full data set and varying the noise parameter ε. The resulting matrix is compared with the LS result $\chi_{\rm full}$ and with the ideal matrix $\chi_{\rm ideal}$. Figure 5.1 shows the

corresponding fidelities $F(\chi_{\rm CS}, \chi_{\rm full})$ and $F(\chi_{\rm CS}, \chi_{\rm ideal})$ as functions of ε. We see that $\chi_{\rm CS}$ coincides with $\chi_{\rm full}$ [so that $F(\chi_{\rm CS}, \chi_{\rm full}) = 1$] at the optimal value $\varepsilon_{\rm opt} = 0.0199$. This is exactly the noise level corresponding to the LS procedure, $||\vec{P}^{\rm exp}_{\rm full} - \Phi\vec{\chi}_{\rm full}||_{\ell_2}/\sqrt{M} = 0.0199$. With ε increasing above this level, the relative fidelity between $\chi_{\rm CS}$ and $\chi_{\rm full}$ decreases, but it remains above 0.95 for $\varepsilon < 0.028$. Correspondingly, the process fidelity reported by $\chi_{\rm CS}$, i.e. $F(\chi_{\rm CS}, \chi_{\rm ideal})$, also changes. It starts with $F(\chi_{\rm CS}, \chi_{\rm ideal}) = F(\chi_{\rm full}, \chi_{\rm ideal}) = 0.907$ for $\varepsilon = 0.0199$, then increases with increasing ε, then remains flat above $\varepsilon = 0.025$, and then decreases at $\varepsilon > 0.032$. We note that for another set of experimental data (for a CZ gate realized with phase qubits) there was no increasing part of this curve, and the dependence of $F(\chi_{\rm CS}, \chi_{\rm ideal})$ on ε remained practically flat for a wide range of ε, up to $4\varepsilon_{\rm opt}$, see Fig. 5.2. The process fidelity for this CZ gate was very low, $F(\chi_{\rm full}, \chi_{\rm ideal}) = 0.51$. One more set of experimental data for phase qubits with a little better process fidelity of $F(\chi_{\rm full}, \chi_{\rm ideal}) = 0.62$ again had the increasing part of the curve $F(\chi_{\rm CS}, \chi_{\rm ideal})$, see Fig. 5.3.

To check how close the result of ℓ_1-optimization Eq. (5.2) is to the upper bound of the condition (5.3), we calculate the numerical value $\varepsilon_{\rm num} = ||\vec{P}^{\rm exp}_{\rm full} - \Phi\vec{\chi}_{\rm CS}||_{\ell_2}/\sqrt{M}$ as a function of ε. The results for all three CZ gates are shown in the insets of Figs. 5.1 – 5.3, we see that $\varepsilon_{\rm num}$ is quite close to ε for all three CZ gates. The CVX-SeDuMi package does not solve the optimization problem for values of the noise parameter ε below the optimal value $\varepsilon_{\rm opt}$.

Finding a proper value of ε to be used in the CS method is not a trivial problem, since for the reduced data set we cannot find $\varepsilon_{\rm opt}$ in the way we used. Therefore the value of ε should be estimated either from some prior information about the noise level in the system or by trying to solve the ℓ_1-minimization problem with varying value of ε. Note that the noise level $||\vec{P}^{\rm exp} -$

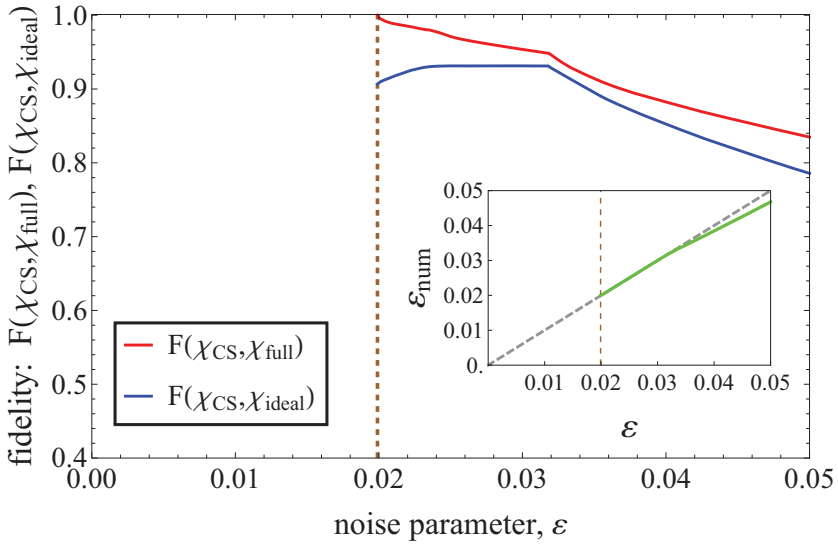

Figure 5.1: The CS QPT procedure, applied to the full data set for the superconducting Xmon qubit, with varying noise parameter ε. The red (upper) line shows the fidelity $F(\chi_{\rm CS}, \chi_{\rm full})$ between the process matrix $\chi_{\rm CS}$ obtained using the compressed-sensing method and the matrix $\chi_{\rm full}$ obtained using the least-squares method. The blue (lower) line shows the process fidelity $F(\chi_{\rm CS}, \chi_{\rm ideal})$, i.e., compared with the matrix $\chi_{\rm ideal}$ of the ideal unitary process. The vertical dashed brown line corresponds to the noise level $\varepsilon_{\rm opt} = ||\vec{P}^{\rm exp}_{\rm full} - \Phi\vec{\chi}_{\rm full}||_{\ell_2}/\sqrt{M} = 0.0199$ obtained in the LS procedure. The inset shows $\varepsilon_{\rm num} = ||\vec{P}^{\rm exp}_{\rm full} - \Phi\vec{\chi}_{\rm CS}||_{\ell_2}/\sqrt{M}$ as a function of ε (green line); for comparison, the dashed line shows the expected straight line, $\varepsilon_{\rm num} = \varepsilon$. The process fidelity $F(\chi_{\rm full}, \chi_{\rm ideal}) = 0.91$. The numerical calculations have been carried out in the Pauli-error basis using CVX-SeDuMi package.

$\Phi\vec{\chi}_{\rm ideal}||_{\ell_2}/\sqrt{M}$ defined by the ideal process is not a good estimate of $\varepsilon_{\rm opt}$; in particular for our full data corresponding to the CZ gate implemented with the Xmon qubit, it is 0.035, which is significantly higher than $\varepsilon_{\rm opt} = 0.0199$.

5.5 Reduced data set, near-optimal noise parameter ε

Now we apply the CS method to a reduced data set, by randomly choosing $m_{\rm conf}$ out of $M_{\rm conf} = 144$ configurations, while using all 4 probabilities for each configuration. (Therefore the number of used probabilities is $m = 4m_{\rm conf}$ instead of $M = 4M_{\rm conf}$ in the full data set.) For the

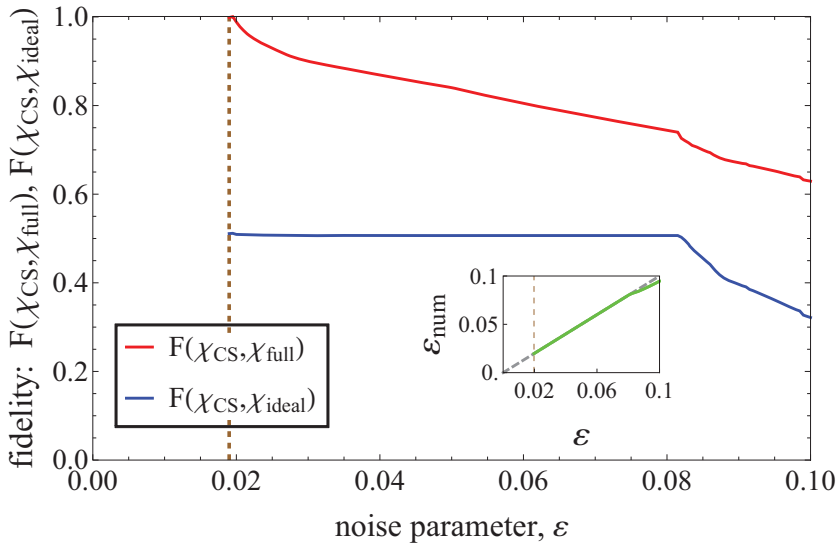

Figure 5.2: Similar to Fig. 5.1, but for the CZ gate, realized with superconducting phase qubits. The process fidelity $F(\chi_{\rm full}, \chi_{\rm ideal}) = 0.51$ is much lower than that for the Xmon qubit gate. The CS QPT procedure, applied to the full data set for the superconducting phase qubit, with varying noise parameter ε. The red (upper) line shows the fidelity $F(\chi_{\rm CS}, \chi_{\rm full})$ between the process matrix $\chi_{\rm CS}$ obtained using the compressed-sensing method and the matrix $\chi_{\rm full}$ obtained using the least-squares method. The blue (lower) line shows the process fidelity $F(\chi_{\rm CS}, \chi_{\rm ideal})$, i.e., compared with the matrix $\chi_{\rm ideal}$ of the ideal unitary process. The vertical dashed brown line corresponds to the noise level $\varepsilon_{\rm opt} = ||\vec{P}^{\rm exp}_{\rm full} - \Phi\vec{\chi}_{\rm full}||_{\ell_2}/\sqrt{M} = 0.0197$ obtained in the LS procedure. The inset shows $\varepsilon_{\rm num} = ||\vec{P}^{\rm exp}_{\rm full} - \Phi\vec{\chi}_{\rm CS}||_{\ell_2}/\sqrt{M}$ as a function of ε (green line); for comparison, the dashed line shows the expected straight line, $\varepsilon_{\rm num} = \varepsilon$. The numerical calculations have been carried out in the Pauli-error basis using CVX-SeDuMi package.

noise level ε we use a value slightly larger than $\varepsilon_{\rm opt}$ [40]. If a value too close to $\varepsilon_{\rm opt}$ is used, then the optimization procedure often does not find a solution; this happens when we choose configurations with a relatively large level of noise in the measured probability values. For the figures presented in this Section, which are obtained from the experimental data for the CZ gate realized with the Xmon qubits, we used $\varepsilon = 0.02015$, which for the full data set corresponds to the fidelity of 0.995 compared with $\chi_{\rm full}$ and to the process fidelity of 0.910 (see Fig. 5.1).

Figure 5.4 shows the fidelities $F(\chi_{\rm CS}, \chi_{\rm full})$ (upper line) and $F(\chi_{\rm CS}, \chi_{\rm ideal})$ (lower line) versus the number $m_{\rm conf}$ of used configurations. For each value of $m_{\rm conf}$ we repeat the procedure

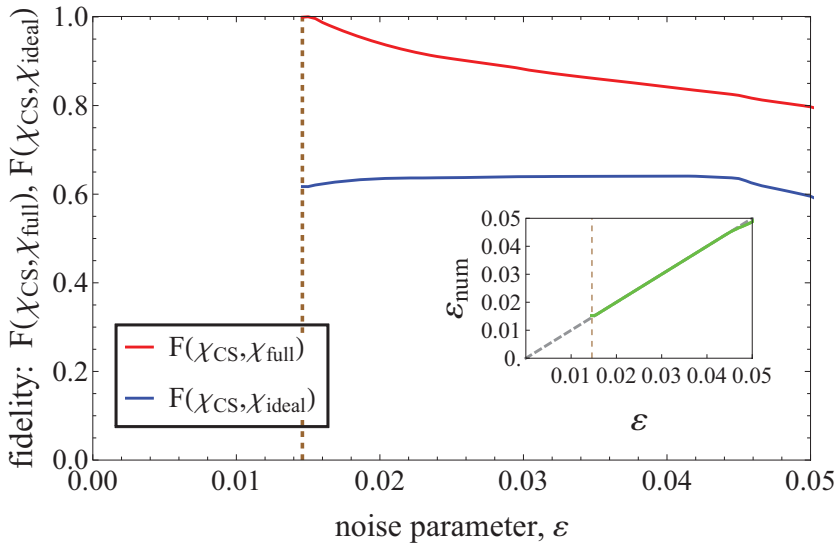

Figure 5.3: Similar to Fig. 5.1, but for the CZ gate, realized with superconducting phase qubits. The process fidelity $F(\chi_{\rm full}, \chi_{\rm ideal}) = 0.62$ is much lower than that for the Xmon qubit gate. The CS QPT procedure, applied to the full data set for the superconducting phase qubit, with varying noise parameter ε. The red (upper) line shows the fidelity $F(\chi_{\rm CS}, \chi_{\rm full})$ between the process matrix $\chi_{\rm CS}$ obtained using the compressed-sensing method and the matrix $\chi_{\rm full}$ obtained using the least-squares method. The blue (lower) line shows the process fidelity $F(\chi_{\rm CS}, \chi_{\rm ideal})$, i.e., compared with the matrix $\chi_{\rm ideal}$ of the ideal unitary process. The vertical dashed brown line corresponds to the noise level $\varepsilon_{\rm opt} = ||\vec{P}^{\rm exp}_{\rm full} - \Phi\vec{\chi}_{\rm full}||_{\ell_2}/\sqrt{M} = 0.0146$ obtained in the LS procedure. The inset shows $\varepsilon_{\rm num} = ||\vec{P}^{\rm exp}_{\rm full} - \Phi\vec{\chi}_{\rm CS}||_{\ell_2}/\sqrt{M}$ as a function of ε (green line); for comparison, the dashed line shows the expected straight line, $\varepsilon_{\rm num} = \varepsilon$. The numerical calculations have been carried out in the Pauli-error basis using CVX-SeDuMi package.

50 times, choosing different random configurations. The error bars in Fig. 5.4 show the standard deviations ($\pm\sigma$) calculated using these 50 numerical experiments, while the central points correspond to the average values.

We see that the upper (red) line starts with fidelity $F(\chi_{\rm CS}, \chi_{\rm full}) = 0.995$ for the full data set ($m_{\rm conf} = 144$) and decreases with decreasing $m_{\rm conf}$. It is important that this decrease is not very strong, so that we can reconstruct the process matrix reasonably accurately, using only a small fraction of the QPT data. We emphasize that the system of equations (4.8) in the standard QPT procedure becomes underdetermined at $m_{\rm conf} < 80$; nevertheless, the CS method reconstructs $\chi_{\rm full}$

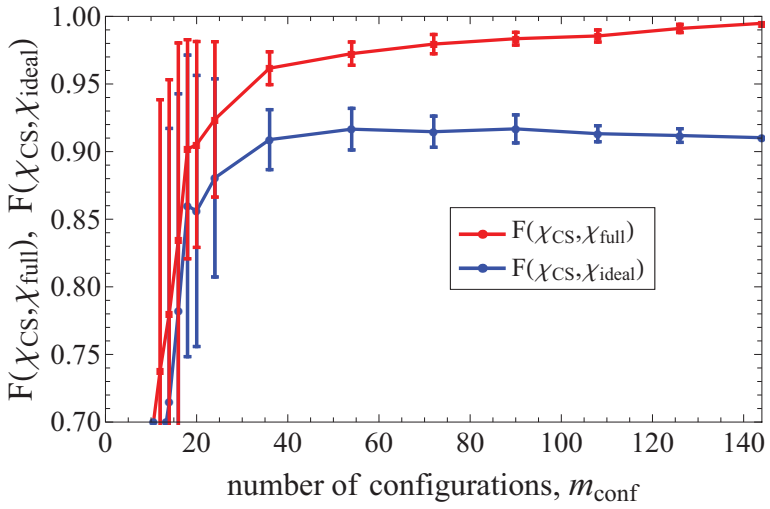

Figure 5.4: The CS method results using a reduced data set with randomly chosen $m_{\rm conf}$ configurations. The red (upper) line shows the fidelity $F(\chi_{\rm CS}, \chi_{\rm full})$ between the CS-estimated process matrix $\chi_{\rm CS}$ and the matrix $\chi_{\rm full}$ obtained from the full data set. The blue (lower) line shows the estimated process fidelity $F_\chi = F(\chi_{\rm CS}, \chi_{\rm ideal})$. The procedure of randomly choosing $m_{\rm conf}$ out of 144 configurations is repeated 50 times; the error bars show the calculated standard deviations. The noise parameter $\varepsilon = 0.02015$ is chosen slightly above $\varepsilon_{\rm opt}$. The calculations are carried out in the Pauli-error basis using CVX-SeDuMi. The experimental data are for the CZ gate realized with Xmon qubits; the process fidelity is $F(\chi_{\rm full}, \chi_{\rm ideal}) = 0.907$.

quite well for $m_{\rm conf} \gtrsim 40$ and still gives reasonable results for $m_{\rm conf} \gtrsim 20$. In particular, for $m_{\rm conf}$ between 40 and 80, the reconstruction fidelity $F(\chi_{\rm CS}, \chi_{\rm full})$ changes between 0.96 and 0.98.

The lower (blue) line in Fig. 5.4 shows that the process fidelity $F_\chi = F(\chi_{\rm CS}, \chi_{\rm ideal})$ can also be found quite accurately, using only $m_{\rm conf} \gtrsim 40$ configurations (the line remains practically flat), and the CS method still works reasonably well down to $m_{\rm conf} \gtrsim 20$. Even though the blue line remains practically flat down to $m_{\rm conf} \simeq 40$, the error bars grow, which means that in a particular experiment with substantially reduced set of QPT data, the estimated process fidelity F_χ may noticeably differ from the actual value. It is interesting that the error bars become very large at approximately the same value ($m_{\rm conf} \simeq 20$), for which the average values for the red and blue lines become unacceptably low.

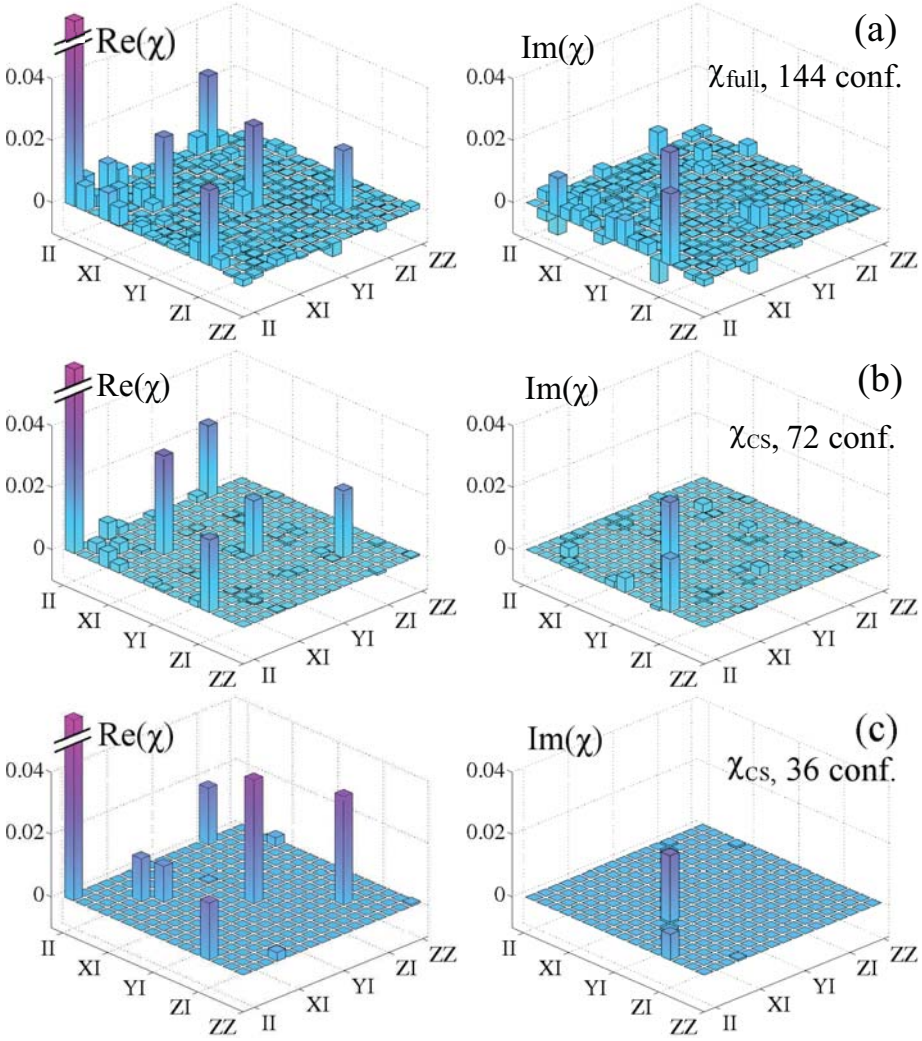

Figure 5.5: (a) The process matrix χ_{full} based on the full data set (144 configurations) and (b,c) the CS-estimated matrices χ_{CS} using a reduced data set: 72 configurations (b) and 36 configurations (c). The process matrices are shown in the Pauli-error basis. The main element $\chi_{II,II}$ (process fidelity) is off the scale and therefore is cut; its height is 0.907, 0.918, and 0.899 for the panels (a), (b), and (c), respectively. All other peaks characterize imperfections. The fidelity $F(\chi_{\text{CS}}, \chi_{\text{full}})$ for the matrices in panels (b) and (c) is equal to 0.981 and 0.968, respectively. The middle and lower panels use the data set, corresponding to underdetermined systems of equations. Experimental data for CZ gate realized with Xmon qubits have been used.

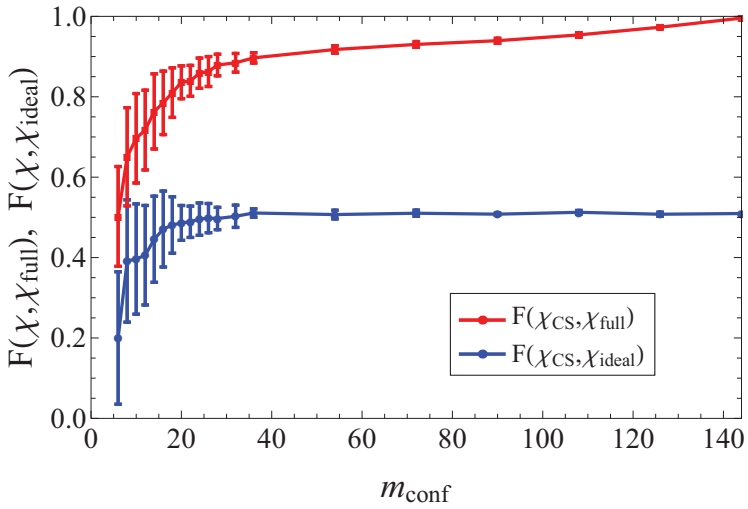

Figure 5.6: Similar to Fig. 5.4, but for the CZ gate realized with superconducting phase qubits. The process fidelity $F(\chi_{\text{full}}, \chi_{\text{ideal}}) = 0.51$ is much lower than that for the Xmon qubit gate. As we see, CS QPT works significantly better for this lower-fidelity gate than for the better gate presented in Fig. 5.4.

Figure 5.5 shows examples of the CS estimated process matrices χ_{CS} for $m_{\text{conf}} = 72$ (middle panel) and $m_{\text{conf}} = 36$ (lower panel), together with the full-data process matrix χ_{full} (upper panel). The process matrices are drawn in the Pauli-error basis to display the process imperfections more clearly. The peak $\chi_{II,II}$ is off the scale and is cut arbitrarily. We see that the CS estimated process matrices are different from the full-data matrix; however the positions of the main peaks are reproduced exactly, and their heights are also reproduced rather well (for a small number of selected configurations the peaks sometimes appear at wrong positions). It is interesting to see that the CS procedure suppressed the height of minor peaks. Note that both presented χ_{CS} are based on the data sets corresponding to underdetermined system of equations.

The computer resources needed for the calculation of results presented in Fig. 5.4 are not demanding. The calculations require about 30 MB of computer memory and 2–4 seconds time for a modest PC per individual calculation (smaller time for smaller number of configurations).

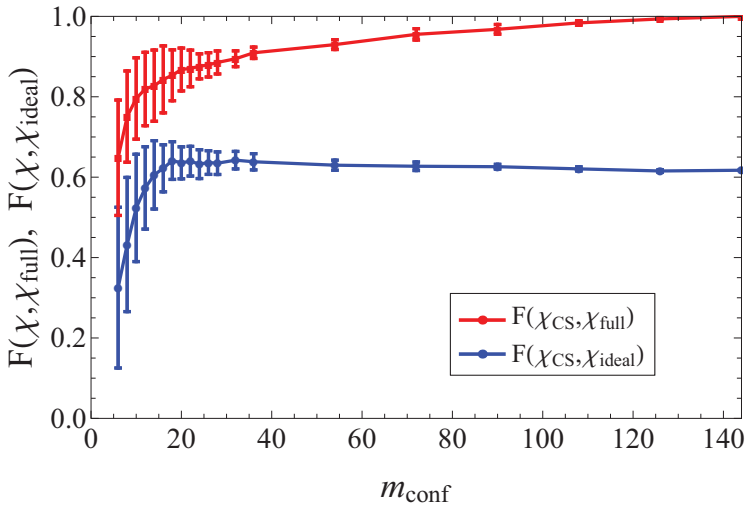

Figure 5.7: Similar to Fig. 5.4, but for the CZ gate realized with superconducting phase qubits. The process fidelity $F(\chi_{\text{full}}, \chi_{\text{ideal}}) = 0.62$ is much lower than that for the Xmon qubit gate. As we see, CS QPT works significantly better for this lower-fidelity gate than for the better gate presented in Fig. 5.4.

Besides the presented results, we have also performed analysis for the CS QPT of two CZ gates based on phase qubits. The results are qualitatively similar, except the process fidelity for phase-qubit gates was significantly lower: 0.62 and 0.51. The results for these gates are presented in Figs. 5.6 and 5.7. Comparing with Fig. 5.4, we see that CS QPT works better for this lower-fidelity gate. In particular, the blue line in Figs. 5.6 and 5.7 are practically flat down to $m_{\text{conf}} \simeq 20$ and the error bars are quite small (in Fig. 5.7 the blue line has even values slightly exceeding the process fidelity $F(\chi_{\text{full}}, \chi_{\text{ideal}}) = 0.62$ for $m_{\text{conf}} \gtrsim 20$. We think that the CS QPT works better for a lower-fidelity gate because experimental imperfections affect the measurement error relatively less in this case than for a higher-fidelity gate.

Thus our results show that for a CZ gate realized with superconducting qubits CS QPT can reduce the number of used QPT configurations by up to a factor of 7 compared with full QPT, and up to a factor of 4 compared with the threshold at which the system of equations for the standard QPT becomes underdetermined.

5.6 Reduced data set, nonoptimal noise parameter ε

As mentioned above, in a QPT experiment with a reduced data set, there is no straightforward way to find the near-optimal value of the noise parameter ε (which we find here from the full data set). Therefore, it is important to check how well the CS method works when a nonoptimal value of ε is used. Figure 5.8 shows the results obtained using the experimental data for the CZ gate realized with Xmon qubits, similar to those in Fig. 5.4, but with several values of the noise parameter: $\varepsilon/\varepsilon_{\text{opt}} = 1.01$, 1.2, 1.4, 1.6, and 1.8. The upper panel shows the fidelity between the matrix χ_{CS} and the full-data matrix χ_{full}; the lower panel shows the process fidelity $F(\chi_{\text{CS}}, \chi_{\text{ideal}})$. We see that the fidelity of the χ matrix estimation, $F(\chi_{\text{CS}}, \chi_{\text{full}})$, becomes monotonously worse with increasing ε, while the estimated process fidelity, $F(\chi_{\text{CS}}, \chi_{\text{ideal}})$, may become larger when a nonoptimal ε is used. This observation agrees with the results presented in Fig. 5.1, where we saw that for the full data set (the most right points in Figs. 5.8) the fidelity between the matrix χ_{CS} and the full-data matrix χ_{full}, $F(\chi_{\text{CS}}, \chi_{\text{full}})$, monotonously decreases for increasing nonoptimal values of ε, while the process fidelity, $F(\chi_{\text{CS}}, \chi_{\text{ideal}})$, may take larger values for nonoptimal values of ε.

Similar results are presented in Fig. 5.9 and Fig. 5.10 for the CZ gate based on phase qubits (see Fig. 5.6 and Fig. 5.7) with the process fidelities of $F(\chi_{\text{full}}, \chi_{\text{ideal}}) = 0.62$ and $F(\chi_{\text{full}}, \chi_{\text{ideal}}) = 0.51$, respectively. The values of the noise parameter were chosen as $\varepsilon/\varepsilon_{\text{opt}} = 1.01$, 2, 3, 4 (and even 5 for the CZ gate with $F(\chi_{\text{full}}, \chi_{\text{ideal}}) = 0.51$). The results presented in Fig. 5.9 and Fig. 5.10 show significantly better tolerance to a nonoptimal choice of ε; in particular, even for $\varepsilon = 3\varepsilon_{\text{opt}}$ the process fidelity practically coincides with the blue lines in Fig. 5.7 and Fig. 5.6 (obtained for $\varepsilon \approx \varepsilon_{\text{opt}}$). This result is in agreement with the behaviour of the process fidelities presented in Fig. 5.3 and Fig. 5.2 for the full data, where we saw that the process fidelity remained practically

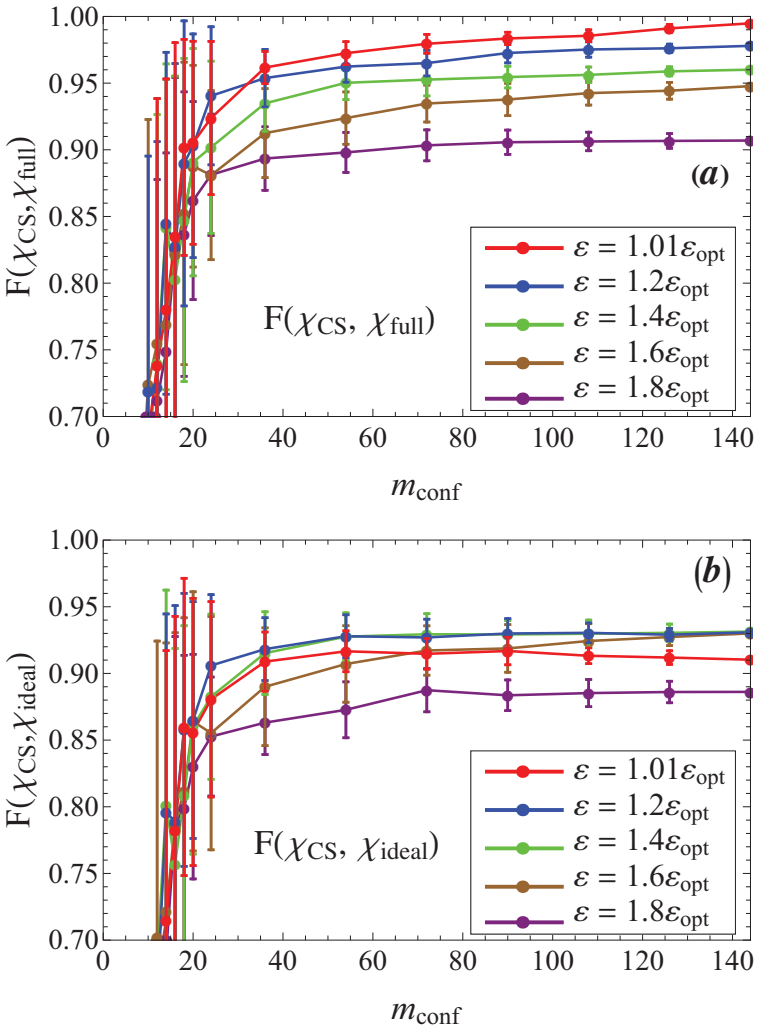

Figure 5.8: (a) Fidelity $F(\chi_{\rm CS}, \chi_{\rm full})$ of the process matrix estimation and (b) the estimated process fidelity $F(\chi_{\rm CS}, \chi_{\rm ideal})$ as functions of the data set size for several values of the noise parameter ε used in the CS optimization: $\varepsilon/\varepsilon_{\rm opt} = 1.01$, 1.2, 1.4, 1.6, and 1.8. Error bars show the standard deviations calculated using 50 random selections of reduced data sets. The red lines are the same as the lines in Fig. 5.4. The experimental data are for the CZ gate realized with Xmon qubits; the process fidelity is $F(\chi_{\rm full}, \chi_{\rm ideal}) = 0.907$.

flat for a wide range of ε, approximately up to $4\varepsilon_{\rm opt}$. We believe the lower gate fidelity for phase qubits is responsible for this relative insensitivity to the choice of ε.

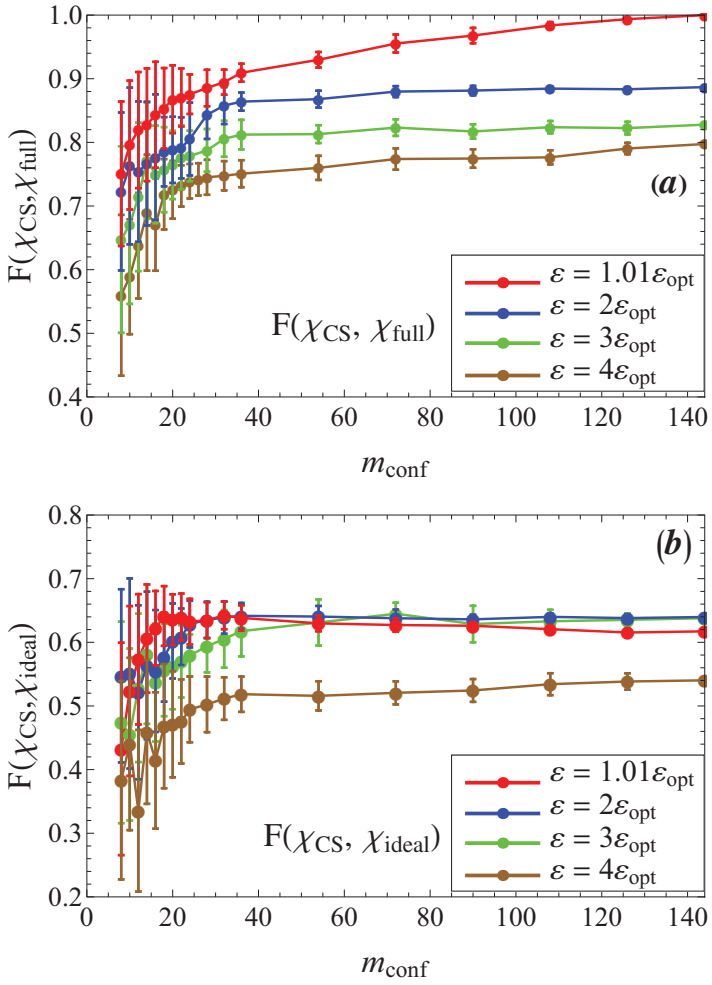

Figure 5.9: (a) Fidelity $F(\chi_{\rm CS}, \chi_{\rm full})$ of the process matrix estimation and (b) the estimated process fidelity $F(\chi_{\rm CS}, \chi_{\rm ideal})$ as functions of the data set size for several values of the noise parameter ε used in the CS optimization: $\varepsilon/\varepsilon_{\rm opt} = 1.01$, 2.0, 3.0, and 4.0. Error bars show the standard deviations calculated using 50 random selections of reduced data sets. The red lines are the same as the lines in Fig. 5.7. The experimental data are for the CZ gate realized with the phase qubits; the process fidelity is $F(\chi_{\rm full}, \chi_{\rm ideal}) = 0.62$.

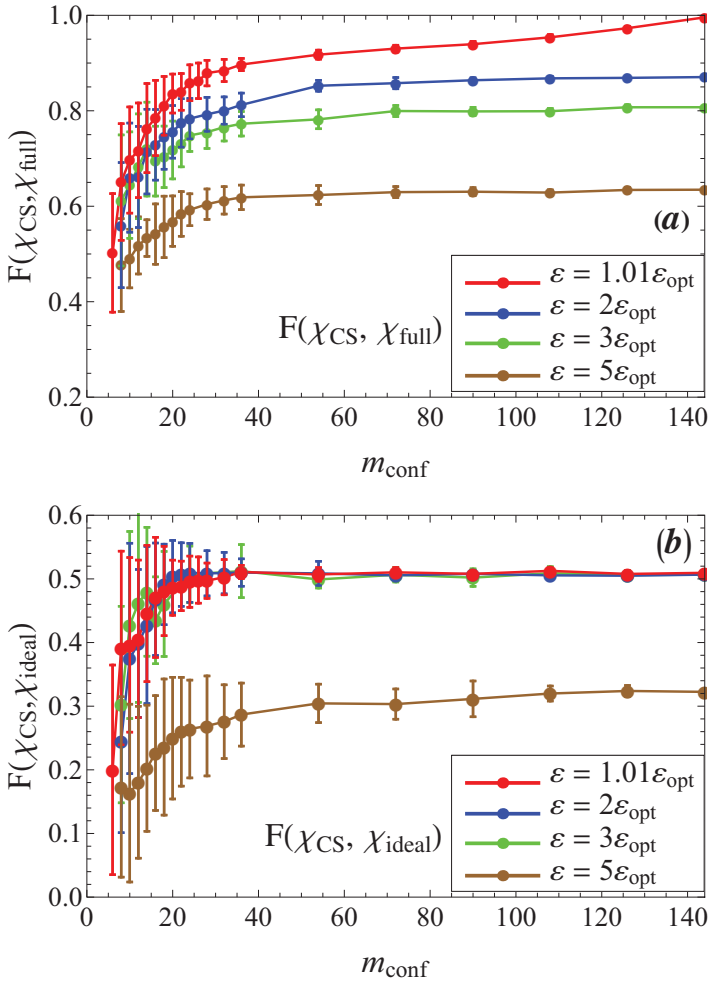

Figure 5.10: (a) Fidelity $F(\chi_{\rm CS}, \chi_{\rm full})$ of the process matrix estimation and (b) the estimated process fidelity $F(\chi_{\rm CS}, \chi_{\rm ideal})$ as functions of the data set size for several values of the noise parameter ε used in the CS optimization: $\varepsilon/\varepsilon_{\rm opt} = 1.01$, 2.0, 3.0, and 5.0. Error bars show the standard deviations calculated using 50 random selections of reduced data sets. The red lines are the same as the lines in Fig. 5.6. The experimental data are for the CZ gate realized with the phase qubits; the process fidelity is $F(\chi_{\rm full}, \chi_{\rm ideal}) = 0.51$.

5.7 Comparison between Pauli-error and SVD bases

So far for the CS method we have used the Pauli-error basis, in which the process matrix χ is expected to be approximately sparse because the ideal process matrix χ_{ideal} contains only one non-zero element, $\chi_{\text{ideal}\,II,II} = 1$. However, there is an infinite number of the operator bases with this property: for example, the SVD basis (see Appendix C) suggested in Refs. [39] and [40]. The process matrices are different in the Pauli-error and SVD bases, therefore the CS optimization should produce different results in two different bases (which entails solving two different optimization problems). To compare the results, we do the CS optimization in the SVD basis, then convert the resulting matrix χ into the Pauli-error basis, and calculate the fidelity $F(\chi_{\text{CS-SVD}}, \chi_{\text{CS}})$ between the transformed process matrix and the matrix χ_{CS} obtained using optimization in the Pauli-error basis directly.

Figure 5.11 shows the comparison of the results obtained in the Pauli-error basis and the SVD-basis, for the experimental data for the CZ gate realized with Xmon qubits. The green line in Fig. 5.11 shows $F(\chi_{\text{CS-SVD}}, \chi_{\text{CS}})$ as a function of the selected size of the data set for the CZ gate realized with Xmon qubits, similar to Fig. 5.4 (the same ε is used). We also show the fidelity between the SVD-basis-obtained matrix $\chi_{\text{CS-SVD}}$ and the full-data matrix χ_{full} as well as the ideal process matrix χ_{ideal}. For comparison we also include the lines shown in Fig. 5.4 (dashed lines), obtained using the Pauli-error basis. As we see, the results obtained in the two bases are close to each other, though the SVD basis seems to work a little better at small data sizes, $m_{\text{conf}} \simeq 20$. We also include a similar plot Fig. 5.12 presenting analogous results for the data, obtained using the experimental data for phase qubits with the much lower process fidelity $F(\chi_{\text{full}}, \chi_{\text{ideal}}) = 0.51$. As we see, for the low-fidelity quantum gate, the results obtained in two bases are very similar to each other for most values of m_{conf}, except for the very low ones below $m_{\text{conf}} \simeq 20$.

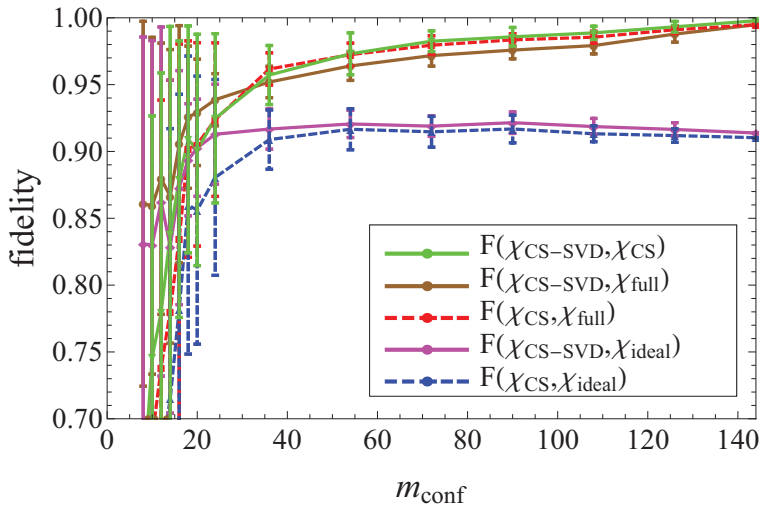

Figure 5.11: Comparison between the CS results obtained in the SVD and Pauli-error bases for the CZ gate realized with Xmon qubits. The green line shows the relative fidelity $F(\chi_{\text{CS-SVD}}, \chi_{\text{CS}})$ as a function of the number m_{conf} of randomly selected configurations. We also show the fidelities $F(\chi_{\text{CS-SVD}}, \chi_{\text{full}})$ (brown line), $F(\chi_{\text{CS}}, \chi_{\text{full}})$ (red dashed line), and process fidelities $F(\chi_{\text{CS-SVD}}, \chi_{\text{ideal}})$ (magenta line) and $F(\chi_{\text{CS}}, \chi_{\text{ideal}})$ (blue dashed line). The dashed lines have been shown in Fig. 5.4. The results using the SVD basis are somewhat more accurate than those for the Pauli-error basis when $m_{\text{conf}} < 40$.

The visual comparison of χ-matrices obtained in these bases (as in Fig. 5.5), presented here at Fig. 5.13 and Fig. 5.14, show that although we worked in two different bases, the process matrices recovered via CS exhibit a similar structure. In addition, the real and imaginary parts of the process matrix χ_{full} reconstructed using the full set of experimental data, are plotted in Fig. 5.15. Experimental data for the CZ gate realized with the phase qubits with the process fidelity $F(\chi_{\text{full}}, \chi_{\text{ideal}}) = 0.51$ have been used in Fig. 5.13 – Fig. 5.15.

It should be noted that the calculations in the SVD basis are somewhat faster ($\sim$2 seconds per point) and require less memory ($\sim$6 MB) than the calculations in the Pauli-error basis. This is because the matrix Φ defined in Eq. (4.8) for the CZ gate contains about half the number of non-zero elements in the SVD basis than in the Pauli-error basis.

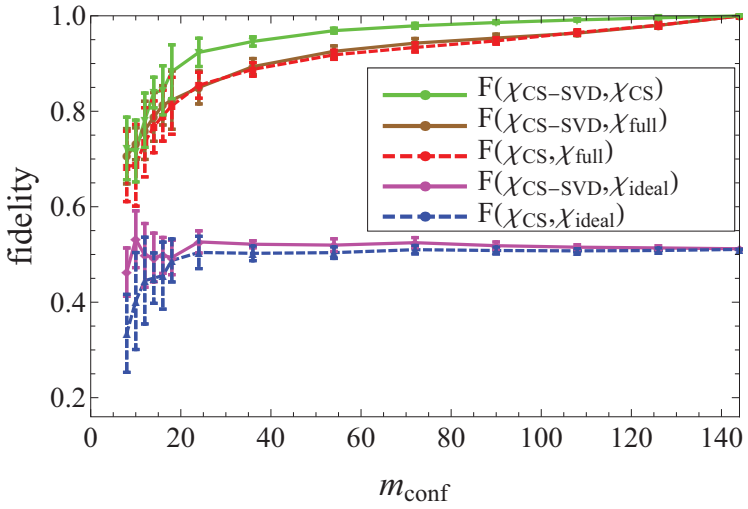

Figure 5.12: Comparison between the CS results obtained in the SVD and Pauli-error bases for the low-fidelity CZ gate realized with phase qubits. The green line shows the relative fidelity $F(\chi_{\text{CS-SVD}}, \chi_{\text{CS}})$ as a function of the number m_{conf} of randomly selected configurations. We also show the fidelities $F(\chi_{\text{CS-SVD}}, \chi_{\text{full}})$ (brown line), $F(\chi_{\text{CS}}, \chi_{\text{full}})$ (red dashed line), and process fidelities $F(\chi_{\text{CS-SVD}}, \chi_{\text{ideal}})$ (magenta line) and $F(\chi_{\text{CS}}, \chi_{\text{ideal}})$ (blue dashed line). The dashed lines have been shown in Fig. 5.6. The results using the SVD basis and the Pauli-error basis are very close to each other for most values of m_{conf}. The experimental data are for the low-fidelity CZ gate realized with the phase qubits; the process fidelity is $F(\chi_{\text{full}}, \chi_{\text{ideal}}) = 0.51$.

All results presented here are obtained using the CVX-SeDuMi package. The results for the CZ gate obtained using the YALMIP-SDPT3 package are similar when the same value of ε is used. Surprisingly, in our realization of computation, the YALMIP-SDPT3 package still finds reasonable solutions when ε is significantly smaller than ε_{opt} (even when ε is zero or negative), so that the problem cannot have a solution; apparently in this case the solver increases the value of ε until a solution is found. This may seem to be a good feature of YALMIP-SDPT3. However, using $\varepsilon < \varepsilon_{\text{opt}}$ should decrease the accuracy of the result (see the next subsection). Moreover, YALMIP-SDPT3 does not work well for the Toffoli gate discussed in Section 5.9. Thus we conclude that CVX-SeDuMi package is better than YALMIP-SDPT3 package for our CS calculations. (Note that this finding may be specific to our system.)

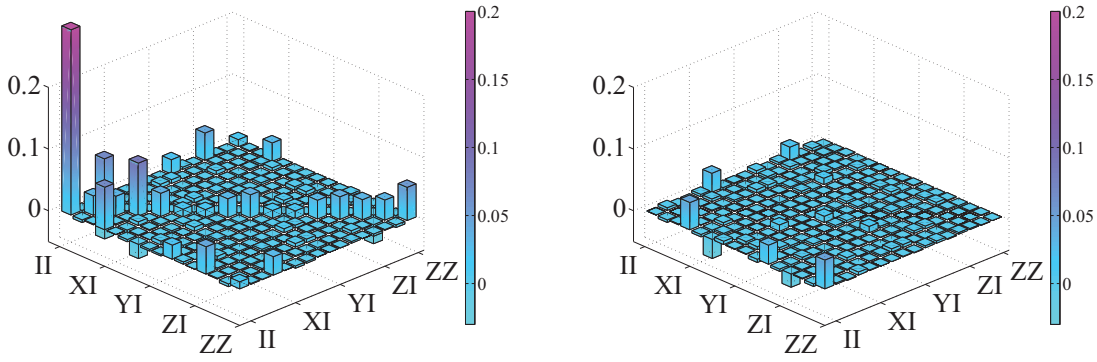

Figure 5.13: Real (left figure) and imaginary (right figure) parts of the process matrix χ_{CS-SVD} for the CZ gate, calculated in the SVD basis and converted into the Pauli-error basis. We used 36 randomly chosen measurement configurations out of the full set of 144. The fidelity $F(\chi_{CS-SVD}, \chi_{\text{full}}) = 0.88$, the fidelity $F(\chi_{CS}, \chi_{CS-SVD}) = 0.94$. The main element $\text{Re}(\chi_{II,II})$ (process fidelity) is off the scale and therefore is cut; its height is 0.52. The experimental data are for the CZ gate realized with the phase qubits; the process fidelity calculated from full data is $F(\chi_{\text{full}}, \chi_{\text{ideal}}) = 0.51$.

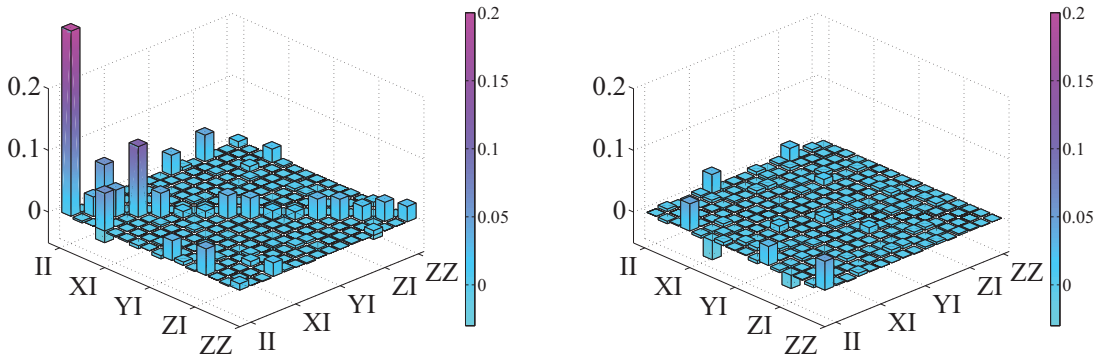

Figure 5.14: Real (left figure) and imaginary (right figure) parts of the process matrix χ_{CS} for the CZ gate calculated directly in the Pauli-error basis. We used 36 randomly chosen measurement configurations out of the full set of 144. The fidelity $F(\chi_{CS}, \chi_{\text{full}}) = 0.91$, the fidelity $F(\chi_{CS}, \chi_{CS-SVD}) = 0.94$. The main element $\text{Re}(\chi_{II,II})$ (process fidelity) is off the scale and therefore is cut; its height is 0.50. The experimental data are for the CZ gate realized with the phase qubits; the process fidelity calculated from full data is $F(\chi_{\text{full}}, \chi_{\text{ideal}}) = 0.51$.

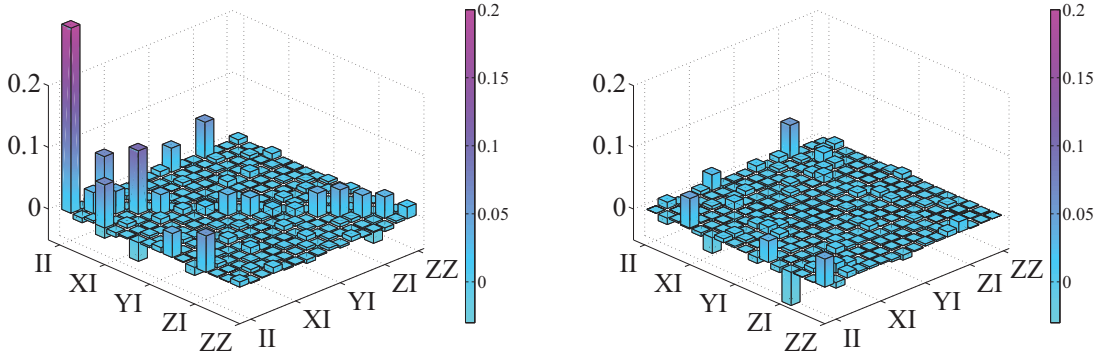

Figure 5.15: Real (left figure) and imaginary (right figure) parts of the process matrix χ_{LS} for the CZ gate calculated from the full data. The main element $\mathrm{Re}(\chi_{II,II})$ (process fidelity) is off the scale and therefore is cut; its height is 0.51. The experimental data are for the CZ gate realized with the phase qubits; the process fidelity is $F(\chi_{\text{full}}, \chi_{\text{ideal}}) = 0.51$.

5.8 Comparison with least-squares minimization

Besides using the CS method for reduced data sets, we also used the LS minimization [with constraints Eq. (4.5) and Eq. (4.6) from Section 4.1 of Chapter 4] for the same reduced sets. Solid lines in Fig. 5.16 show the resulting fidelity $F(\chi_{\text{LS}}, \chi_{\text{full}})$ compared with the full-data process matrix and the estimated process fidelity $F(\chi_{\text{LS}}, \chi_{\text{ideal}})$.

Somewhat surprisingly, the LS method still works (though less well) in a significantly underdetermined regime. Naively, we would expect that in this case Eq. (4.8) from Section 4.1 of Chapter 4 can be satisfied exactly, and there are many exact solutions corresponding to the null space of the selected part of the matrix Φ. However, numerical results show that in reality Eq. (4.8) cannot be satisfied exactly unless the selected data set is very small. The reason is that the matrix χ has to be positive, and the (corrected) experimental probabilities can be close to the limits of the physical range or even outside it.

The problem is that the experimental probabilities are not directly obtained from the experiment, but are corrected for imperfect measurement fidelity [113]. As a result, they may become

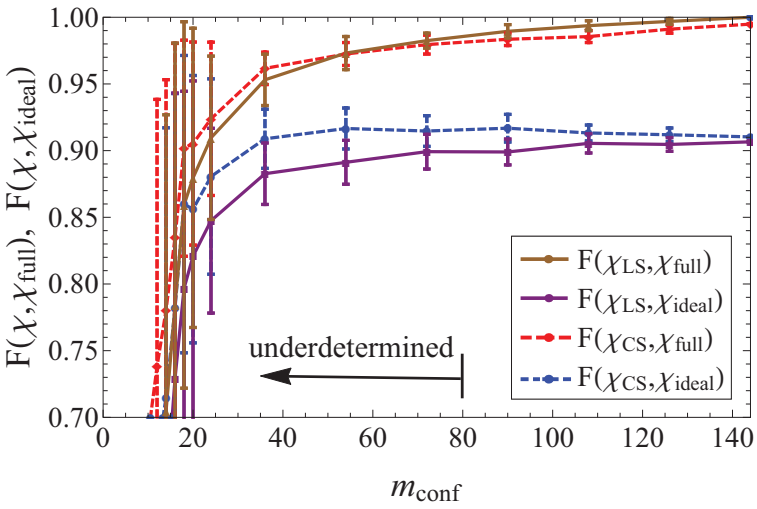

Figure 5.16: Comparison between the results obtained by the LS and CS methods. The solid lines are for the LS method, the dashed lines (same as in Fig. 5.4) are for the CS method. The CS method is more accurate for a substantially reduced data set. The experimental data are for the CZ gate realized with Xmon qubits.

larger than one or smaller than zero. This happens fairly often for high fidelity gates because for an ideal operation the measurement results are often zeros and ones, so the experimental probabilities should also be close to zero or one. Any additional deviation due to imperfect correction for the measurement fidelity may then push the probabilities outside of the physical range. It is obvious that in this case Eq. (4.8) cannot be satisfied exactly for any physical χ. To resolve this problem one could consider rescaling the probabilities in such instances, so that they are exactly one or zero instead of lying outside the range. However, this also does not help much because a probability of one means that the resulting state is pure, so this strongly reduces the number of free parameters in the process matrix χ. As a result, Eq. (4.8) cannot be satisfied exactly, and the LS minimization is formally possible even in the underdetermined case.

Another reason why Eq. (4.8) may be impossible to satisfy in the underdetermined case, is that the randomly selected rows of the matrix Φ can be linearly dependent. Then mathematically

some linear relations between the experimental probabilities must be satisfied, while in reality they are obviously not satisfied exactly.

These reasons make the LS minimization a mathematically possible procedure even in the underdetermined regime. However, as we see from Fig. 5.16, in this case the procedure works less well than the compressed sensing, estimating the process matrix and process fidelity with a lower accuracy. Similar calculations for the CZ gate realized with phase qubits (not presented here) also show that the LS method does not work well at relatively small m_{conf}. The advantage of the compressed sensing in comparison with the LS minimization becomes even stronger for the three-qubit Toffoli gate considered in the next Section. Note though that when the selected data set is large enough to give an overdetermined system of equations Eq. (4.8), the LS method works better than the CS method. Therefore, the compressed sensing is beneficial only for a substantially reduced (underdetermined) data set, which is exactly the desired regime of operation.

5.9 Three-qubit CS QPT for Toffoli gate

In this Section we apply the compressed sensing method to simulated tomographic data corresponding to a three-qubit Toffoli gate [2, 157, 158, 113, 130]. As discussed in Section 4.2 of Chapter4, the process matrix of a three-qubit gate contains $16^3 - 4^3 = 4032$ independent real parameters, while the full QPT requires $M_{\text{conf}} = 12^3 = 1728$ measurement configurations yielding a total of $M = 12^3 \times 2^3 = 13824$ experimental probabilities, if we use $n_{\text{in}} = 4$ initial states and $n_{\text{R}} = 3$ measurement rotations per qubit, with all qubits measured independently. If we work with a partial data set, the system of equations Eq. (4.8) becomes underdetermined if the number m_{conf} of used configurations is less than $4032/7 = 576$. In such a regime the traditional maximum likelihood or LS methods are not expected to provide a good estimate of the process matrix. In this section we demonstrate that for our simulated Toffoli gate the compressed sensing method works well even for a much smaller number of configurations, $m_{\text{conf}} \ll 576$.

For the analysis we have simulated experimental data corresponding to a noisy Toffoli gate by adding truncated Gaussian noise with a small amplitude to each of $M = 13824$ ideal measurement probabilities P_i^{ideal}. We assumed the set of experimental probabilities in Eq. (4.8) to be of the form $P_i^{\text{exp}} = P_i^{\text{ideal}} + \Delta P_i$, where ΔP_i are random numbers sampled from the normal distribution with zero mean and a small standard deviation σ. By choosing different values of the standard deviation σ we can change the process fidelity of the simulated Toffoli gate: a smaller value of σ makes the process fidelity closer to 1. After adding the Gaussian noise ΔP_i to the ideal probabilities P_i^{ideal}, we check whether the resulting simulated probabilities P_i^{exp} are in the interval $[0, 1]$. If a P_i^{exp} happens to be outside the interval $[0, 1]$, we repeat the procedure until the condition $P_i^{\text{exp}} \in [0, 1]$ is satisfied. Finally, we renormalize each set of 8 probabilities corresponding to the

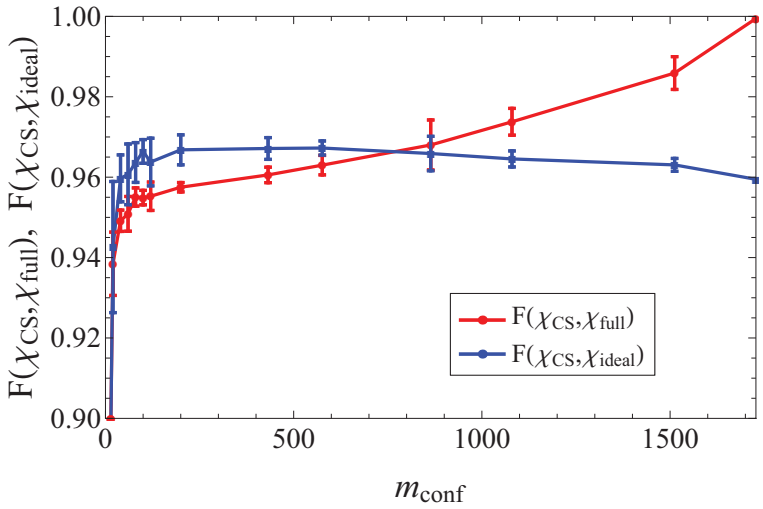

Figure 5.17: CS QPT for a simulated Toffoli gate. Red line: fidelity $F(\chi_{\text{CS}}, \chi_{\text{full}})$ of the process matrix estimation, blue line: the estimated process fidelity $F(\chi_{\text{CS}}, \chi_{\text{ideal}})$, both as functions of the data set size, expressed as the number m_{conf} of randomly selected configurations. The full QPT corresponds to 1728 configurations. The system of equations becomes underdetermined when $m_{\text{conf}} < 576$. The error bars show the standard deviations calculated by repeating the procedure of random selections 7 times.

same measurement configuration so that these probabilities add up to 1.

Thus the simulated imperfect quantum process is defined by $M = 13824$ probabilities, corresponding to $M_{\text{conf}} = 1728$ configurations; the process fidelity for a particular realization (used here) with $\sigma = 0.01$ is $F_\chi = F(\chi_{\text{full}}, \chi_{\text{ideal}}) = 0.959$. We then test efficiency of the compressed sensing method by randomly selecting $m_{\text{conf}} \leq 1728$ configurations, finding the corresponding process matrix χ_{CS}, and comparing it with the full-data matrix χ_{full} by calculating the fidelity $F(\chi_{\text{CS}}, \chi_{\text{full}})$. We also calculate the process fidelity $F(\chi_{\text{CS}}, \chi_{\text{ideal}})$ given by χ_{CS}.

The red line in Fig. 5.17 shows the fidelity $F(\chi_{\text{CS}}, \chi_{\text{full}})$ as a function of the number m_{conf} of randomly selected configurations. The value of ε is chosen to be practically equal to $\varepsilon_{\text{opt}} = ||(\vec{P}^{\text{exp}}_{\text{full}} - \Phi\vec{\chi}_{\text{full}})||_{\ell_2}/\sqrt{M} = 0.01146$ (the relative difference is less than 10^{-3}). The ℓ_1-minimization is done using the CVX-SeDuMi package. The error bars are calculated by repeating

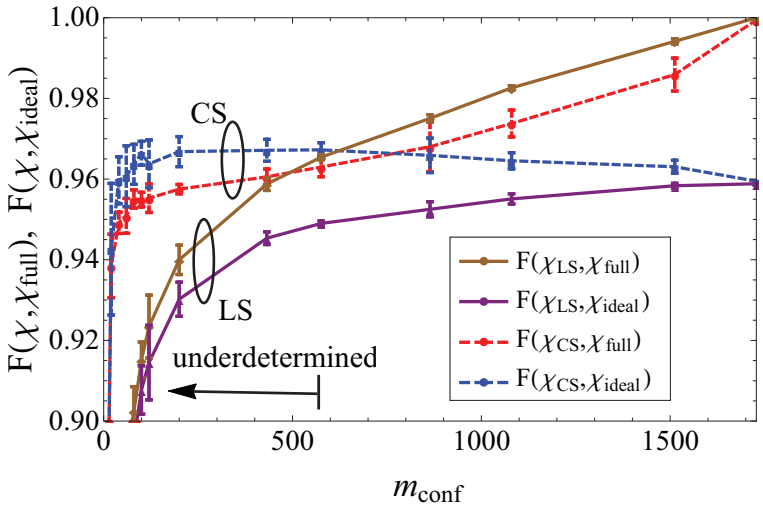

Figure 5.18: Comparison between the calculations using CS and LS methods for the simulated Toffoli gate. Solid lines are for the LS method, dashed lines (the same as in Fig. 5.17) are for the CS method. In the underdetermined regime the CS-method results are much better than the LS-method results.

the procedure of random selection 7 times. We see a reasonably high fidelity $F(\chi_{\text{CS}}, \chi_{\text{full}})$ of the reconstructed process matrix even for small numbers of selected configurations. For example, $F(\chi_{\text{CS}}, \chi_{\text{full}}) = 0.95$ for only $m_{\text{conf}} = 40$ configurations, which represents a reduction by more than a factor of 40 compared with the full QPT and approximately a factor of 15 compared with the threshold of the underdetermined system of equations.

The blue line in Fig. 5.17 shows the process fidelity $F(\chi_{\text{CS}}, \chi_{\text{ideal}})$ calculated by the CS method. We see that it remains practically flat down to $m_{\text{conf}} \gtrsim 40$, which means that χ_{CS} can be used efficiently to estimate the actual process fidelity.

Figure 5.18 shows similar results calculated using the LS method (for comparison the lines from Fig. 5.17 are shown by dashed lines). We see that the LS method still works in the underdetermined regime ($m_{\text{conf}} < 576$); however, it works significantly worse than the CS method. As an example, for $m_{\text{conf}} = 40$ the fidelity of the process matrix estimation using the LS method

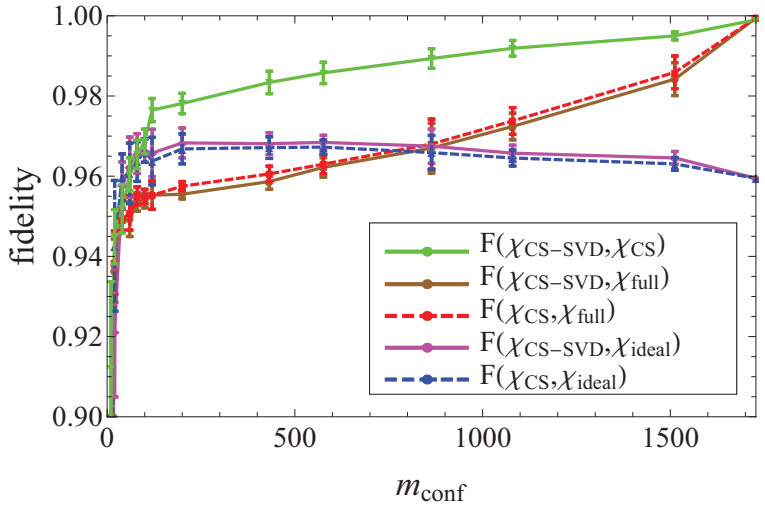

Figure 5.19: Comparison between the CS results obtained in the SVD and Pauli-error bases for the simulated Toffoli gate. The green line shows the relative fidelity $F(\chi_{\text{CS-SVD}}, \chi_{\text{CS}})$ as a function of the number m_{conf} of randomly selected configurations. We also show the fidelities $F(\chi_{\text{CS-SVD}}, \chi_{\text{full}})$ (brown line), $F(\chi_{\text{CS}}, \chi_{\text{full}})$ (red dashed line), and process fidelities $F(\chi_{\text{CS-SVD}}, \chi_{\text{ideal}})$ (magenta line) and $F(\chi_{\text{CS}}, \chi_{\text{ideal}})$ (blue dashed line). The dashed lines have been shown in Fig. 5.17. The results using the SVD basis and the Pauli-error basis are very close to each other for most values of m_{conf}. The error bars show the standard deviations calculated using 7 random selections of reduced data sets.

is $F(\chi_{\text{LS}}, \chi_{\text{full}}) = 0.86$, which is significantly less than $F(\chi_{\text{CS}}, \chi_{\text{full}}) = 0.95$ for the CS method. Similarly, for $m_{\text{conf}} = 40$ the process fidelity obtained via the CS method, $F(\chi_{\text{CS}}, \chi_{\text{ideal}}) = 0.96$ is close to the full-data value of 0.959, while the LS-method value, $F(\chi_{\text{LS}}, \chi_{\text{ideal}}) = 0.85$, is quite different.

Besides using the Pauli-error basis for the results shown in Fig. 5.17, we have also performed the calculations using the SVD basis. The results are presented in Fig. 5.19 and as we see, they are very close to those in Fig. 5.17. In particular, we see that the relative fidelity $F(\chi_{\text{CS-SVD}}, \chi_{\text{CS}})$ is above 0.98 for $m_{\text{conf}} > 200$ and above 0.95 for $m_{\text{conf}} > 40$.

We have also performed the calculations using non-optimal values of the noise parameter ε. In comparison with the results for CZ gate shown in Fig. 5.8, the results for the Toffoli gate

(shown in Fig. 5.20) are more sensitive to the variation of ε. In particular, the fidelity $F(\chi_{\text{CS}}, \chi_{\text{full}})$ is about 0.94 for $\varepsilon = 1.2\varepsilon_{\text{ opt}}$ (not significantly depending on m_{conf} for $m_{\text{conf}} > 40$) and the process fidelity $F(\chi_{\text{CS}}, \chi_{\text{ideal}})$ for $\varepsilon = 1.2\varepsilon_{\text{opt}}$ is approximately 0.94 instead of the actual value 0.96. Note that in contrast to Fig. 5.8, error bars showing the standard deviations at Fig. 5.20 were calculated using only 7 random selections of reduced data sets, because it takes significantly longer time to perform calculations for the three-qubit gate compared with the two-qubit gate.

Compared with the two-qubit case, it takes significantly more computing time and memory to solve the ℓ_1-minimization problem for three qubits. In particular, our calculations in the Pauli-error basis took about 8 hours per point on a personal computer for $m_{\text{conf}} \simeq 1500$ and about 1.5 hours per point for $m_{\text{conf}} \simeq 40$; this is three orders of magnitude longer than for two qubits. The amount of used computer memory was 3–10 GB, which is two orders of magnitude larger than for two qubits. (The calculations in the SVD basis for the Toffoli gate took 1–3 hours per point and $\sim$2 GB of memory.) Such a strong scaling of required computer resources with the number of qubits seems to be the limiting factor in extending the CS QPT beyond three qubits, unless a more efficient algorithm is found. (Note that LS calculations required similar amount of memory, but the computation time was much shorter.)

The presented results have been obtained using the CVX-SeDuMi package. We also attempted to use the YALMIP-SDPT3 package. However, in our realization of computation the calculation results were very unreliable for $m_{\text{conf}} < 200$ using the SVD basis, and even worse when the Pauli-error basis was used. Therefore we decided to use only the CVX-SeDuMi package for the 3-qubit CS procedure.

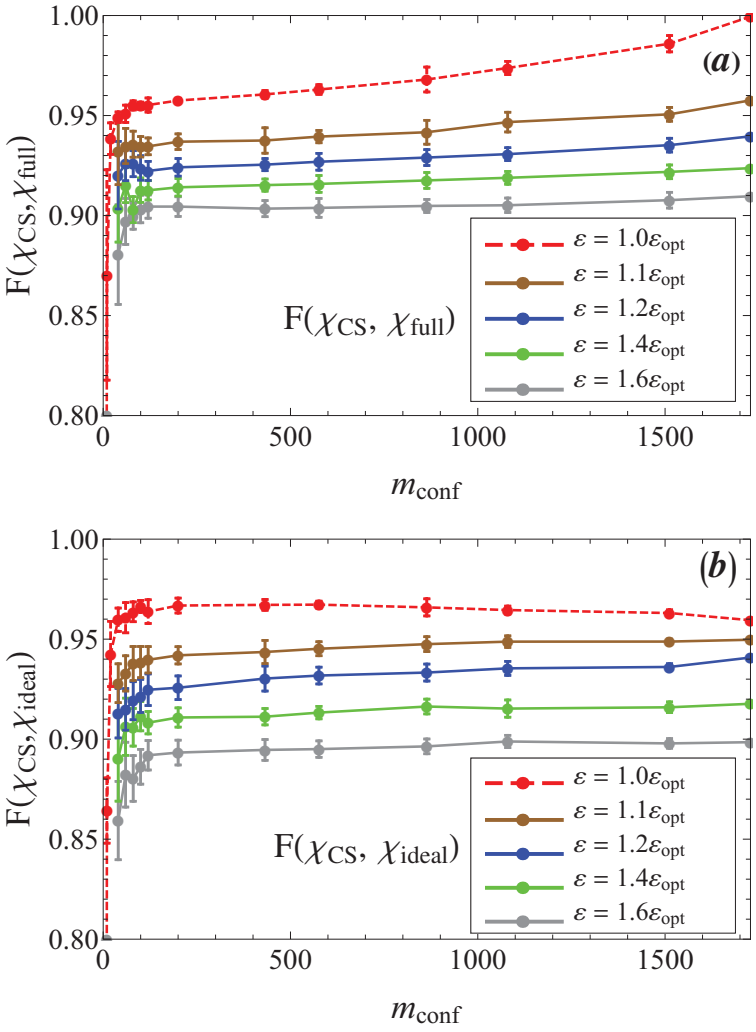

Figure 5.20: (a) Fidelity $F(\chi_{\rm CS}, \chi_{\rm full})$ of the process matrix estimation for the simulated Toffoli gate and (b) the estimated process fidelity $F(\chi_{\rm CS}, \chi_{\rm ideal})$ as functions of the data set size for several values of the noise parameter ε used in the CS optimization: $\varepsilon/\varepsilon_{\rm opt} = 1.0$, 1.1, 1.2, 1.4 and 1.6. The error bars show the standard deviations calculated using 7 random selections of reduced data sets. The red lines are the same as the lines in Fig. 5.17. The process fidelity is $F(\chi_{\rm full}, \chi_{\rm ideal}) = 0.96$.

Chapter 6

Standard deviation of state fidelity

In this Chapter, we consider an equivalent to the process fidelity F_χ characteristic of a quantum gate, called average state fidelity $\overline{F_{\text{st}}}$, which is sometimes also called the "gate fidelity". Since neither of these two fidelities provides any information about fluctuations in the gate fidelity (i.e., how the errors vary over input states), it proves useful to be able to calculate the standard deviation of the average state fidelity. We provide in Section 6.1 the results of the calculation of the average state fidelity and the standard deviation of the state fidelity for both the two-qubit CZ gate, realized with Xmon qubits, and for the three-qubit Toffoli gate with the numerically added noise. In addition we provide the detailed derivation of the formula for the standard deviation of the average state fidelity in Section 6.2, and to confirm our results, we perform the Monte Carlo numerical simulation of the standard deviation of the average state fidelity in Section 6.3.

6.1 Standard deviation of state fidelity

As shown in the previous Chapter, the process matrices χ_{CS} obtained via the CS method allow us to estimate reliably the process fidelity $F_\chi = F(\chi, \chi_{\text{ideal}})$ of a gate using just a small fraction of the full experimental data. While F_χ is the most widely used characteristic of an experimental gate accuracy, it is not the only one. An equivalent characteristic (usually used in randomized benchmarking) is the average state fidelity, defined as $\overline{F_{\text{st}}} = \int \text{Tr}(\rho_{\text{actual}}\rho_{\text{ideal}})\, d|\psi_{\text{in}}\rangle / \int d|\psi_{\text{in}}\rangle$, where the integration is over the initial pure states $|\psi_{\text{in}}\rangle$ (using the Haar measure; it is often assumed that $\int d|\psi_{\text{in}}\rangle = 1$), while the states ρ_{ideal} and ρ_{actual} are the ideal and actual final states for the initial state $|\psi_{\text{in}}\rangle$. The average state fidelity $\overline{F_{\text{st}}}$ is sometimes called the "gate fidelity" [33] and can can naturally be measured via randomized benchmarking [31, 32, 33] ($F_{\text{RB}} = \overline{F_{\text{st}}}$); it is linearly related [159, 160] to the process fidelity F_χ through the formula,

$$\overline{F_{\text{st}}} = \frac{F_\chi d + 1}{d + 1}, \tag{6.1}$$

where $d = 2^N$ is the Hilbert space dimension.

Besides the average state fidelity, an obviously important characteristic of a gate operation is the worst-case state fidelity $F_{\text{st, min}}$, which is minimized over the initial state. Unfortunately, finding the minimum state fidelity is a hard problem from the computational point view, even when the process matrix χ is known. Another natural characteristic is the standard deviation of the state fidelity,

$$\Delta F_{\text{st}} = \sqrt{\overline{F_{\text{st}}^2} - \overline{F_{\text{st}}}^2}, \tag{6.2}$$

where $\overline{F_{\text{st}}^2} = \int [\text{Tr}(\rho_{\text{actual}}\rho_{\text{ideal}})]^2\, d|\psi_{\text{in}}\rangle / \int d|\psi_{\text{in}}\rangle$ is the average square of the state fidelity. The advantage of ΔF_{st} in comparison with $F_{\text{st, min}}$ is that $\overline{F_{\text{st}}^2}$ and ΔF_{st} can be calculated from χ in a

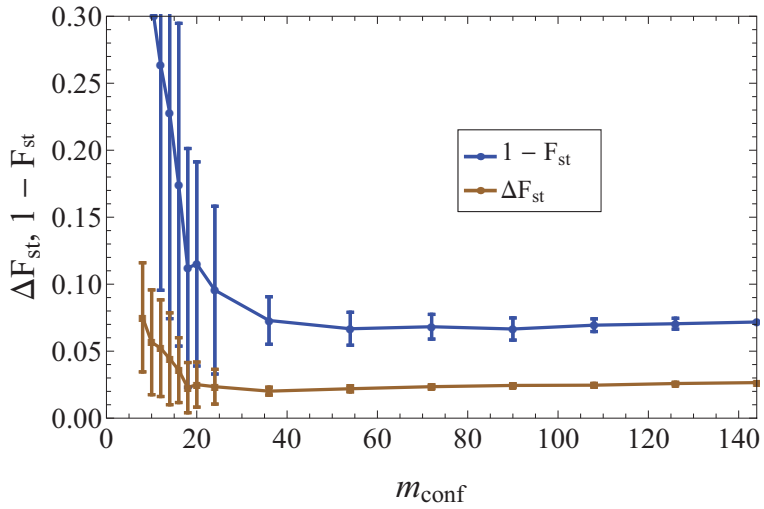

Figure 6.1: Blue (upper) line: average state infidelity $1 - \overline{F_{\rm st}}$ for the CS-estimated process matrix $\chi_{\rm CS}$ as a function of the selected data set size for the experimental CZ gate, realized with Xmon qubits (this line is linearly related to the blue line in Fig. 5.4). Brown (lower) line: the standard deviation of the state fidelity $\Delta F_{\rm st}$, defined via variation of the initial state, Eq. (6.2), using the same $\chi_{\rm CS}$. The error bars are computed by repeating the procedure 50 times with different random selections of used configurations.

straightforward way [161, 162]. Our way of calculating $\overline{F_{\rm st}^2}$ is described in the next Section 6.2 [see Eq. (6.14)].

We have analyzed numerically how well the CS QPT estimates $\Delta F_{\rm st}$ from the reduced data set, using the previously calculated process matrices $\chi_{\rm CS}$ for the experimental CZ gate and the simulated Toffoli gate (considered in Secs. 5.5 and 5.9 of Chapter 5). The results are presented in Figs. 6.1 and 6.2. We show the average state infidelity, $1-\overline{F_{\rm st}}$, and the standard deviation of the state fidelity, $\Delta F_{\rm st}$, as functions of the number of selected configurations, $m_{\rm conf}$. The random selection of used configurations is repeated 50 times for Fig. 6.1 (7 times for Fig. 6.2), the error bars show the statistical variation, while the dots show the average values.

As it may be seen in Figs. 6.1 and 6.2, the CS method estimates reasonably well not only the average state fidelity $\overline{F_{\rm st}}$ (which is equivalent to F_χ presented in Figs. 5.4 and 5.17), but also its

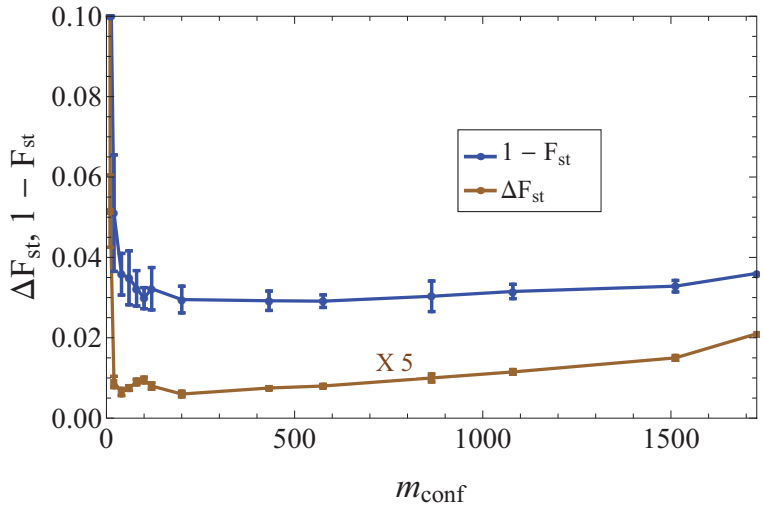

Figure 6.2: The same as in Fig. 6.1, but for the simulated Toffoli gate. The random selection of configurations is repeated 7 times for each point. The results for the standard deviation $\Delta F_{\rm st}$ are multiplied by the factor of 5 for clarity.

standard deviation $\Delta F_{\rm st}$. It is interesting to note that $\Delta F_{\rm st}$ is significantly smaller than the infidelity $1 - \overline{F_{\rm st}}$, which means that the state fidelity $\mathrm{Tr}(\rho_{\rm acual}\rho_{\rm ideal})$ does not vary significantly for different initial states [the ratio $\Delta F_{\rm st}/(1 - \overline{F_{\rm st}})$ is especially small for the simulated Toffoli gate, though this may be because of our particular way of simulation].

6.2 Details of the formula for average square of state fidelity

In this Section we present a detailed derivation of an explicit formula for the squared state fidelity $\overline{F_{\rm st}^2}$, averaged over all pure initial states, for a quantum operation, represented via Kraus operators. We follow the same steps as in Ref. [162], where a closed-form expression for $\overline{F_{\rm st}^2}$ in terms of the process matrix χ was presented. Although our approach is not new, we show it here for completeness. In our derivation we use the theory of permutation operators, symmetric subspaces

and we compute averages over Haar measures. We start this Section from the review of some basic definitions from the theory of permutations.

Definition. A function $f : A \to B$ is called "one-to-one", or "injective", if each element of B appears at most once as the image of an element of A. A function $f : A \to B$ is called "onto", or "surjective", if $f(A) = B$. That is, if each element of B is the image of at least one element of A. A function that is both injective and surjective is called "bijective".

Definition. A "permutation" of a set A is a function $\alpha : A \to A$ that is bijective (i.e. both one-to-one and onto).

Unlike calculus, where most functions are defined on infinite sets and given by formulas, permutations of finite sets are usually given by simply listing where each value goes. For example, we can define a permutation α of the set $S = \{1, 2, 3\}$ by stating:

$$\alpha(1) = 2, \quad \alpha(2) = 1, \quad \alpha(3) = 3.$$

A slightly more convenient way to represent this permutation α is by the so-called array notation or Cauchy's two-line notation:

$$\alpha \longleftrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix},$$

where one lists the elements of the set S in the first row, and for each element its image under the permutation below it in the second row. Also the techniques of arrow diagrams and cycle-arrow diagrams are nice visual ways of representing the permutation (not explained here).

Definition. The "identity" permutation is the permutation that does nothing:

$$\alpha \leftrightarrow \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}.$$

Definition. An k-cycle is a permutation which cyclically permutes (rearranges) k elements of a set A while leaving the rest elements unchanged. For example ($k = 3$),

$$\begin{pmatrix} 1 & 2 & 3 & 4 & ... & n \\ & & & & & \\ 2 & 3 & 1 & 4 & ... & n \end{pmatrix}.$$

For the first three numbers every number moves to the right and the third number, $k = 3$, cycles around back to 1.

An example of a 2-cycle is a "transposition", which is an exchange of two elements of a set A with all others staying the same. In other words a transposition is a permutation of two elements.

Definition. The set of all permutations of n elements is called the "the symmetric group of degree n", and is denoted by S_n. For clarity, we can write this definition as $S_n = \{\alpha : \alpha$ is a permutation of $\mathbb{Z}_n\}$. Some authors use notation $Sym(n)$ instead of S_n.

One of the basic properties of permutations is that every permutation can be written as a product of "disjoint cycles", that is where the various cycles have no numbers in common. The algorithm to determine the cycle form of the permutation can be illustrated in the following example. Suppose we need to determine the cycle form of the permutation

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ & & & & & & & & & \\ 5 & 1 & 6 & 8 & 4 & 10 & 7 & 2 & 9 & 3 \end{pmatrix}.$$

We start with the smallest number in the set, in our case it is 1. Since $\alpha(1) = 5$ we begin the cycle by writing $(1, 5, \ldots) \ldots$, then as $\alpha(5) = 4$, so we continue building the cycle as $(1, 5, 4, \ldots) \ldots$ until we reach the mapping $\alpha(2) = 1$, which closes the first cycle $(1, 5, 4, 8, 2) \ldots$. Now we need to find the smallest number that doesn't appear in any previously constructed cycle, number 3 in our

example. In a similar fashion, we build the second cycle $(1,5,4,8,2)(3,6,10)\ldots$. Now we pick the smallest number that is not contained in any previously constructed cycle, 7 in our example, and as it maps to itself, we get $(1,5,4,8,2)(3,6,10)(7)\ldots$. The only remaining number is 9, and as it maps to itself, we get the following cycle form of permutation α: $(1,5,4,8,2)(3,6,10)(7)(9)$. Following the usual convention, we omit the 1-cycles, and simply write α as $(1,5,4,8,2)(3,6,10)$. Thus, in our case, α is the product of a 3-cycle and a 5-cycle.

After this brief review of the theory of permutation operators, we will explain the derivation of the formula for the squared state fidelity $\overline{F_{\rm st}^2}$.

We begin by writing the quantum operation as $\mathcal{E} = \mathcal{U} \circ \tilde{\mathcal{E}}$ [see Eq. (B.2)], where $\mathcal{U}$ corresponds to the ideal (desired) unitary operation, while the map $\tilde{\mathcal{E}}$ accounts for the errors in the actual gate, and $\circ$ denotes composition. Let

$$\tilde{\mathcal{E}}(\rho) = \sum_n A_n \rho A_n^\dagger \tag{6.3}$$

be the operator-sum representation of $\tilde{\mathcal{E}}$, where $\{A_n\}_{n=1}^{d^2}$ are Kraus operators satisfying the trace-preservation condition $\sum_n A_n^\dagger A_n = \mathbb{I}$. The Kraus operators can be easily obtained from the process matrix $\chi_{\alpha\beta}$ describing the operation $\mathcal{E}$. Note that by diagonalizing χ, i.e., $\chi = VDV^\dagger$, where V is unitary and $D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots)$ with $\lambda_n \geq 0$, we can express the Kraus operators in Eq. (6.3) as $A_n = \sqrt{\lambda_n}\, U^\dagger \sum_\alpha E_\alpha V_{\alpha n}$, where U is the desired unitary.

The state fidelity F_ϕ (assuming a pure initial state $|\phi\rangle$) can be written in terms of Kraus operators $\{A_n\}$ as follows:

$$F_\phi \equiv \langle\phi|\, \tilde{\mathcal{E}}(\phi)\, |\phi\rangle = \sum_n \langle\phi|\, A_n\, |\phi\rangle\, \langle\phi|\, A_n^\dagger\, |\phi\rangle\,. \tag{6.4}$$

Obviously, this expression depends on the state $|\phi\rangle$. Using the identity $\mathrm{Tr}(A \otimes B) = \mathrm{Tr}(A)\,\mathrm{Tr}(B)$,

one can rewrite the above expression for F_ϕ as

$$F_\phi = \sum_n \mathrm{Tr}\left[A_n \left|\phi\right\rangle \left\langle\phi\right|\right] \mathrm{Tr}\left[A_n^\dagger \left|\phi\right\rangle \left\langle\phi\right|\right] = \sum_n \mathrm{Tr}\left[(A_n \otimes A_n^\dagger)(\left|\phi\right\rangle \left\langle\phi\right|^{\otimes 2})\right], \tag{6.5}$$

where the notation $\left|\phi\right\rangle \left\langle\phi\right|^{\otimes k} \equiv \underbrace{\left|\phi\right\rangle \left\langle\phi\right| \otimes \left|\phi\right\rangle \left\langle\phi\right| \ldots \otimes \left|\phi\right\rangle \left\langle\phi\right|}_{k}$ means that the state is copied in k identical Hilbert spaces (k equals 2 in formulae for the state fidelity). The expression Eq. (6.5) is an inner product between a term including all Kraus operators and a term including all the state-dependence. Similarly, one can express the squared state fidelity as

$$\begin{aligned} F_\phi^2 &= \sum_{n,m} \left\langle\phi\right| A_n \left|\phi\right\rangle \left\langle\phi\right| A_n^\dagger \left|\phi\right\rangle \left\langle\phi\right| A_m \left|\phi\right\rangle \left\langle\phi\right| A_m^\dagger \left|\phi\right\rangle \\ &= \sum_{n,m} \mathrm{Tr}\left[(A_n \otimes A_n^\dagger \otimes A_m \otimes A_m^\dagger)(\left|\phi\right\rangle \left\langle\phi\right|^{\otimes 4})\right]. \end{aligned} \tag{6.6}$$

The average state fidelity $\overline{F_{\mathrm{st}}}$ of a quantum operation $\tilde{\mathcal{E}}$ is defined as follows:

$$\overline{F_{\mathrm{st}}} \equiv \int d\phi \left\langle\phi\right| \tilde{\mathcal{E}}(\phi) \left|\phi\right\rangle = \int F_\phi \, d\phi, \tag{6.7}$$

where the integral is over the uniform (Haar) measure $d\phi$ on the initial pure states, normalized so $\int d\phi = 1$. Obviously, the averaging over the states $\left|\phi\right\rangle$ should be performed on the second term in Eq. (6.5) only, therefore we need to calculate the average of the type $\overline{\left|\phi\right\rangle\left\langle\phi\right| \otimes \left|\phi\right\rangle\left\langle\phi\right|}$.

In order to compute the average state fidelity $\overline{F_{\mathrm{st}}} = \int F_\phi \, d\phi$, the average square of the state fidelity

$$\overline{F_{\mathrm{st}}^2} = \int F_\phi^2 \, d\phi, \tag{6.8}$$

and higher moments of F_{st} (we assume the normalized integration over the initial pure states, $\int d\phi = 1$), one can use the following result [163]

$$\int \left|\phi\right\rangle \left\langle\phi\right|^{\otimes k} d\phi = \frac{1}{\binom{k+d-1}{d-1}} \Pi_k, \quad \Pi_k \equiv \frac{1}{k!} \sum_{\sigma \in S_k} P_\sigma. \tag{6.9}$$

Here σ is an element of the permutation group S_k (the $k!$ permutations of k objects), $\binom{k+d-1}{d-1} = \frac{d(d+1)(d+2)...(d+k-1)}{k(k-1)...1}$ is a binomial coefficient "$(k+d-1)$ choose $(d-1)$", Π_k is the projector onto the symmetric subspace of $\mathcal{H}^{\otimes k}$, and the operator P_σ is the representation of permutation σ in $\mathcal{H}^{\otimes k} = \underbrace{\mathcal{H} \otimes \ldots \mathcal{H}}_{k}$, i.e.,

$$P_\sigma(|\phi_1\rangle \otimes |\phi_2\rangle \ldots \otimes |\phi_k\rangle) = |\phi_{\sigma(1)}\rangle \otimes |\phi_{\sigma(2)}\rangle \ldots \otimes |\phi_{\sigma(k)}\rangle \,. \tag{6.10}$$

(The operator P_σ acts on the wavefunction of kN qubits by permuting k blocks, each containing N qubits.)

In view of the above discussion, we see that the kth moment $\overline{F_{\rm st}^k} \equiv \int F_\phi^k \, d\phi$ can be expressed as a sum of $(2k)!$ terms corresponding to the elements in S_{2k} [note that k in Eqs. (6.9) and (6.10) is now replaced with $2k$],

$$\overline{F_{\rm st}^k} = \frac{\displaystyle\sum_{n_1 \ldots n_k} \sum_{\sigma \in S_{2k}} \mathrm{Tr}[(A_{n_1} \otimes A_{n_1}^\dagger \otimes \ldots A_{n_k} \otimes A_{n_k}^\dagger) P_\sigma]}{\binom{2k+d-1}{d-1}(2k)!}. \tag{6.11}$$

For example, the average state fidelity $\overline{F_{\rm st}}$ is determined by the sum over S_2,

$$\begin{aligned} \mathrm{Tr}(A_n \otimes A_n^\dagger \, \Pi_2) &= \frac{1}{2} \sum_{\sigma \in S_2} \mathrm{Tr}(A_n \otimes A_n^\dagger P_\sigma) \\ &= \frac{1}{2} \sum_{\sigma \in S_2} \sum_{i_1, i_2} \langle i_1, i_2 | \, A_n \otimes A_n^\dagger \, |\sigma(i_1), \sigma(i_2)\rangle \\ &= \frac{1}{2} \Big[\underbrace{\mathrm{Tr}(A_n)\,\mathrm{Tr}(A_n^\dagger)}_{\text{identity}} + \underbrace{\mathrm{Tr}(A_n A_n^\dagger)}_{\text{transposition}} \Big], \end{aligned} \tag{6.12}$$

which yields the well-known result [160]

$$\overline{F_{\rm st}} = \frac{1}{d(d+1)} \left(\sum_n |\,\mathrm{Tr}(A_n)|^2 + d \right). \tag{6.13}$$

In order to express the average square of the state fidelity $\overline{F_{\rm st}^2}$ in terms of Kraus operators, it is convenient to write each element of the group S_4 as a product of disjoint cycles. Using the so-

called cycle notation for permutations, the 24 elements of the permutation groups S_4 can be grouped as follows:

• Identity (1 element): (1)(2)(3)(4) (this notation means that no change of position occurs for all numbers in the sequence 1234);

• Transpositions (6 elements): (12), (13), (14), (23), (24), and (34) (this notations means that only the specified numbers in the sequence are exchanged):

$$(12) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 2 & 1 & 3 & 4 \end{pmatrix}, \quad (13) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 3 & 2 & 1 & 4 \end{pmatrix}, \quad (14) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 4 & 2 & 3 & 1 \end{pmatrix},$$

$$(23) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 1 & 3 & 2 & 4 \end{pmatrix}, \quad (24) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 1 & 4 & 3 & 2 \end{pmatrix}, \quad (34) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 1 & 2 & 4 & 3 \end{pmatrix};$$

$$(23) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 1 & 3 & 2 & 4 \end{pmatrix}, \quad (24) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 1 & 4 & 3 & 2 \end{pmatrix}, \quad (34) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 1 & 2 & 4 & 3 \end{pmatrix};$$

• 3-cycles (8 elements): (123), (132), (124), (142), (134), (143), (234), and (243) [here the notation (123) means the permutation 1→2→3→1, while the remaining number does not change]:

$$(123) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 2 & 3 & 1 & 4 \end{pmatrix} = (123)(4), \quad (132) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 3 & 1 & 2 & 4 \end{pmatrix} = (132)(4),$$

$$(124) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 2 & 4 & 3 & 1 \end{pmatrix} = (124)(3), \quad (142) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 4 & 1 & 3 & 2 \end{pmatrix} = (142)(3),$$

$$(134) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 3 & 2 & 4 & 1 \end{pmatrix} = (134)(2), \quad (143) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 4 & 2 & 1 & 3 \end{pmatrix} = (143)(2),$$

$$(234) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 1 & 3 & 4 & 2 \end{pmatrix} = (234)(1), \quad (243) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 1 & 4 & 2 & 3 \end{pmatrix} = (243)(1);$$

• Products of transpositions (3 elements): (12)(34), (13)(24), and (14)(23) (two pairs of numbers exchange):

$$(12)(34) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 2 & 1 & 4 & 3 \end{pmatrix}, \quad (13)(24) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 3 & 4 & 1 & 2 \end{pmatrix}, \quad (14)(23) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 4 & 3 & 2 & 1 \end{pmatrix};$$

• 4-cycles (6 elements): (1234), (1243), (1324), (1342), (1423), and (1432) [here (1234) means the permutation 1→2→3→4→1]:

$$(1234) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 2 & 3 & 4 & 1 \end{pmatrix}, \quad (1243) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 2 & 4 & 1 & 3 \end{pmatrix}, \quad (1324) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 3 & 4 & 2 & 1 \end{pmatrix},$$

$$(1342) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 3 & 1 & 4 & 2 \end{pmatrix}, \quad (1423) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 4 & 3 & 1 & 2 \end{pmatrix}, \quad (1432) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & & & \\ 4 & 1 & 2 & 3 \end{pmatrix}.$$

This classification simplifies keeping track of the terms $N_\sigma \equiv \sum_{n,m} \mathrm{Tr}\left[\left(A_n \otimes A_n^\dagger \otimes A_m \otimes\right.\right.$

$A_m^\dagger)P_\sigma\big]$ in Eq. (6.11). The corresponding contributions to the sum $\sum_{\sigma\in S_4} N_\sigma$ are the following:

Identity:

$$\big(\sum_n |\operatorname{Tr}(A_n)|^2\big)^2.$$

Transpositions:

$$2d\sum_n |\operatorname{Tr}(A_n)|^2 + 2\sum_{n,m}\operatorname{Tr}(A_nA_m^\dagger)\operatorname{Tr}(A_n^\dagger)\operatorname{Tr}(A_m)$$
$$+\sum_{n,m}\Big[\operatorname{Tr}(A_nA_m)\operatorname{Tr}(A_n^\dagger)\operatorname{Tr}(A_m^\dagger)+\operatorname{Tr}(A_n^\dagger A_m^\dagger)\operatorname{Tr}(A_n)\operatorname{Tr}(A_m)\Big].$$

3-cycles:

$$4\sum_n |\operatorname{Tr}(A_n)|^2 + 2\sum_{n,m}\Big[\operatorname{Tr}(A_nA_n^\dagger A_m)\operatorname{Tr}(A_m^\dagger)+\operatorname{Tr}(A_nA_n^\dagger A_m^\dagger)\operatorname{Tr}(A_m)\Big].$$

Products of transpositions:

$$d^2+\sum_{n,m}\Big[|\operatorname{Tr}(A_nA_m)|^2+|\operatorname{Tr}(A_nA_m^\dagger)|^2\Big].$$

4-cycles:

$$3d+\sum_{n,m}\operatorname{Tr}(A_nA_n^\dagger A_mA_m^\dagger)+2\sum_{n,m}\operatorname{Tr}(A_nA_mA_n^\dagger A_m^\dagger).$$

(We used the trace-preservation condition $\sum_n A_n^\dagger A_n=\mathbb{I}$). Substituting the above terms in Eq. (6.11)

(with $k = 2$), we finally obtain the average square of the state fidelity,

$$\begin{aligned}\overline{F_{\rm st}^2} = \frac{1}{d(d+1)(d+2)(d+3)}\Big\{ & d^2 + 3d \\ & +2(d+2)\sum_n |\operatorname{Tr}(A_n)|^2 + \Big[\sum_n |\operatorname{Tr}(A_n)|^2\Big]^2 \\ & +\sum_{n,m}\left(|\operatorname{Tr}(A_n A_m)|^2 + |\operatorname{Tr}(A_n A_m^\dagger)|^2\right) \\ & +2\sum_{n,m}\operatorname{Tr}(A_n A_m A_n^\dagger A_m^\dagger) + \sum_{n,m}\operatorname{Tr}(A_n A_n^\dagger A_m A_m^\dagger) \\ & +2\sum_{n,m}\operatorname{Tr}(A_n A_m^\dagger)\operatorname{Tr}(A_n^\dagger)\operatorname{Tr}(A_m) \\ & +2\sum_{n,m}\operatorname{Re}[\operatorname{Tr}(A_n A_m)\operatorname{Tr}(A_n^\dagger)\operatorname{Tr}(A_m^\dagger)] \\ & +4\sum_{n,m}\operatorname{Re}[\operatorname{Tr}(A_n A_n^\dagger A_m^\dagger)\operatorname{Tr}(A_m)]\Big\}. \end{aligned} \tag{6.14}$$

This is the formula we used in Section 6.1 to calculate $\overline{F_{\rm st}^2}$.

6.3 Monte Carlo numerical calculations

In order to verify the correctness of the formula Eq. (6.14), we have performed numerical calculations of the average state fidelity $\overline{F_{\rm st}}$ and the average square of the state fidelity $\overline{F_{\rm st}^2}$, defined in Eq. (6.7) and Eq. (6.8), using a Monte Carlo integration method. The idea of Monte Carlo integration is to numerically evaluate the integrals Eq. (6.7) and Eq. (6.8) over the uniform Haar measure $d\phi$ by evaluating the discrete versions of the integrals for $\overline{F_{\rm st}}$ and $\overline{F_{\rm st}^2}$ with randomly chosen points. For this purpose we calculate in a sum over index r the values of the state fidelity and its square for randomly generated quantum states $\{|\Phi_r\rangle\} = \{|\Phi_1\rangle, |\Phi_2\rangle, \ldots, |\Phi_R\rangle\}$, uniformly distributed on a surface of a multidimensional sphere with a unit radius, and then average the results. Our implementation of the algorithm for an N-qubit system is the following. 1) Using random numbers from a uniform distribution, generate a random N-qubit quantum state $|\Phi_r\rangle$ on the surface

of the (2×2^N)–dimensional sphere of unit radius, 2) at each iteration of r calculate the propagation of the state $|\Phi_r\rangle$ through the quantum channel described by a process matrix χ, and 3) calculate the quantum process characteristics such as the state fidelity $F_{\phi,\mathrm{r}}$ and the square of the state fidelity $F^2_{\phi,\mathrm{r}}$ for each random state $|\Phi_r\rangle$. After averaging over r we obtain the numerical results for the average state fidelity $\overline{F_{\mathrm{st}}}$ and for the average square of the state fidelity $\overline{F^2_{\mathrm{st}}}$. We observe that in the limit of a large number R of random states $|\Phi_r\rangle$, used in the procedure, our numerical results for the average characteristics asymptotically approach the values calculated using the analytic expressions Eq. (6.13) and Eq. (6.14). We have for the state fidelity

$$\overline{F}_R = \frac{1}{R}\sum_{r=1}^{R} F(\Phi_r) \xrightarrow[R\to\infty]{} \overline{F} = \int d\phi F(\phi),$$

and for the square of the state fidelity

$$\overline{F^2}_R = \frac{1}{R}\sum_{r=1}^{R} F^2(\Phi_r) \xrightarrow[R\to\infty]{} \overline{F^2} = \int d\phi F^2(\phi).$$

These numerical results also make it possible to evaluate numerically the standard deviation of the state fidelity. We perform Monte Carlo numerical simulations for two-qubit and three-qubit quantum channels, which were considered in Chapter 5.

Below follows a more detailed description of our implementation of the Monte Carlo algorithm. The steps of the algorithm at each iteration of the sum over r are the following.

1. First, we generate a random N-qubit quantum state (a vector) $|\Phi_r\rangle$, which is described by (2×2^N) independent real coefficients in the computational basis of 2^N basis states. For example,

for the system of two qubits ($N = 2$) the computational basis consists of 4 states

$$|00\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \ |01\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \ |10\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \ |11\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

and the two-qubit quantum state is described by 8 real coefficients, corresponding to 4 complex numbers, as follows:

$$|\Phi_r\rangle = (\alpha_{00} + i\beta_{00})\,|00\rangle + (\alpha_{01} + i\beta_{01})\,|01\rangle + (\alpha_{10} + i\beta_{10})\,|10\rangle + (\alpha_{11} + i\beta_{11})\,|11\rangle\,, \quad (6.15)$$

where the eight coefficients $\alpha_{k,m}$ and $\beta_{k,m}$ ($k, m = 0, 1$) satisfy the normalization condition

$$\sum_{k,m=0,1} |\alpha_{k,m}|^2 + \sum_{k,m=0,1} |\beta_{k,m}|^2 = 1. \quad (6.16)$$

Eq. (6.16) is the equation of a (2×2^N)–dimensional sphere of unit radius (8-dimensional sphere for two qubits or 16-dimensional sphere for three qubits), centered at the origin of coordinates. Therefore, the task of generating a uniformly distributed random N-qubit quantum state is equivalent to the task of generating a uniformly distributed point on a surface of a (2×2^N)–dimensional sphere of unit radius.

To generate a uniformly distributed random point on the surface of a multidimensional sphere, we use the following acceptance-rejection algorithm [164, 165]. We first generate a random point within a (2×2^N)–dimensional hypercube, the center of which is at the origin of coordinates, and the length of each of its sides equals 2. For this purpose we generate (2×2^N) random numbers (denote them as $x_{k,m}$ and $y_{k,m}$, where $k, m = 0, 1$) uniformly distributed on the interval $[-1, +1]$ (for example, generate 8 random numbers for a two-qubit system). These (2×2^N) random numbers

determine a point inside a (2×2^N)-dimensional cube. We then reject the points that are within this (2×2^N)–dimensional cube, but outside the (2×2^N)–dimensional sphere of unit radius, keeping only the points that are within the multidimensional sphere. We need this step in order to ensure that the distribution of random points on the surface of the multidimensional sphere is uniform. Also in order to reduce the possibility of the numerical (precision) errors, we discard the sets of randomly generated points that happen to be in the immediate vicinity of the origin. Hence we check if the following conditions

$$\eta \leq \left(\sum_{k,m} |x_{k,m}|^2 + \sum_{k,m} |y_{k,m}|^2 \right) \leq 1 \tag{6.17}$$

are satisfied (most of the time the "right" condition in Eq. (6.17) is violated). The parameter η is a small cut-off parameter, its value can be set arbitrarily. For our simulation we used $\eta = 0.01$. In the rare case when the condition Eq. (6.17) is satisfied, we rescale the numbers $x_{k,m}$ and $y_{k,m}$, that is, we compute

$$\alpha_{k,m} = \frac{x_{k,m}}{\sqrt{\left[\sum_{k,m} |x_{k,m}|^2 + \sum_{k,m} |y_{k,m}|^2 \right]}}, \quad \beta_{k,m} = \frac{y_{k,m}}{\sqrt{\left[\sum_{k,m} |x_{k,m}|^2 + \sum_{k,m} |y_{k,m}|^2 \right]}},$$

which ensures that the vector with (2×2^N) components $\alpha_{k,m}$ and $\beta_{k,m}$ lies on the surface of the multidimensional unit sphere, defined by Eq. (6.16).

Once we have, successfully, generated a random N-qubit state $|\Phi_r\rangle$, we can compute $F_{\rm st}(|\Phi\rangle)$ and $F^2_{\rm st}(|\Phi\rangle)$.

It is not difficult to estimate the efficiency of the described algorithm by estimating the ratio of "successful" random point generation events, when the points happen to be inside of the (2×2^N)–dimensional ball, to the total number of randomly generated sets of points inside of a (2×2^N)–dimensional cube. This ratio is equal to the ratio of the volume of the (2×2^N)–

dimensional ball with radius R, $V_B = \dfrac{\pi^{[2^N]} R^{[2\times 2^N]}}{(2^N)!}$, to the volume of the (2×2^N)–dimensional cube with the length of its side $2R$, $V_C = (2R)^{[2\times 2^N]}$. In particular, for the two-qubit system ($N = 2$) we have the ratio of volumes of an 8-dimensional ball and an 8-dimensional cube equal to $\dfrac{V_B}{V_C} = \dfrac{\pi^4 R^8}{2^8\, 4!\, R^8} = 0.016$, which is not a very small number, so this algorithm is relatively efficient. Roughly speaking, on average, we succeed in generating a random 2-qubit state after one hundred attempts. In case of a three-qubit system ($N = 3$), this ratio of volumes of a 16-dimensional ball and cube is $\dfrac{V_B}{V_C} = \dfrac{\pi^8 R^{16}}{2^{16}\, 8!\, R^{16}} = 3.6 \times 10^{-6}$, and the algorithm still works, although not as fast as for the two-qubit situation: we succeed only one time in approximately one million attempts.

2. After we generated the random input pure state $|\Phi_r\rangle$ described by Eq. (6.15), and its density matrix $|\Phi_r\rangle\langle\Phi_r|$, we calculate the corresponding output "experimental" density matrix $\rho_r^{\rm exp}$ at the output of the quantum gate by using the standard definition Eq. (4.4) of Chapter 4 of the evolution of a quantum state propagating through the quantum gate described by a process matrix χ. While performing this calculation, we substitute into Eq. (4.4) the process matrix χ calculated either from the full data by the Least-Squares method, or from the partial data by the compressed sensing method. Also, using the same randomly generated input state $|\Phi_r\rangle$, we easily calculate the "ideal" output density matrix $\rho_r^{\rm ideal}$, corresponding to the unitary evolution of the input state $|\Phi_r\rangle$ described by Eq. (2.18) of Chapter 2.

3. Using the results for $\rho_r^{\rm exp}$ and $\rho_r^{\rm ideal}$, calculated at the $r^{\rm th}$ iteration of the loop over various randomly generated states, we calculate the statistical characteristics of the channel after r iterations of the loop in a cumulative way. We define the cumulative variables $F_{\phi,{\rm r}}$ and $F^2_{\phi,{\rm r}}$ (which are initialized to zero before the iterative loop over r), as follows: $F_{\phi,{\rm r}} = F_{\phi,{\rm r}-1} + {\rm Tr}(\rho_r^{\rm exp}\rho_r^{\rm ideal})$ and $F^2_{\phi,{\rm r}} = F^2_{\phi,{\rm r}-1} + [{\rm Tr}(\rho_r^{\rm exp}\rho_r^{\rm ideal})]^2$. Then the average state fidelity $\overline{F_{\rm st}}(r)$ after r realizations can

be numerically calculated as

$$\overline{F_{\rm st}}(r) = F_{\phi,{\rm r}}/r, \tag{6.18}$$

the average square of state fidelity $\overline{F_{\rm st}^2}(r)$ after r realizations as

$$\overline{F_{\rm st}^2}(r) = F_{\phi,{\rm r}}^2/r, \tag{6.19}$$

and the standard deviation of the fidelity $\Delta F(r)$ after r realizations as

$$\Delta F(r) = \sqrt{\frac{F_{\phi,{\rm r}}^2 - r(\overline{F_{\rm st}}(r))^2}{r}}. \tag{6.20}$$

These results Eqs. (6.18) – (6.20), obtained by Monte - Carlo numerical simulations, lead to the numerical values of the statistical characteristics of the quantum channel (such as the average state fidelity $\overline{F_{\rm st}}$, the average square of state fidelity $\overline{F_{\rm st}^2}$, and the standard deviation of the state fidelity ΔF), which are in a good agreement with the analytical results obtained using the exact formulae in Eq. (6.13) and Eq. (6.14), for both two and three qubits.

To demonstrate this agreement between the numerical and analytical results, consider for simplicity the two-qubit case. We show that the difference between the numerical results for the average state fidelity $\overline{F_{\rm st}}(r)$ or its average square $\overline{F_{\rm st}^2}(r)$ and their analytic values $\overline{F_{\rm st}}$ or $\overline{F_{\rm st}^2}$ approaches zero, as the number r of random states used in the Monte-Carlo procedure increases. As the values of $\overline{F_{\rm st}}(r)$ or $\overline{F_{\rm st}^2}(r)$ can fluctuate slightly for different randomly generated states used, we repeat the Monte-Carlo procedure $K = 100$ times for every specific number r, and denote as $\overline{F_{\rm st,k}}(r)$ and $\overline{F_{\rm st,k}^2}(r)$ the values for the statistical characteristics obtained at each of such repetitions (index k changes from 1 to 100). Now we introduce the following two functions $\sigma_{_F}(r)$ and $\sigma_{_{F^2}}(r)$,

$$\sigma_{_F}(r) = \sqrt{\frac{1}{K}\sum_{k=1}^{K}\big(\overline{F_{\rm st,k}}(r) - \overline{F_{\rm st}}\big)^2}, \tag{6.21}$$

and

$$\sigma_{F^2}(r) = \sqrt{\frac{1}{K}\sum_{k=1}^{K}\left(\overline{F^2_{\mathrm{st,k}}}(r) - \overline{F^2_{\mathrm{st}}}\right)^2}, \tag{6.22}$$

which have the meaning of the standard deviation of the numerical values for $\overline{F_{\mathrm{st}}}(r)$ or $\overline{F^2_{\mathrm{st}}}(r)$ from their analytical values $\overline{F_{\mathrm{st}}}$ or $\overline{F^2_{\mathrm{st}}}$.

We show in Fig. 6.3 and Fig. 6.4 the dependence of these functions $\sigma_F(r)$ and $\sigma_{F^2}(r)$ on the number of random states r used in the Monte-Carlo procedure, plotted in a log-log scale (the number r varies from 1 to 10^5). We see that as we increase the number of random states r used in the Monte-Carlo procedure, these functions $\sigma_F(r)$ and $\sigma_{F^2}(r)$ approach zero. This means that the results of the numerical estimation of the average state fidelity by Eq. (6.18) and the average square of state fidelity by Eq. (6.19) asymptotically approach the corresponding results obtained from the exact analytic formulae Eq. (6.13) and Eq. (6.14). Also, as we see, these functions $\sigma_F(r)$ and $\sigma_{F^2}(r)$ have the $1/\sqrt{r}$ dependence, which corresponds to the straight line on the log-log plot (shown by the green dashed lines). The presented Fig. 6.3 and Fig. 6.4 have been obtained using the experimental data for the two-qubit CZ gate realized with the phase qubits.

These results confirm the correctness of the analytic formulae Eq. (6.13) and Eq. (6.14), which have been used in our analytic calculations of the average state fidelity, the average square of the state fidelity, and the standard deviation of the fidelity.

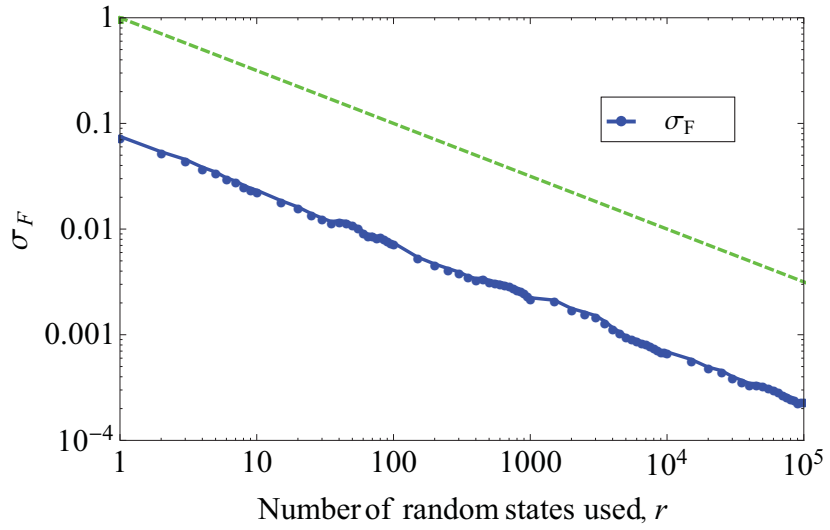

Figure 6.3: Verification of the formula for the average state fidelity by the Monte-Carlo method. The blue curve shows the function $\sigma_F(r)$ defined in Eq. (6.21), for the various number r of the random states used in the Monte-Carlo procedure. For comparison, the green dashed line shows the $1/\sqrt{r}$ dependence. Experimental data for the two-qubit CZ gate realized with the phase qubits have been used.

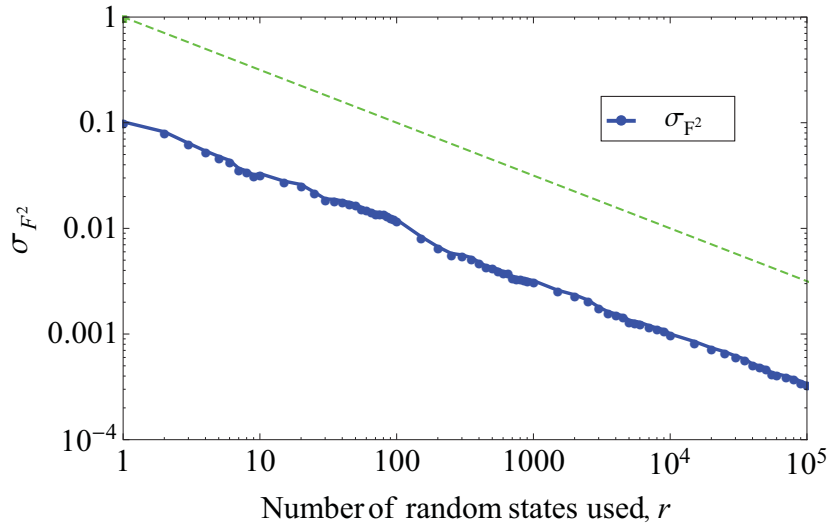

Figure 6.4: Verification of the formula for the average square of the state fidelity by Monte-Carlo method. The blue curve shows the function $\sigma_{F^2}(r)$ defined in Eq. (6.22), for the various number r of the random states used in the Monte-Carlo procedure. For comparison, the green dashed line shows the $1/\sqrt{r}$ dependence.

Chapter 7

Conclusion

In this dissertation we explored quantum process tomography (QPT), which is a technique for fully characterizing a quantum operation. We first presented results for the analysis of the two-qubit and multi-qubit quantum gates based on superconducting qubits using three distinct methods of standard quantum process tomography: linear inversion, maximum likelihood, and least-squares. A well-known problem of standard QPT is the exponential scaling of the resources, i.e., the number of the required experimental configurations needed to characterize the evolution of a multi-qubit system increases exponentially with the number of qubits. We showed that the method of compressed sensing quantum process tomography (CS QPT), applied to the two-qubit and three-qubit quantum gates based on superconducting qubits, offers a significant reduction of the needed amount of experimental data.

In Chapter 2, we gave a brief review of the main concepts of quantum information and different types of superconducting qubits. In Chapter 3, we presented the results of the state estimation using two methods for quantum state tomography (QST), which is the procedure of experimentally

determining all the elements of the density matrix of an unknown quantum state. We wrote Matlab code for the estimation of the density matrices of a system of two superconducting phase qubits by the method of direct linear inversion, as well as by the method of maximum likelihood, and we presented the resulting density matrices and their comparison. While the method of linear inversion may lead to nonphysical artifacts, such as negative probabilities, the maximum likelihood method guarantees the density matrix to be theoretically valid while giving the closest fit to the measured experimental data.

In Chapter 4, we gave an overall introduction to the quantum process tomography, discussed the experimental details of QPT with superconducting qubits, and presented the results of the estimation for the process matrix by the method of linear inversion. We discussed computational resources required for the implementation of the linear inversion for various numbers of qubits, and we concluded that the maximum number of qubits in the system that allows the exact computation of the process matrix on an average personal computer is six. We also discussed that as the constraints for the process matrix to be physical were not imposed in the linear inversion method, the resulting process matrix may be nonphysical. The methods of maximum likelihood or least-squares, formulated as the convex optimization problems, result in a legitimate physical process matrix. We discussed both these methods, and we presented the results of the two-qubit process matrix calculations by the least-squares method. The method of the least-squares was used in the subsequent Chapters 5 and 6 to obtain a fully physical process matrix, against which the results of the CS QPT process matrix reconstruction were benchmarked.

In Chapter 5, we numerically analyzed the efficiency of CS QPT applied to superconducting qubits. We used experimental data for two-qubit controlled-Z (CZ) gates realized with Xmon

and phase qubits, and also used simulated data for the three-qubit Toffoli gate with numerically added noise. We showed that CS QPT permits a reasonably high fidelity estimation of the process matrix from a substantially reduced data set compared to the full QPT. In particular, for the CZ gate (Fig. 5.4) the amount of data can be reduced by a factor of $\sim$7 compared to the full QPT (which is a factor of $\sim$4 compared to the threshold of underdetermined system of equations). For the Toffoli gate (Fig. 5.17) the data reduction factor is $\sim$40 compared to the full QPT ($\sim$15 compared to the threshold of underdeterminacy).

We primarily used two fidelity characteristics in our analysis: first, the fidelity $F(\chi_{\rm CS}, \chi_{\rm full})$ of the CS QPT-estimated process matrix $\chi_{\rm CS}$ compared with the matrix $\chi_{\rm full}$ calculated from the full data set, and second, the fidelity $F(\chi_{\rm CS}, \chi_{\rm ideal})$ of $\chi_{\rm CS}$ compared with the ideal process matrix $\chi_{\rm ideal}$. Besides these two characteristics, we also calculated in Chapter 6 the standard deviation of the average state fidelity $\Delta F_{\rm st}$. We established that the CS method estimates $\Delta F_{\rm st}$ from a reduced data set (Figs. 6.1 and 6.2) with a high accuracy. The dependence of the standard deviation of the average state fidelity $\Delta F_{\rm st}$ on the number of measurement configurations was evaluated both analytically and numerically using a Monte Carlo simulation technique. In order to obtain an analytic expression for the standard deviation of the state fidelity $\Delta F_{\rm st}$, a detailed derivation of the formula for calculating the second moment $\overline{F_{\rm st}^2}$ of the average state fidelity $\overline{F_{\rm st}}$ for the CS QPT-estimated process matrix $\chi_{\rm CS}$ compared with $\chi_{\rm ideal}$ was presented.

We also showed that the results of the compressed sensing method depend on the choice of the basis, in which the process matrix should be approximately sparse. We used two bases in this work: the Pauli-error basis and the singular value decomposition (SVD) basis. We found that the results obtained in both bases are similar to each other, though the SVD basis required fewer computational resources.

We also performed the comparison of the CS method with the least squares (LS) optimization. We showed that even though the LS method formally works, it gives a less accurate estimate of the process matrix χ than the CS method in the significantly underdetermined regime (although it does give a better estimate in the overdetermined regime). The advantage of the CS method over the LS method was more pronounced for the Toffoli gate (Fig. 5.18).

Overall, several different methods of performing quantum process tomography for the quantum gates based on superconducting qubits were presented in this dissertation, and in particular we showed that the compressed sensing method of QPT offers efficient estimation of process matrices of superconducting two-qubit and three-qubit logic gates.

Appendices

A Appendix A. List of publications and presentations by Andrey Rodionov

Journal Papers

1. Andrey V. Rodionov, Andrzej Veitia, R. Barends, J. Kelly, Daniel Sank, J. Wenner, John M. Martinis, Robert L. Kosut, and Alexander N. Korotkov, "Compressed sensing quantum process tomography for superconducting quantum gates", Physical Review B, vol. **90**, article 144504, Oct. 2014. (This dissertation is based mostly on this paper.)

2. A. O′Toole, F. E. Peña Arellano, A. V. Rodionov, M. Shaner, E. Sobacchi, V. Dergachev, R. DeSalvo, M. Asadoor, A. Bhawal, P. Gong, C. Kim, A. Lottarini, Y. Minenkov, and C. Murphy, "Design and initial characterization of a compact, ultra high vacuum compatible, low frequency, tilt accelerometer", Review of Scientific Instruments, vol. **85**, article 075003, July 2014.

3. V. Dergachev, R. DeSalvo, M. Asadoor, A. Bhawal, P. Gong, C. Kim, A. Lottarini, Y. Minenkov, C. Murphy, A. O′Toole, F. E. Peña Arellano, A. V. Rodionov, M. Shaner, and E. Sobacchi, "A high

precision, compact electromechanical ground rotation sensor", Review of Scientific Instruments, vol. **85**, article 054502, May 2014.

4. A. V. Rodionov and A. S. Chirkin, "Entangled photon states in consecutive nonlinear optical interactions", JETP Letters, vol. **79**, no. 6, pp. 253 - 256, 2004 (Translated from Pis′ma v Zhurnal Eksperimental′noi i Teoreticheskoi Fiziki, vol. **79**, No. 6, 2004, pp. 311 - 314). Erratum: JETP Letters, vol. **79**, no. 11, p. 582, 2004.

5. A. V. Rodionov and A. S. Chirkin, "Photon statistics upon consecutive parametric interactions of light waves with nonmultiple frequencies", Optics and Spectroscopy, vol. **96**, no. 5, pp. 721 - 726, 2004 (Translated from Optika i Spektroskopiya, vol. **96**, No. 5, 2004, pp. 790 - 795).

6. A. S. Chirkin and A. V. Rodionov, "Quantum Zeno effect in two coupled nonlinear optical processes", Journal of Russian Laser Research, vol. **26**, issue 2, pp. 83 - 93, 2005.

Conference Presentations

Andrey V. Rodionov, Alexander N. Korotkov, Robert L. Kosut, Matteo Mariantoni, Daniel Sank, J. Wenner, and John M. Martinis, "Compressed Sensing Quantum Process Tomography of Superconducting Qubit Gates", *APS March Meetings* (Baltimore, MD, March 18 – 22, 2013), Bulletin of APS, vol. **58**, no. 1, abstract G26.003.

Posters

1. Andrey Rodionov, Alexander N. Korotkov, Robert L. Kosut, Matteo Mariantoni, Daniel Sank, James Wenner, and John M. Martinis, “Quantum Process Tomography of Superconducting Qubit Gates via Compressed Sensing”, Southwest Quantum Information and Technology (SQUINT), Fifteenth Annual Meeting, February 21 – 23, 2013, Santa Barbara, CA.

2. Andrey Rodionov, Alexander N. Korotkov, Robert L. Kosut, Matteo Mariantoni, Daniel Sank, James Wenner, and John M. Martinis, “Quantum Process Tomography of Superconducting Qubit Gates via Compressed Sensing”,
IARPA MQCO PI Meeting, January 30 – 31, 2013, San Diego, CA.

B Appendix B. Pauli-error basis

In this Appendix we discuss the definition of the Pauli-error basis used in this paper. The detailed theory of the QPT in the Pauli-error basis is presented in Ref. [18].

Let us start with description of a quantum process $\mathcal{E}$ in the Pauli basis $\{\mathcal{P}_\alpha\}$,

$$\rho^{\text{in}} \mapsto \mathcal{E}(\rho^{\text{in}}) = \sum_{\alpha,\beta=1}^{d^2} \chi_{\alpha\beta} \mathcal{P}_\alpha \rho^{\text{in}} \mathcal{P}_\beta^\dagger, \tag{B.1}$$

where for generality $\mathcal{P}$ is not necessarily Hermitian (to include the modified Pauli basis, in which $Y = -i\sigma_y$). Recall that $d = 2^N$ is the dimension of the Hilbert space for N qubits, and in N-qubit systems the elements of the Pauli basis $\{\mathcal{P}_\alpha\}$ are built as the Kronecker (tensor) products of N single-qubit Pauli operators from the set $\{\sigma_0, \sigma_x, \sigma_y, \sigma_z\}$, introduced in Eq. (2.12) of Section 2.1.3 of Chapter 2.

In order to compare the process $\mathcal{E}$ with a desired unitary rotation U [i.e. with the map $\mathcal{U}(\rho^{\text{in}}) = U\rho^{\text{in}}U^\dagger$], let us formally apply the inverse unitary $U^{-1} = U^\dagger$ after the process $\mathcal{E}$. The resulting composed process

$$\tilde{\mathcal{E}} = \mathcal{U}^{-1} \circ \mathcal{E} \tag{B.2}$$

characterizes the error: if $\mathcal{E}$ is close to the desired $\mathcal{U}$, then $\tilde{\mathcal{E}}$ is close to the identity (memory) operation. The process matrix $\tilde{\chi}$ of $\tilde{\mathcal{E}}$ in the Pauli basis is what we call in this paper the process matrix in the Pauli-error basis.

The process matrix $\tilde{\chi}$ obviously satisfies the relation

$$\sum_{\alpha,\beta} \tilde{\chi}_{\alpha\beta} \mathcal{P}_\alpha \rho^{\text{in}} \mathcal{P}_\beta^\dagger = U^{-1} \left(\sum_{\alpha,\beta} \chi_{\alpha\beta} \mathcal{P}_\alpha \rho^{\text{in}} \mathcal{P}_\beta^\dagger \right) U, \tag{B.3}$$

which can be rewritten as

$$\sum_{\alpha,\beta} \tilde{\chi}_{\alpha\beta} (U\mathcal{P}_\alpha) \rho^{\text{in}} (U\mathcal{P}_\beta)^\dagger = \sum_{\alpha,\beta} \chi_{\alpha\beta} \mathcal{P}_\alpha \rho^{\text{in}} \mathcal{P}_\beta^\dagger. \tag{B.4}$$

Therefore the error matrix $\tilde{\chi}$ is formally the process matrix of the original map $\mathcal{E}$, expressed in the operator basis

$$E_\alpha = U\mathcal{P}_\alpha. \tag{B.5}$$

This is the Pauli-error basis used in our paper. (Another obvious way to define the error basis is to use $E_\alpha = \mathcal{P}_\alpha U$ [18]; however, we do not use this second definition here.) The Pauli-error basis matrices E_α have the same normalization as the Pauli matrices,

$$\langle E_\alpha | E_\beta \rangle = \mathrm{Tr}(E_\alpha^\dagger E_\beta) = d\,\delta_{\alpha\beta}. \tag{B.6}$$

The matrices χ and $\tilde{\chi}$ (in the Pauli and Pauli-error bases) are related via unitary transformation,

$$\tilde{\chi} = V\chi V^\dagger, \;\; V_{\alpha\beta} = \mathrm{Tr}(\mathcal{P}_\alpha^\dagger U^\dagger \mathcal{P}_\beta)/d. \tag{B.7}$$

The matrix $\tilde{\chi}$ has a number of convenient properties [18]. It has only one large element, which is at the upper left corner and corresponds to the process fidelity, $\tilde{\chi}_{\mathcal{II}} = F_\chi = F(\chi, \chi_{\text{ideal}})$. All other non-zero elements of $\tilde{\chi}$ describe imperfections. In particular, the imaginary elements in the left column (or upper row) characterize unitary imperfections (assuming the standard non-modified Pauli basis), other off-diagonal elements are due to decoherence, and the diagonal elements correspond to the error probabilities in the Pauli-twirling approximation.

C Appendix C. Singular value decomposition (SVD) basis

The SVD basis used in this paper is introduced following Ref. [39]. Let us start with the so-called natural basis for $d \times d$ matrices, which consists of matrices E_α^{nat}, having one element equal to one, while other elements are zero. The numbering corresponds to the vectorized form obtained by stacking the columns: for $\alpha = (d-1)i + j$ the matrix is $(E_\alpha^{\text{nat}})_{lk} = \delta_{il}\delta_{jk}$. For a desired unitary rotation U, the process matrix χ^{nat} in the natural basis can be obtained by expanding U in the natural basis, $U = \sum_\alpha u_\alpha E_\alpha^{\text{nat}}$, and then constructing the outer product,

$$\chi_{\alpha\beta}^{\text{nat}} = u_\alpha u_\beta^*. \tag{C.1}$$

For example, for the ideal CZ gate the components u_α are $(1,0,0,0,0,1,0,0,0,0,1,0,0,0,0,-1)$, and χ^{nat} has 16 non-zero elements, equal to ± 1. Note that χ^{nat} is a rank-1 matrix with $\text{Tr}(\chi^{\text{nat}}) = \sum_\alpha |u_\alpha|^2 = d$.

We then apply numerical procedure of the SVD decomposition, which diagonalizes the matrix χ^{nat} for the desired unitary process,

$$\chi^{\text{nat}} = V \text{diag}(d, 0, \ldots, 0) V^\dagger, \tag{C.2}$$

where V is a unitary $d^2 \times d^2$ matrix and the only non-zero eigenvalue is equal to d because $\text{Tr}(\chi^{\text{nat}}) = d$. The columns of thus obtained transformation matrix V are the vectorized forms of thus introduced SVD-basis matrices E_α^{SVD},

$$E_\alpha^{\text{SVD}} = \sum_{\beta=1}^{d^2} V_{\beta\alpha}\, E_\beta^{\text{nat}}. \tag{C.3}$$

Note that the notation V used in Appendix B has a different meaning.

The matrices of the SVD basis introduced via Eqs. (C.2) and (C.3) have the different

normalization compared with the Pauli basis,

$$\mathrm{Tr}(E_\alpha^{\mathrm{SVD}\dagger} E_\beta^{\mathrm{SVD}}) = \delta_{\alpha\beta}. \tag{C.4}$$

Correspondingly, the normalization of the process matrix χ^{SVD} in the SVD basis is $\mathrm{Tr}\chi^{\mathrm{SVD}} = d$ (for a trace-preserving process). For the ideal unitary process the matrix χ^{SVD} has one non-zero (top left) element, which is equal to $\sqrt{d}$. For an imperfect realization of the desired unitary operation the top left element is related to the process fidelity as $\chi_{11}^{\mathrm{SVD}} = F_\chi d$.

Note that when the numerical SVD procedure (C.2) is applied to χ^{nat} of ideal CZ and/or Toffoli gates, many (most) of the resulting SVD-basis matrices E_α^{SVD} coincide with the matrices of the natural basis E_α^{nat}. Since these matrices contain only one non-zero element, the matrix Φ in Eq. (4.8) is simpler (has more zero elements) than for the Pauli or Pauli-error basis. (The number of non-zero elements of Φ in the SVD basis is crudely twice less for the CZ gate and 4 times less for the Toffoli gate.) As the result, from the computational point of view it is easier to use the SVD basis than the Pauli-error basis: less memory and less computational time are needed.

Bibliography

[1] A. V. Rodionov, A. Veitia, R. Barends, J. Kelly, D. Sank, J. Wenner, J. M. Martinis, R. L. Kosut, and A. N. Korotkov, "Compressed sensing quantum process tomography for superconducting quantum gates", Phys. Rev. B **90**, 144504 (2014).

[2] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information* (Cambridge University Press, Cambridge, England, 2000).

[3] P. Shor, "Algorithms for quantum computation: Discrete logarithms and factoring", in Proc. 35th Annu. Symp. on the Foundations of Computer Science (ed. Shafi Goldwasser), pp. 124 - 134, (IEEE Computer Society Press, Los Alamitos, California, 1994).

[4] P. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer", SIAM J. Comput. **26**, pp. 1484 - 1509 (1997), also arXiv:quant-ph/9598927v2.

[5] A. Ekert and R. Jozsa, "Quantum computation and Shor's factoring algorithm", Rev. Mod. Phys. **68 (3)**, pp. 733 – 753 (1996).

[6] L. M. K. Vandersypen, M. Steffen, G. Breyta, C. S. Yannoni, M. H. Sherwood, and I. L. Chuang, "Experimental realization of Shor's quantum factoring algorithm using nuclear magnetic resonance", Nature **414 (6866)**, pp. 883 – 887 (2001).

[7] R. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems", Communications of the ACM **21**, pp. 120 – 126 (1978).

[8] L. Grover, "Quantum mechanics helps in searching for a needle in a haystack", Phys. Rev. Lett. **79**, pp. 325 – 328 (1997), see also arXiv:quant-ph/9605043.

[9] A. K. Ekert, "Quantum cryptography based on Bells theorem", Phys. Rev. Lett. **67**, pp. 661 - 663 (1991).

[10] D. Deutsch, "Quantum Computational Networks", Proc. R. Soc. Lond. A **425**, pp. 73 - 90 (1989).

[11] R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. C. White, J. Mutus, A. G. Fowler, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, C. Neill, P. O'Malley, P.

Roushan, A. Vainsencher, J. Wenner, A. N. Korotkov, A. N. Cleland, and John M. Martinis, "Superconducting quantum circuits at the surface code threshold for fault tolerance", Nature **508**, 500 (2014).

[12] 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 (1997).

[13] I. L. Chuang and M. A. Nielsen. "Prescription for experimental determination of the dynamics of a quantum black box", Journal of Modern Optics, **44**, pp. 2455 – 2467 (1997).

[14] N. Boulant, T. F. Havel, M. A. Pravia, and D. G. Cory, "Robust method for estimating the Lindblad operators of a dissipative quantum process from measurements of the density operator at multiple time points", Phys. Rev. A **67**, 042322 (2003).

[15] A. Bendersky, F. Pastawski, and J. P. Paz, "Selective and Efficient Estimation of Parameters for Quantum Process Tomography", Phys. Rev. Lett. **100**, 190403 (2008).

[16] M. Mohseni and A. T. Rezakhani, "Equation of motion for the process matrix: Hamiltonian identification and dynamical control of open quantum systems", Phys. Rev. A **80**, 010101 (2009).

[17] A. G. Kofman and A. N. Korotkov, "Two-qubit decoherence mechanisms revealed via quantum process tomography", Phys. Rev. A **80**, 042103 (2009).

[18] A. N. Korotkov, "Error matrices in quantum process tomography", arXiv:1309.6405.

[19] D. W. Leung, "Choi's Proof and Quantum Process Tomography", J. Math. Phys. **44**, 528 (2003).

[20] G. M. D'Ariano and P. Lo Presti, "Imprinting Complete Information about a Quantum Channel on its Output State", Phys. Rev. Lett. **91**, 047902 (2003).

[21] J. Emerson, M. Silva, O. Moussa, C. Ryan, M. Laforest, J. Baugh, D. G. Cory, and R. Laamme, "Symmetrized characterization of noisy quantum processes", Science **317**, 1893 (2007).

[22] M. Mohseni and D. A. Lidar, "Direct Characterization of Quantum Dynamics", Phys. Rev. Lett. **97**, 170501 (2006).

[23] M. M. Wolf, J. Eisert, T. S. Cubitt, and J. I. Cirac, "Assessing Non-Markovian Quantum Dynamics", Phys. Rev. Lett. **101**, 150402 (2008).

[24] Yu. I. Bogdanov, G. Brida, M. Genovese, S. P. Kulik, E. V. Moreva, and A. P. Shurupov, "Statistical Estimation of the Efficiency of Quantum State Tomography Protocols", Phys. Rev. Lett. **105**, 010404 (2010).

[25] A. G. White, A. Gilchrist, G. J. Pryde, J. L. OBrien, M. J. Bremner, and N. K. Langford, "Measuring two-qubit gates", JOSA B **24**, pp. 172-183 (2007).

[26] K. Banaszek, G. M. D'Ariano, M. G. A. Paris and M. F. Sacchi, "Maximum likelihood estimation of the density matrix". Phys. Rev. A **61**, 010304(R) (1999).

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

[28] M. G. A. Paris, G. M. D'Ariano and M. F. Sacchi, "Maximum-likelihood method in quantum estimation", in *Bayesian inference and maximum entropy methods in science and engineering*, **568** of AIP Conf. Proc., page 456 (2001).

[29] R. Kosut, I. A. Walmsley, H. Rabitz, "Optimal Experiment Design for Quantum State and Process Tomography and Hamiltonian Parameter Estimation", arXiv:quant-ph/0411093.

[30] M. Mohseni, A. T. Rezakhani, and D. A. Lidar, "Quantum-process tomography: Resource analysis of different strategies", Phys. Rev. A **77**, 032322 (2008).

[31] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, "Randomized benchmarking of quantum gates", Phys. Rev. A **77**, 012307 (2008).

[32] J. Emerson, R. Alicki, and K. Zyczkowski, "Scalable Noise Estimation with Random Unitary Operators", J. Opt. B **7**, S347 (2005).

[33] E. Magesan, J. M. Gambetta, B. R. Johnson, C. A. Ryan, J. M. Chow, S. T. Merkel, M. P. da Silva, G. A. Keefe, M. B. Rothwell, T. A. Ohki, M. B. Ketchen, and M. Steffen, "Efficient Measurement of Quantum Gate Error by Interleaved Randomized Benchmarking", Phys. Rev. Lett. **109**, 080505 (2012).

[34] S. T. Flammia and Y.-K. Liu, "Direct Fidelity Estimation from Few Pauli Measurements", Phys. Rev. Lett. **106**, 230501 (2011).

[35] M. P. da Silva, O. Landon-Cardinal, and D. Poulin, "Practical Characterization of Quantum Devices without Tomography", Phys. Rev. Lett. **107**, 210404 (2011).

[36] J. M. Chow, J. M. Gambetta, L. Tornberg, J. Koch, L. S. Bishop, A. A. Houck, B. R. Johnson, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, "Randomized Benchmarking and Process Tomography for Gate Errors in a Solid-State Qubit", Phys. Rev. Lett. **102**, 090502 (2009).

[37] A. D. Córcoles, J. M. Gambetta, J. M. Chow, J. A. Smolin, M. Ware, J. Strand, B. L. T. Plourde, and M. Steffen, "Process verification of two-qubit quantum gates by randomized benchmarking", Phys. Rev. A **87**, 030301(R) (2013).

[38] L. Steffen, M. P. da Silva, A. Fedorov, M. Baur, and A. Wallraff, "Experimental Monte Carlo Quantum Process Certification", Phys. Rev. Lett. **108**, 260506 (2012).

[39] R. Kosut, "Quantum Process Tomography via L1-norm Minimization", arXiv:0812.4323.

[40] A. Shabani, R. L. Kosut, M. Mohseni, H. Rabitz, M. A. Broome, M. P. Almeida, A. Fedrizzi, and A. G. White, "Efficient Measurement of Quantum Dynamics via Compressive Sensing", Phys. Rev. Lett. **106**, 100401 (2011).

[41] E. J. Candès, J. K. Romberg, T. Tao, "Stable signal recovery from incomplete and inaccurate measurements", Comm. Pure Appl. Math. **59**, 1207 (2006).

[42] E. J. Candès, "The restricted isometry property and its implications for compressed sensing", C. R. Acad. Sci. Paris, Ser. I **346**, 589 (2008).

[43] D. L. Donoho, "Compressed sensing", IEEE Trans. Inf. Theory **52**, 1289 (2006).

[44] E. J. Candes, M. B. Wakin, "An Introduction To Compressive Sampling", IEEE Signal Proc. Mag. **25**, 21 (2008).

[45] Y. H. Shih and C. O. Alley, "New Type of Einstein-Podolsky-Rosen-Bohm Experiment Using Pairs of Light Quanta Produced by Optical Parametric Down Conversion", Phys. Rev. Lett. **61**, pp. 2921-2924 (1988).

[46] T. E. Kiess, Y. H. Shih, A. V. Sergienko and C. O. Alley,"Einstein-Podolsky-Rosen-Bohm experiment using pairs of light quanta produced by type-II parametric down-conversion", Phys. Rev. Lett. **71**, pp. 3893-3897 (1993).

[47] P. G. Kwiat, K. Mattle, H. Weinfurter, A. Zeilinger, A. V. Sergienko and Y. H. Shih, "New High-Intensity Source of Polarization-Entangled Photon Pairs", Phys. Rev. Lett. **75**, pp. 4337 – 4341 (1995).

[48] P. G. Kwiat, E. Waks, A. G. White, I. Appelbaum and P. H. Eberhard, "Ultrabright Source of Polarization Entanglement Photons", Phys. Rev. A **60**, pp. R773 – R776 (1999).

[49] Y.-H. Kim, S. P. Kulik and Y. Shih, "Bell-State Preparation Using Pulsed Nondegenerate Two-Photon Entanglement", Phys. Rev. A **63**, 060301R (2001).

[50] Y.-H. Kim, S. P. Kulik and Y. Shih,"Quantum Teleportation of a Polarization State with a Complete Bell State Measurement", Phys. Rev. Lett. **86**, pp. 1370 – 1373 (2001).

[51] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, "Measurement of qubits", Phys. Rev. A, **64**, 052312 (2001).

[52] I. J. Cirac and P. Zoller, "Quantum computations with cold trapped ions", Phys. Rev. Lett., **74**, pp. 4091 - 4094 (1995).

[53] D. J. Wineland, C. Monroe, W. M. Itano, D. Leibfried, B. E. King, D. M. Meekhof, "Experimental issues in coherent quantum-state manipulation of trapped atomic ions", J. Res. Natl. Inst. Stand. Technol. **103**, pp. 259 - 328 (1998); also available in arXiv:9710025.

[54] R. Blatt and D. Wineland, "Entangled states of trapped atomic ions", Nature **453**, pp. 1008-1015 (2008).

[55] D. G. Cory, M. D. Price, and T. F. Havel, "Nuclear Magnetic Resonance Spectroscopy: an experimentally accessible paradigm for quantum computing", arxiv quant-ph/9709001v2.

[56] N. A. Gershenfeld and I. L. Chuang, "Bulk spin-resonance quantum computation", Science **275**, pp 350 - 356 (1997).

[57] D. G. Cory, M. D. Price, W. Maas, E. Knill, R. Laflamme, W. H. Zurek, T. F. Havel, and S. S. Somaroo, "Experimental Quantum Error Correction", Phys. Rev. Lett. **81**, pp. 2152 - 2155 (1998).

[58] E. Charron, E. Tiesinga, F. Mies, and C. Williams, "Optimizing a Phase Gate Using Quantum Interference", Phys. Rev. Lett. **88**, 077901 (2002).

[59] K. Eckert, J. Mompart, X. X. Yi, J. Schliemann, D. Bruss, G. Birkl, and M. Lewenstein, "Quantum computing in optical microtraps based on the motional states of neutral atoms", Phys. Rev. A **66**, 042317 (2002).

[60] O. Morsch and M. Oberthaler, "Dynamics of Bose-Einstein condensates in optical lattices", Rev. Mod. Phys. **78**, pp. 179 - 215 (2006).

[61] F. Jelezko, T. Gaebel, I. Popa, A. Gruber, and J. Wrachtrup, "Observation of Coherent Oscillations in a Single Electron Spin", Phys. Rev. Lett. **92**, 076401 (2004).

[62] M. V. Gurudev Dutt, L. Childress, L. Jiang, E. Togan, J. Maze, F. Jelezko, A. S. Zibrov, P. R. Hemmer, M. D. Lukin, "Quantum Register Based on Individual Electronic and Nuclear Spin Qubits in Diamond", Science **316**, pp. 1312 – 1316 (2007).

[63] L. Robledo, L. Childress, H. Bernien, B. Hensen, P. F. A. Alkemade, and R. Hanson, "High-fidelity projective read-out of a solid-state spin quantum register", Nature **477**, pp. 574 – 578 (2011).

[64] A. Gali, "Theory of the neutral nitrogen-vacancy center in diamond and its application to the realization of a qubit", Phys. Rev. B **79**, 235210 (2009).

[65] V. Bouchiat, D. Vion, P. Joyez, D. Esteve, and M. H. Devoret, "Quantum coherence with a single cooper pair", Physica Scripta , **T76**, pp. 165 – 170 (1998).

[66] J. E. Mooij, T. P. Orlando, L. Levitov, L. Tian, C. H. van der Wal, and S. Lloyd. "Josephson persistent-current qubit". Science, **285(5430)**, pp. 1036 – 1039 (1999).

[67] R. W. Simmonds, K. M. Lang, D. A. Hite, S. Nam, D. P. Pappas, and J. M. Martinis. "Decoherence in Josephson Phase Qubits from Junction Resonators". Phys. Rev. Lett., **93**, 077003 (2004).

[68] K. B. Cooper, M. Steffen, R. McDermott, R. W. Simmonds, S. Oh, D. A. Hite, D. P. Pappas, and J. M. Martinis. "Observation of Quantum Oscillations between a Josephson Phase Qubit and a Microscopic Resonator Using Fast Readout", Phys. Rev. Lett. **93**, 180401 (2004).

[69] R. Jozsa, "Fidelity for mixed quantum states", J. Mod. Opt. **41**, pp. 2315 – 2323 (1994).

[70] K. Zyczkowski and I. Bengtsson, "An Introduction to Quantum Entanglement: a Geometric Approach", arXiv:quant-ph/0606228 (2006).

[71] J. A. Miszczak, Z. Puchała, P. Horodecki, A. Uhlmann, K. Życzkowski, "Sub– and super–fidelity as bounds for quantum fidelity", arXiv:0805.2037v2 (2008).

[72] A. K. Rajagopal, A. R. Usha Devi, and R. W. Rendell, "Kraus representation of quantum evolution and fidelity as manifestations of Markovian and non-Markovian forms", Phys. Rev. A **82**, 042107 (2010).

[73] B. D. Josephson, "Possible new effects in superconductive tunnelling", Physics Letters **1**, pp. 251 – 253 (1962).

[74] P. W. Anderson and J. M. Rowell, "Probable Observation of the Josephson Superconducting Tunneling Effect", Phys. Rev. Lett. **10**, pp. 230 – 232 (1963).

[75] Y. Nakamura, Y. A. Pashkin, and J. S. Tsai, "Coherent control of macroscopic quantum states in a single-cooper-pair box", Nature **398**, pp. 786 - 788 (1999).

[76] J. M. Martinis, S. Nam, J. Aumentado, and C. Urbina. "Rabi oscillations in a large Josephson-junction qubit", Phys. Rev. Lett. **89**, 117901 (2002).

[77] M. Devoret, A. Wallraff, and J. Martinis. "Superconducting qubits: A short review", arXiv:cond-mat/0411174v1 (2004).

[78] J. M. Martinis, "Superconducting Phase Qubits", Quantum Information Processing **8**, 81 (2009).

[79] D. Vion, A. Aassime, A. Cottet, P. Joyez, H. Pothier, C. Urbina, D. Esteve, M. H. Devoret, "Manipulating the Quantum State of an Electrical Circuit", Science **296**, pp. 886 – 889 (2002).

[80] J. Koch, T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, "Charge-insensitive qubit design derived from the Cooper pair box", Phys. Rev. A **76**, 042319 (2007).

[81] J. A. Schreier, A. A. Houck, J. Koch, D. I. Schuster, B. R. Johnson, J. M. Chow, J. M. Gambetta, J. Majer, L. Frunzio, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, "Suppressed charge noise decoherence in superconducting charge qubits", Phys. Rev. B **77**, 180502 (2008).

[82] V. E. Manucharyan, J. Koch, L. I. Glazman, M. H. Devoret, "Fluxonium: Single Cooper-Pair Circuit Free of Charge Offsets", Science **326**, pp. 113 – 116 (2009).

[83] T. Duty, D. Gunnarsson, K. Bladh, and P. Delsing, "Coherent dynamics of a Josephson charge qubit", "Coherent dynamics of a Josephson charge qubit", Phys. Rev. B **69**, 140503(R) (2004).

[84] A. Guillaume, J. F. Schneiderman, P. Delsing, H. M. Bozler, and P. M. Echternach, "Free evolution of superposition states in a single Cooper pair box", Phys. Rev. B **69**, 132504 (2004).

[85] J. Clarke, "Principles and Applications of SQUIDs", Proc. IEEE **77**, pp. 1208 – 1223 (1989).

[86] J. R. Friedman, V. Patel, W. Chen, S. K. Tolpygo, and J. E. Lukens, "Quantum superposition of distinct macroscopic states", Nature **406**, pp. 43 – 46 (2000).

[87] C. H. van der Wal, A. C. J. ter Haar, F. K. Wilhelm, R. N. Schoten, C. J. P. M. Harmans, T. P. Orlando, S. Lloyd, and J. E. Mooij, "Quantum Superposition of Macroscopic Persistent-Current States", Science **290**, pp. 773 – 777 (2000).

[88] I. Chiorescu, Y. Nakamura, C. J. P. M. Harmans, J. E. Mooij, "Coherent Quantum Dynamics of a Superconducting Flux Qubit", Science **299**, pp. 1869 – 1871 (2003).

[89] D. I. Schuster, A. A. Houck, J. A. Schreier, A. Wallraff, J. M. Gambetta, A. Blais, L. Frunzio, J. Majer, B. R. Johnson, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, "Resolving photon number states in a superconducting circuit", Nature (London) **445**, 515 (2007).

[90] A. A. Houck, D. I. Schuster, J. M. Gambetta, J. A. Schreier, B. R. Johnson, J. M. Chow, L. Frunzio, J. Majer, M. H. Devoret, S. M. Girvin, and R/ J. Schoelkopf, "Generating single microwave photons in a circuit", Nature (London) **449**, 328 (2007).

[91] J. Majer, J. M. Chow, J. M. Gambetta, J. Koch, B. R. Johnson, J. A. Schreier, L. Frunzio, D. I. Schuster, A. A. Houck, A. Wallraff, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, "Coupling superconducting qubits via a cavity bus", Nature (London) **449**, 443 (2007).

[92] A. Blais, R.-S. Huang, A. Wallraff, S. M. Girvin, and R. J. Schoelkopf, "Cavity quantum electrodynamics for superconducting electrical circuits: An architecture for quantum computation", Phys. Rev. A **69**, 062320 (2004).

[93] A. Wallraff, D. I. Schuster, A. Blais, L. Frunzio, R.- S. Huang, J. Majer, S. Kumar, S. M. Girvin, and R. J. Schoelkopf, "Strong coupling of a single photon to a superconducting qubit using circuit quantum electrodynamics", Nature bf 431, pp. 162 – 167 (2004).

[94] D. I. Schuster, A. Wallraff, A. Blais, L. Frunzio, R.-S. Huang, J. Majer, S. M. Girvin, and R. J. Schoelkopf, "ac Stark shift and dephasing of a superconducting qubit strongly coupled to a cavity field", Phys. Rev. Lett. **94**, 123602 (2005).

[95] A. Wallraff, D. I. Schuster, A. Blais, L. Frunzio, J. Majer, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, "Approaching unit visibility for control of a superconducting qubit with dispersive readout", Phys. Rev. Lett. **95**, 060501 (2005).

[96] M. Hatridge, S. Shankar, M. Mirrahimi, F. Schackert, K. Geerlings, T. Brecht, K. M. Sliwa, B. Abdo, L. Frunzio, S. M. Girvin, R. J. Schoelkopf, and M. H. Devoret, "Quantum Back-Action of an Individual Variable-Strength Measurement", Science **339**, pp. 178 – 181 (2013).

[97] S. Filipp, P. Maurer, P. J. Leek, M. Baur, R. Bianchetti, J. M. Fink, M. Göppl, L. Steffen, J. M. Gambetta, A. Blais, and A. Wallraff, "Two-qubit state tomography using a joint dispersive read-out", Phys. Rev. Lett. **102**, 200402 (2009).

[98] R. Barends, J. Kelly, A. Megrant, D. Sank, E. Jeffrey, Y. Chen,Y. Yin, B. Chiaro, J. Mutus, C. Neill, P. O'Malley, P. Roushan, J. Wenner, T. C. White, A. N. Cleland, and John M. Martinis, "Coherent Josephson Qubit Suitable for Scalable Quantum Integrated Circuits", Phys. Rev. Lett. **111**, 080502 (2013).

[99] J. Kelly, R. Barends, A. G. Fowler, A. Megrant, E. Jeffrey, T. C. White, D. Sank, J. Y. Mutus, B. Campbell, Yu Chen, Z. Chen, B. Chiaro, A. Dunsworth, I. − C. Hoi, C. Neill, P. J. J. O′Malley, C. Quintana, P. Roushan, A. Vainsencher, J. Wenner, A. N. Cleland, and John M. Martinis, "State preservation by repetitive error detection in a superconducting quantum circuit", arXiv:quant-ph/1411.7403 (2014).

[100] A. G. Kofman, Q. Zhang, J. M. Martinis, and A. N. Korotkov. "Theoretical analysis of measurement crosstalk for coupled josephson phase qubits", Phys. Rev. B **75**, 014524 (2007).

[101] M. Steffen, M. Ansmann, R. McDermott, N. Katz, R. C. Bialczak, E. Lucero, M. Neeley, E. M. Weig, A. N. Cleland, and J. M. Martinis, "State Tomography of Capacitively Shunted Phase Qubits with High Fidelity", Phys. Rev. Lett. **97**, 050502 (2006).

[102] J. Clarke and A. I. Braginsky, *The SQUID Handbook: Fundamentals and Technology of SQUIDs and SQUID Systems* (Wiley-VCH, Weinheim, 2004).

[103] D. Leibfried, D. M. Meekhof, B. E. King, C. Monroe, W. M. Itano, and D. J. Wineland, "Experimental determination of the motional quantum state of a trapped atom", Phys. Rev. Lett. **77**, pp. 4281 – 4285 (1996).

[104] J. F. Poyatos, R. Walser, J. I. Cirac, P.Zoller, and R. Blatt, "Motion tomography of a single trapped ion", Phys. Rev. A **53**, pp. R1966 – R1969 (1996).

[105] I. L. Chuang, L. M. K. Vandersypen, X. Zhou, D. W. Leung, S. Lloyd, "Experimental realization of a quantum algithm", Nature **393**, pp. 143 – 146 (1998).

[106] I.L. Chuang, N. Gershenfeld, M. Kubinee, "Experimental Implementation of Fast Quantum Searching", Phys. Rev. Lett. **80**, 3408 (1998).

[107] G. Klose, G. Smith, and P.S. Jessen, "Measuring the quantum state of a large angular momentum", Phys. Rev. Lett. **86**, pp. 4721 – 4724 (2001).

[108] F. A. Bonk, R. S. Sarthour, E. R. de Azevedo, J. D. Bulnes, G. L. Mantovani, J. C. C. Freitas, T. J. Bonagamba, A. P. Guimaraes, I. S. Oliveira, "Quantum-state tomography for quadrupole nuclei and its application on a two-qubit system", Phys. Rev. A **69**, 042322 (2004).

[109] J. D. Bulnes, F. A. Bonk, R. S. Sarthour, E. R. de Azevedo, J. C. C. Freitas, T. J. Bonagamba, I. S. Oliveira, "Quantum Information Processing through Nuclear Magnetic Resonance", Brazilian Journal of Physics, **35**, 617 (2005).

[110] G. G. Stokes, "On the composition and resolution of streams of polarized light from different sources", Trans. Cambridge. Philos. Soc. **9**, pp. 399-416 (1852), Reprinted in *Mathematical and Physical Papers*, **Vol. III**, pp. 233 – 258, Johnson Reprint Corp. (1966).

[111] M. Steffen, M. Ansmann, R. McDermott, N. Katz, Radoslaw C. Bialczak, Erik Lucero, Matthew Neeley, E.M. Weig, A.N. Cleland, John M. Martinis, "State tomography of capacitively shunted phase qubits with high fidelity", Phys. Rev. Lett. **97**, 050502 (2006).

[112] M. Steffen, M. Ansman, R. C. Bialczak, N. Katz, E. Lucero, R. McDermott, M. Neeley, E. M. Weig, A. N. Cleland, J. M. Martinis, "Measurement of the entanglement of two superconducting qubits via state tomography", Science **313**, pp. 1423 – 1425 (2006).

[113] M. Mariantoni, H. Wang, T. Yamamoto, M. Neeley, R. C. Bialczak, Y. Chen, M. Lenander, E. Lucero, A. D. O'Connell, D. Sank, M. Weides, J. Wenner, Y. Yin, J. Zhao, A. N. Korotkov, A. N. Cleland, and J. M. Martinis, "Implementing the Quantum von Neumann Architecture with Superconducting Circuits", Science **334**, 61 (2011).

[114] T. Yamamoto, M. Neeley, E. Lucero, R. C. Bialczak, J. Kelly, M. Lenander, M. Mariantoni, A. D. O'Connell, D. Sank, H. Wang, M. Weides, J. Wenner, Y. Yin, A. N. Cleland, and J. M. Martinis, "Quantum process tomography of two-qubit controlled-Z and controlled-NOT gates using superconducting phase qubits", Phys. Rev. B **82**, 184515 (2010).

[115] See for example: M. H. Kutner, C. J. Nachtsheim, and J. Neter, "Applied Linear Regression Models", 4th Edn. Boston: McGraw-Hill Irwin. ISBN: 978-0-07-301344-2, 2004. (page 25).

[116] Z. Hradil, "Quantum-State Estimation", Phys. Rev. A **55**, R1561(R) (1997).

[117] M. Jezek, J. Fiurasek, and Z. Hradil, "Quantum inference of states and processes", Phys. Rev. A **68**, 012305 (2003).

[118] A. M. Childs, I. L. Chuang, and D. W. Leung, "Realization of quantum process tomography in NMR", Phys. Rev. A **64**, 012314 (2001).

[119] Y. S. Weinstein, T. F. Havel, J. Emerson, N. Boulant, M. Saraceno, S. Lloyd, and D. G. Cory, "Quantum process tomography of the quantum Fourier transform", J. Chem. Phys. **121**, pp. 6117 – 6133 (2004).

[120] N. Boulant, T. F. Havel, M. A. Pravia, and D. G. Cory, "Robust method for estimating the Lindblad operators of a dissipative quantum process from measurements of the density operator at multiple time points", Phys. Rev. A **67**, 042322 (2003).

[121] M.W. Mitchell, C.W. Ellenor, S. Schneider, and A.M. Steinberg, "Diagnosis, Prescription, and Prognosis of a Bell-State Filter by Quantum Process Tomography", Phys. Rev. Lett. **91**, 120402 (2003).

[122] J. L. O′Brien, G. J. Pryde, A. Gilchrist, D. F. V. James, N. K. Langford, T. C. Ralph, and A.G.White, "Quantum Process Tomography of a Controlled-NOT Gate", Phys. Rev. Lett. **93**, 080502 (2004).

[123] M. Riebe, K. Kim, P. Schindler, T. Monz, P. O. Schmidt, T. K. Körber, W. Hänsel, H. Häffner, C. F. Roos, and R. Blatt, "Process Tomography of Ion Trap Quantum Gates", Phys. Rev. Lett. **97**, 220407 (2006).

[124] D. Hanneke, J. P. Home, J. D. Jost, J. M. Amini, D. Leibfried, and D. J.Wineland, "Realization of a programmable two-qubit quantum processor", Nature Phys. **6**, pp. 13 – 16 (2010).

[125] M. Neeley, M. Ansmann, R.C. Bialczak, M. Hofheinz, N. Katz, E. Lucero, A. O'Connell, H. Wang, A.N. Cleland, J.M. Martinis, "Process tomography of quantum memory in a Josephson-phase qubit coupled to a two-level state", Nature Physics **4**, pp. 523 – 526 (2008).

[126] R. C. Bialczak, M. Ansmann, M. Hofheinz, E. Lucero, M. Neeley, A. D. OConnell, D. Sank, H. Wang, J. Wenner, M. Steffen, A. N. Cleland, and J. M. Martinis, "Quantum Process Tomography of a Universal Entangling Gate Implemented with Josephson Phase Qubits", Nature Physics **6**, pp. 409 – 413 (2010).

[127] J. M. Chow, A. D. Córcoles, J. M. Gambetta, C. Rigetti, B. R. Johnson, J. A. Smolin, J. R. Rozen, G. A. Keefe, M. B. Rothwell, M. B. Ketchen, and M. Steffen, "Simple All-Microwave Entangling Gate for Fixed-Frequency Superconducting Qubits", Phys. Rev. Lett. **107**, 080502 (2011).

[128] A. Dewes, F. R. Ong, V. Schmitt, R. Lauro, N. Boulant, P. Bertet, D. Vion, and D. Esteve, "Characterization of a Two-Transmon Processor with Individual Single-Shot Qubit Readout", Phys. Rev. Lett. **108**, 057002 (2012).

[129] M. D. Reed, L. DiCarlo, S. E. Nigg, L. Sun, L. Frinzio, S. M. Girvin and R. J. Schoelkopf, "Realization of Three-Qubit Quantum Error Correction with Superconducting Circuits", Nature **482**, pp. 382 – 385 (2012).

[130] A. Fedorov, L. Steffen, M. Baur, M. P. da Silva, and A. Wallraff, "Implementation of a Toffoli gate with superconducting circuits", Nature **481**, pp. 170 – 172 (2012).

[131] J. M. Chow, J. M. Gambetta, A. D. Córcoles, S. T. Merkel, J. A. Smolin, C. Rigetti, S. Poletto, G. A. Keefe, M. B. Rothwell, J. R. Rozen, M. B. Ketchen, and M. Steffen, "Universal Quantum Gate Set Approaching Fault-Tolerant Thresholds with Superconducting Qubits", Phys. Rev. Lett. **109**, 060501 (2012).

[132] J. M. Chow, J. M. Gambetta, A. W. Cross, S. T. Merkel, C. Rigetti and M. Steffen, "Microwave-activated conditional-phase gate for superconducting qubits", New J. Phys. **15**, 115012 (2013).

[133] M. Ježek, J. Fiurášek, and Z. Hradil, "Quantum inference of states and processes", Phys. Rev. A **68**, 012305 (2003).

[134] M. Mičuda, M. Sedlák, I. Straka, M. Miková, M. Dušek, M. Ježek, and J. Fiurášek, "Process-fidelity estimation of a linear optical quantum-controlled-Z gate: A comparative study", Phys. Rev. A **89**, 042304 (2014).

[135] Yu. I. Bogdanov, M. V. Chekhova, L. A. Krivitsky, S. P. Kulik, A. N. Penin, A. A. Zhukov, L. C. Kwek, C. H. Oh, and M. K. Tey, "Statistical reconstruction of qutrits", Phys. Rev. A **70**, 042303 (2004).

[136] G. D. Fuchs, A. L. Falk, V. V. Dobrovitski, and D. D. Awschalom, "Spin Coherence during Optical Excitation of a Single Nitrogen-Vacancy Center in Diamond", Phys. Rev. Lett. **108**, 157602 (2012).

[137] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, 2004).

[138] J. Löfberg, in Proceedings of the IEEE CACSD Symposium, p. 284 (Taipei, Taiwan, 2004); the package YALMIP is available at http://users.isy.liu.se/johanl/yalmip.

[139] M. C. Grant and S. P. Boyd, "CVX: Matlab software for disciplined convex programming", http://cvxr.com/cvx.

[140] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker and J. Eisert, "Quantum State Tomography via Compressed Sensing", Phys. Rev. Lett. **105**, 150401 (2010).

[141] S. T. Flammia, D. Gross, Y.-K. Liu and J. Eisert, "Quantum Tomography via Compressed Sensing: Error Bounds, Sample Complexity, and Efficient Estimators", New J. Phys. **14**, 095022 (2012).

[142] C. H. Baldwin, A. Kalev, and I. H. Deutsch, "Quantum process tomography of unitary and near-unitary maps", Phys. Rev. A **90**, 012110 (2014).

[143] W.-T. Liu, T. Zhang, J.-Y. Liu, P.-X. Chen, and J.-M. Yuan, "Experimental Quantum State Tomography via Compressed Sampling", Phys. Rev. Lett. **108**, 170403 (2012).

[144] A. Smith, C. A. Riofrio, B. E. Anderson, H. Sosa-Martinez, I. H. Deutsch, and P. S. Jessen, "Quantum state tomography by continuous measurement and compressed sensing", Phys. Rev. A **87**, 030102(R) (2013).

[145] A. Jamiołkowski, "Linear transformations which preserve trace and positive semide finiteness of operators", Rep. Math. Phys. **3**, 275 (1972).

[146] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The Application of Compressed Sensing for Rapid MR Imaging", Magn. Reson. Med. **58**, 1182 (2007).

[147] R. G. Baraniuk, "Compressive sensing", IEEE Signal Process. Mag. **24**, No. 4, 118 (2007).

[148] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robot face recognition via sparse representation", IEEE Trans. Pattern Anal. Mach. Intell. **31**, 210 (2009).

[149] D. J. Brady, K. Choi, D. L. Marks, R. Horisaki, and S. Lim, "Compressive holography", Optics Exp. **17**, 13040 (2009).

[150] F. J. Herrmann and G. Hennenfent, "Non-parametric seismic data recovery with curvelet frames", Geophys. J. Int. **173**, pp. 233 – 248 (2008).

[151] F. W. Strauch, P. R. Johnson, A. J. Dragt, C. J. Lobb, J. R. Anderson, and F. C. Wellstood, "Quantum Logic Gates for Coupled Superconducting Phase Qubits", Phys. Rev. Lett. **91**, 167005 (2003).

[152] J. Kelly, R. Barends, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, I.-C. Hoi, E. Jeffrey, A. Megrant, J. Mutus, C. Neill, P. J. J. O'Malley, C. Quintana, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White, A. N. Cleland, and J. M.

Martinis, "Optimal Quantum Control Using Randomized Benchmarking", Phys. Rev. Lett. **112**, 240504 (2014).

[153] A. Uhlmann, "The Transition Probability in the State Space of a *-Algebra", Rep. Math. Phys. **9**, pp. 273 - 279 (1976).

[154] A. Gilchrist, N. K. Langford, and M. A. Nielsen, "Distance measures to compare real and ideal quantum processes", Phys. Rev. A **71**, 062310 (2005).

[155] J. F. Sturm, Optimization Meth. & Soft., **11 & 12**, 625 (1999); the solver SeDuMi is available at http://sedumi.ie.lehigh.edu.

[156] K. C. Toh, M. J. Todd, and R. H. Tütüncü, Optimization Meth. & Soft., **11 & 12**, 545 (1999); the solver SDPT3 is available at http://www.math.nus.edu.sg/∼mattohkc/sdpt3.html.

[157] D. G. Cory, M. D. Price, W. Maas, E. Knill, R. Laflamme, W. H. Zurek, T. F. Havel, and S. S. Somaroo, "Experimental Quantum Error Correction", Phys. Rev. Lett. **81**, 2152 (1998).

[158] T. Monz, K. Kim, W. Hänsel, M. Riebe, A. S. Villar, P. Schindler, M. Chwalla, M. Hennrich, and R. Blatt, "Realization of the Quantum Toffoli Gate with Trapped Ions", Phys. Rev. Lett. **102**, 040501 (2009).

[159] M. Horodecki, P. Horodecki, R. Horodecki, "General teleportation channel, singlet fraction, and quasidistillation", Phys. Rev. A **60**, 1888 (1999).

[160] M. A. Nielsen, "A simple formula for the average gate fidelity of a quantum dynamical operation", Phys. Lett. A **303**, 249 - 252 (2002).

[161] L. H. Pedersen, N. M. Møller and K. Mølmer, "The distribution of quantum fidelities", Phys. Lett. A **372**, pp. 7028 – 7032 (2008).

[162] E. Magesan, R. Blume-Kohout and J. Emerson, "Gate fidelity fluctuations and quantum process invariants", Phys. Rev. A **84**, 012309 (2011).

[163] D. Poulin, R. Blume-Kohout, R. Laflamme, and H. Ollivier, "Exponential speedup with a single bit of quantum information: measuring the average fidelity decay", Phys. Rev. Lett. **92**, 177906 (2004).

[164] C. P. Robert and G. Casella, *"Monte Carlo Statistical Methods" (second edition)* (New York: Springer-Verlag, 2004).

[165] J. von Neumann, "Various techniques used in connection with random digits. Monte Carlo methods", Nat. Bureau Standards, **12**, pp. 36 - 38 (1951).