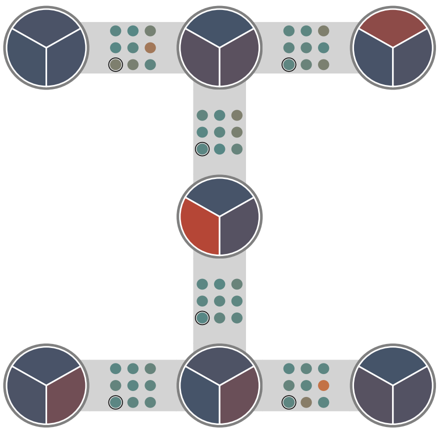
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Noise Characterization of Near-Term Quantum Devices

Master's Thesis



Noah Kaufmann

October, 2022

Trapped Ion Quantum Information Group
Institute for Quantum Electronics
Departement of Physics, D-PHYS
ETH Zürich

Supervisors:
Ivan Rojkov
Dr. Florentin Reiter
Prof. Dr. Jonathan Home

Title page

Illustration of the coherent errors that originate from executing a single gate-layer on a 7-qubit superconducting quantum processor of IBM Quantum. The characterization protocol applied to obtain this result is an integral part of this thesis.

Abstract

Noise affects many potential applications of current existing quantum processors. The characterization of near-term devices generates information about the sources of those disturbances. These insights facilitate error mitigation, error correction, and hardware improvements. However, characterization is challenging, as the total amount of potential noise processes becomes intractable with the increasing size of existing hardware platforms. Therefore, noise structure assumptions are necessary for the scalability of characterization protocols. This thesis introduces a new model approach for processes influenced by incoherent noise. Furthermore, we extend an existing incoherent noise model to coherent errors. We demonstrate the characterization and mitigation of the proposed model on a superconducting hardware platform. The developed characterization procedure to estimate coherent errors provides guidance for device calibration and future hardware development.

I want to thank the whole TIQI group, and especially Prof. Dr. Jonathan Home, for welcoming me into the group. I really enjoyed my time and got an insight into how much it takes to engineer new quantum computing systems. Moreover, the group retreat at Bettmeralp was truly a highlight of my master's. It was great being part of the theory subgroup, and I will miss the detailed and entangled discussions with David, Kristina, Elias, Ivan, and Florentin. I especially want to thank Dr. Florentin Reiter for contextualizing what I am doing, promoting this work beyond my reach, and finding a great balance between providing guidance and letting me explore different topics. Special thanks to Ivan Rojkov for simply being a fantastic supervisor! I benefited hugely from our discussions, his constructive feedback, and the countless inputs. Finally, much love to my flatmates, Eva and Nina, for constantly encouraging me, proofreading this thesis, and keeping track of the fan base.

Contents

Chapter 1

Introduction

From the wooden beads of the first abacuses to the electrons in novel quantum computers, the size of the fundamental computational unit has decreased by about ten orders of magnitude. From sliding around 10^{24} electrons with our bare hands, we have advanced to manipulating individual electrons with particles of light. And while even a trembling person can operate an abacus, already cosmic rays may interfere with the calculations on a quantum computer. In this evolution of shrinking system sizes, a major difficulty is to isolate the increasingly delicate computing units from disturbances. Quantum computers have yet to prove that they can sufficiently overcome this challenge and eventually allow us to compute tasks that so far were incalculable.

The current phase of quantum computing, where noise hinders many potential applications, is referred to as the noisy intermediate-scale quantum (NISQ) era [1]. On the path to building less error-prone quantum computers, it is critical to learn more about the sources of the disturbing processes and understand the limiting factors of available platforms better. Therefore, *characterization*, which denotes the analysis of existing devices, stands at the core of future technological progress.

1.1. Characterization

Current existing quantum computers are highly affected by errors. The errors originate from interactions of the quantum computer with its environment, unwanted dynamics between the qubits, or imperfect control signals. This is not a unique quantum problem, as classical computers are also exposed to such disturbances, and methods had to be developed to protect bits from the influence of external magnetic fields. However, unlike classical digital computers, the information in quantum computers is encoded as a continuous superposition of states. Therefore, while for classical digital computers the disturbances have to overcome some threshold to affect the state of the computer, in quantum computers disturbances can influence computations in a continuous fashion [2]. Characterization aims to find error sources and quantify their effect on the performance of a computing processor [3]. The information obtained from the characterization of a device can then be used for the development of a less error-prone hardware platform [4]. Moreover, the insights allow the implementation of schemes for correcting specific errors [5] and the application of other procedures that mitigate the effect of the noise [6]. Characterization procedures gain information about a process by analyzing the measurement statistic of an ensemble of circuits. The set of executed circuits and the applied post-processing steps are called *characterization protocol* [3].

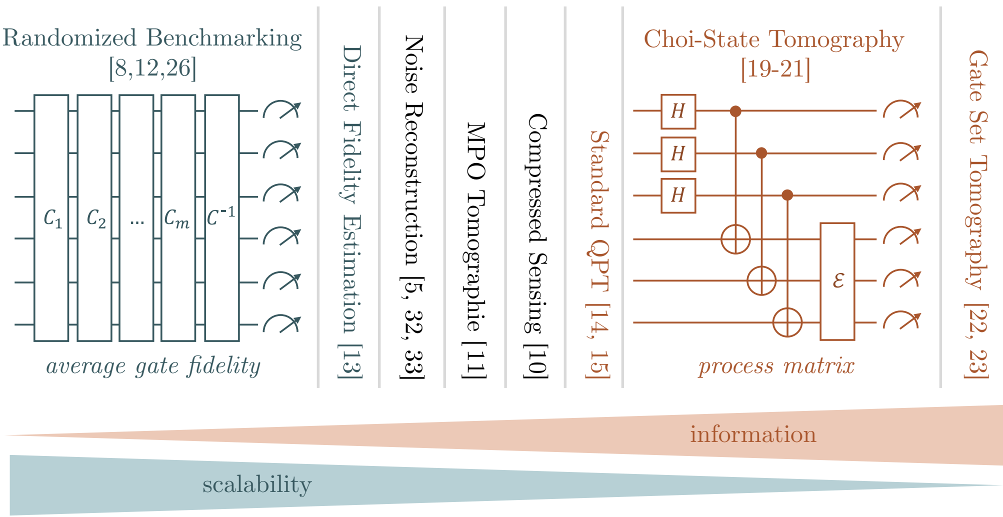


Figure 1.1.: *Characterization methods* - Different characterization methods are sorted relative to each other concerning the amount of information the characterization reveals and how scalable they are. Benchmark protocols are marked in blue, and full quantum process tomography protocols appear in orange. The black protocols introduce different assumptions on the process to lower the characterization's complexity.

Noise can be seen as some unknown process that overlaps the intended operations. A complete description of a general n-qubit process involves $\mathcal{O}(16^n)$ real values [7]. Hence, estimating a complete representation of a noise process is intractable for systems that can run interesting quantum computing algorithms. A characterization protocol is considered *scalable* if it can be used to characterize processes involving many qubits [8]. Towards scalable characterization protocols, there are two ways forward, either introduce a noise model of limited complexity and fit the real process to this model or limit the characterization to certain metrics of the process. Either way, it is a trade-off between limiting complexity and gaining information [9].

Protocols of the first approach introduce assumptions on the process structure to limit the number of degrees of freedom in the estimation. Such assumptions are, for example, that the noise can be expressed by a mixture of a few simple structured processes [10] or that the noise is only correlated among spatially close qubits [11]. Benchmark protocols, on the other side, focus on the estimation of heuristic properties of the system such as the average error rate of a gate [12] or the average fidelity in state preparation [13]. A sorted overview of different characterization methods is provided in Fig. 1.1. The following subsections introduce several protocols located at the extremes of the explained trade-off.

1.1.1. Quantum Process Tomography

A characterization protocol that leads to a process description that facilitates the estimation of the output of the process for any input state is called *quantum process tomography* (QPT) [7, 14, 15]. Full QPT refers to protocols with no prior assumptions about the structure of the process and they are, therefore, not scalable. Generally, QPT approaches can be divided into direct and indirect methods [16].

Indirect methods map information about the quantum process on a set of states and then estimate experimentally a description of this state to retain the properties of the process. Probably the most straightforward indirect approach is to map a quantum process to a state in a Hilbert space of a higher dimension [17, 18]. Various methods can then be applied to characterize the corresponding state [19–21]. Fig. 1.1 refers to this method as Choi-state tomography and shows an exemplary circuit scheme. Another example of indirect QPT is standard QPT [14, 15]. Introduced in 1997, it was the first QPT approach and can be seen as the starting point for quantum process characterization. In this protocol, states that form a complete basis of the corresponding Hilbert space are evolved by the examined process. From an estimated representation of the evolved states, the full description of the examined process is reconstructed [7]. One caveat of indirect QPT is that the required experiment runs of these protocols scale exponentially with the number of qubits. Furthermore, they rely on precise knowledge of the input and output states of the process and, consequently, are susceptible to state preparation and measurement errors (SPAM) [22].

Gate-set tomography (GST) [22, 23], on the other hand, does not rely on the accurate preparation of states, and is therefore considered a direct characterization method. It is by construction calibration-free, meaning SPAM error resistant [22]. Due to this advantage, GST has become the standard method for performing complete characterization of quantum processes [24]. GST completely characterizes a set of gates (used here as a synonym for operation) by running a protocol that includes sequences of different lengths of these gates [25]. In GST, the examined gates are applied not just once per circuit but multiple times. This technique amplifies the action of the gates and improves the estimation accuracy. As GST entirely characterizes an ensemble of operations, its complexity also scales exponentially with the number of qubits.

1.1.2. Randomized Benchmarking

Randomized benchmarking (RB) [8, 12, 26] is the most widely used benchmarking protocol. The characterization of a process is limited to estimating one or a few summary statistics. By applying similar concepts as GST, it is also resistant to SPAM errors. Most RB protocols execute circuits of various lengths and consist of randomly selected unitary gates (Fig. 1.1) [26]. The circuits are constructed such that they correspond to identity operations in the absence of noise meaning that the input state is left invariant. However, in the non-ideal case, the fidelity between the input and output states drops with increasing noise level. The average gate error rate is estimated by analyzing the decay of this fidelity with increasing circuit length [26]. While initially derived for benchmarking some specific, so called Clifford, gates [8], extensions to other operations exist [27, 28]. Since its proposal in 2005, various RB-related schemes have been developed that characterize different metrics [28–30]. As RB focuses on the characterization of single heuristics instead of a complete process description, the number of parameters to estimate does not scale exponentially with an increasing number of involved qubits. Therefore, RB protocols are generally scalable [12].

1.1.3. Noise Reconstruction

As stated above, the primary motivation for quantum process characterization is to gain physical insight into noise sources, in order to refine the hardware or improve error correction schemes. GST and full QPT schemes result in a complete process description, including all the noise, but they do not scale to the characterization of a whole gate layer on a NISQ device [31]. Furthermore, it is often difficult to link the results of QPT to specific noise sources. Conversely, RB protocols are scalable and provide physically intuitive results. However, the obtained information from RB does often not suffice for the mentioned applications of characterization [5].

An approach that seeks a middle ground between RB and full QPT is called *noise reconstruction* [9]. These schemes introduce a noise model with a limited set of parameters. They then assume that the real process can be modeled by concatenating the ideal process with this noise model. Existing protocols mainly focus on modeling the noise as a particular type of incoherent noise, called Pauli noise [5, 32, 33]. Incoherent noise corresponds to the loss of information stored in the qubits. This often arises from interactions with the environment. As not all noise can be modeled by Pauli noise, the mentioned protocols have to enforce their model assumption by applying a twirling technique [34] that projects arbitrary undesired processes onto Pauli noise without introducing new errors [5]. Then, the protocols applied to estimate the model parameters are similar to the schemes used in GST and RB and are also resistant to SPAM error [5, 32]. The noise reduction protocols of Refs. [32] and [5] are scalable under physically motivated assumptions on the noise structure [5] and facilitate error mitigation schemes [32]. However, their application of twirling eliminates certain noise components [31], which leads to an incomplete picture of the actual noise processes. This thesis contributes towards resolving this issue by expanding the noise reduction approach to a noise model involving coherent errors. Coherent errors are related to incomplete knowledge about the dynamics of the qubits. Coherent and incoherent errors affect the performance of quantum processors differently [35]. This observation further motivates the proposed model expansion.

1.2. Mitigation

In the NISQ era, where fault tolerance is not reached, and error correction schemes are hard to implement experimentally, *quantum error mitigation* is one approach to simulate quantum circuits on noisy devices [32]. Quantum error mitigation protocols aim to minimize the noise-induced bias in expectation values measured on noisy hardware platforms [36]. Such protocols usually contain involved post-processing schemes that combine the measurement statistic of an ensemble of modified circuits [36]. It is important to note that, different from error correction schemes, every single circuit of the ensemble is noisy and only the examination of the measurements of the whole ensemble in post-processing leads to a noise-free estimate of the expectation value [36]. The complexity of executing quantum error mitigation increases with increasing noise rates, and its scalability to deep circuits is, therefore, limited [36]. Examples of quantum error mitigation schemes are probabilistic error cancellation (PEC) [6], quantum error extrapolation [37], and stabilizer emulation [38]. Based on the characterized noise, this thesis applies quantum error mitigation techniques to prove the validity and usefulness of the proposed characterization methods.

1.3. Thesis Outline

This thesis aims to explore scalable noise characterization protocols that provide physical insights which help advance existing quantum computing technologies. We start in chapter **2** by introducing fundamental concepts of quantum information theory in the framework of Pauli operators. In chapter **3**, we present different noise models. We propose a new Pauli noise based model in section 3.1 that models the noisy process as a probabilistic mixture of the ideal process and a Pauli channel. Furthermore, we extend a model similar to the one used in Refs. [32] and [5] to coherent noise in section 3.2. Chapter **4** presents protocols to estimate the parameters of the introduced models. Our main new contributions are the extension of the Pauli channel estimation protocol in Ref. [39] to the combination of Pauli channel and decay, and the introduction of a characterization protocol for coherent noise. Finally, we demonstrate in chapter **5** the capability of the introduced coherent noise model to capture the main noise contributions of a process executed on an existing hardware platform. We do this by mitigating the noise based on the results of the characterization protocol. For this step, we develop a gate-level correction scheme for coherent errors.

Chapter 2

Fundamentals

The Bloch sphere picture represents a single qubit geometrically in $\mathbb{R}^3$. This representation can be used to create an intuitive illustration of the action of a channel on a qubit. While this geometrical picture loses its simplicity when more than one qubit is involved, the algebraic framework behind the interpretation remains a powerful tool. This chapter introduces a notation in terms of Pauli matrices for states and channels of single and multiple qubits.

2.1. Qubit States

Qubits are two-level systems. While there are many ways to implement qubits physically, the Hilbert space that represents their states mathematically is hardware-independent. This chapter elaborates on the density matrix formalism in the basis of the Pauli matrices for single-qubit and multi-qubit systems.

2.1.1. Single Qubit

An isolated single qubit is the simplest quantum mechanical system [7]. The state ψ of a single-qubit system has two degrees of freedom. They can be interpreted as the excitation of the qubit in some basis and its phase with respect to this basis. States of such qubit systems that are independent of the environment and exactly known are called pure states, and they are represented by vectors $|\psi\rangle$ in the Hilbert space $\mathcal{H}_2 = \mathbb{C}^2$. Partially unknown states or states of qubits that are entangled with the environment are called mixed states. They can not be expressed by a vector $|\psi\rangle \in \mathcal{H}_2$. Instead, they are described by a mixture of pure states. The density operator formalism describes the states of a qubit subsystem by mapping them to linear operators $\rho \in \mathcal{L}(\mathcal{H}_2)$ that act on the Hilbert space. The density operator of a pure state simply corresponds to $|\psi\rangle\langle\psi|$. Density operators of mixed states are expressed by some mixture $\sum_i p(i) |\psi_i\rangle\langle\psi_i|$. Unlike the statevector description of pure states, the set of valid density operators has three degrees of freedom. The additional degree of freedom reflects the purity of the state and describes to what extent the qubit is mixed. The density operators map to the positive-semidefinite complex Hermitian matrices of size two and trace one. This set of matrices can be expressed as a real superposition of the extended Pauli basis, consisting of the three Pauli matrices and the identity,

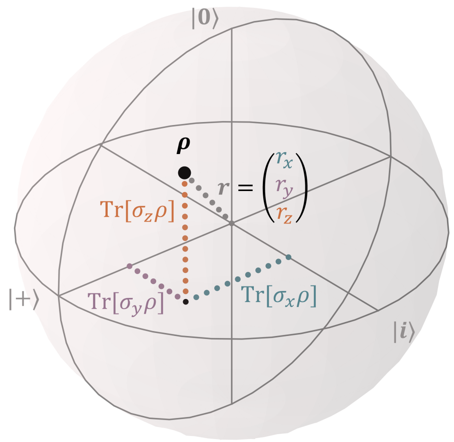


Figure 2.1.: *Bloch sphere representation* In the Bloch sphere representation of a single-qubit state ρ, the expectation values of the Pauli operators are used to represent ρ in $\mathbb{R}^3$. All physical states lie within a unit sphere, where pure states correspond to states on the surface of the sphere. $|+\rangle$, $|i\rangle$, and $|0\rangle$ are the eigenstates of the operators σ_x, σ_y, respectively σ_z.

$$\begin{aligned} \sigma_0 = \mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ \sigma_2 = \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned} \tag{2.1}$$

The expansion of the density matrix ρ in the Pauli basis,

$$\rho = \frac{1}{2}\left(\mathbb{I} + r_1\sigma_1 + r_2\sigma_2 + r_3\sigma_3\right) = \frac{1}{2}\left(\mathbb{I} + r \cdot \sigma\right), \tag{2.2}$$

creates a unique relation between a quantum state and a vector $r \in \mathbb{R}^3$, where σ is a vector containing the Pauli matrices. From the condition that density matrices must be positive-semidefinite, it follows that $|r| \leq 1$. Therefore, single-qubit states can be geometrically interpreted as points inside the unit sphere. Points on the sphere correspond to pure states, and points located inside the sphere represent mixed states. This relates to the convex nature of the set of states, with the pure states being the boundary of the set and the mixed states being a mixture of pure states. The completely mixed state is mapped to the origin of the sphere. This sphere is referred to as the Bloch sphere [40], shown in Fig. 2.1.

The Pauli matrices themselves represent operators in the Hilbert space. In the geometrical representation, their action on a particular state corresponds to the rotation around the corresponding axis by π. The points where the sphere intersects the positive x, y, and z-axis, mirror the pure states $|+\rangle$, $|i\rangle$, and $|0\rangle$. They are eigenstates of the Pauli operators σ_x, σ_y, respectively, σ_z with eigenvalue 1. The states $|-\rangle$, $|-i\rangle$, and $|1\rangle$, which are the eigenstates of the Pauli operators with eigenvalue -1, are mapped to the intersection of the sphere with the negative part of the axis.

Choosing one axis as the computational basis, the projection of r onto this axis corresponds to excitation of the state in the computational basis, the magnitude of r relates to the purity of the state, and the angle with respect to the chosen axis represents the phase of the state. Since the matrix product $\sigma_i\sigma_j$ is traceless for $i \neq j$ and the identity for $i = j$, the cartesian coordinates r_i of a state in the geometrical picture correspond to $r_i = \mathrm{Tr}[\sigma_i \rho]$.

In the case of the qubit modeling a spin system with $s = 1/2$, by the Born rule and the proportionality of the Pauli matrices to the spin operators, $r_i = \mathrm{Tr}[\sigma_i \rho]$ is directly related to the spin expectation values of the system.

2.1.2. Multiple Qubits

For systems of multiple qubits, no obvious generalization of the Bloch sphere picture exists. While multiple separate Bloch spheres can represent separable states, the representation of entangled states is more involved [41]. There are different approaches to illustrate states of multi-qubit systems geometrically [42], for example, using Gell-Mann matrices [43], Hopf fibrations [44], or steering ellipsoids [45, 46]. In this project, the applications focus on the characterization of Clifford circuits. As the conjugation with a Clifford unitary maps the group of Pauli operators to itself [47], the Pauli basis is convenient in the Clifford context. Therefore, this thesis continues to describe states of multi-qubit systems in the Pauli framework.

The Hilbert space that describes an n-qubit system is 2^n-dimensional. The density matrix representation of such a system consists of Hermitian matrices of dimension $d = 2^n \times 2^n$. A basis for those Hermitian matrices is given by all tensor products of length n between the extended Pauli matrices $P_i = \sigma_{i_1} \otimes \sigma_{i_2} \otimes ... \otimes \sigma_{i_n} = \bigotimes_{m=1}^{n} \sigma_{i_m}$, denoted by $\{P_i\}_{i \in \{0,1,2,3\}^n}$. Where $P_{\{0\}^n}$ corresponds to the identity. This thesis refers to these matrices as n-qubit Pauli matrices. For two qubits, this set of matrices is sometimes called Dirac-matrices [42].

Each multi-qubit state can be represented in the n-qubit Pauli basis according to,

$$\rho = \frac{1}{2^n} \left(\mathbb{I}_n + \sum_{i \in \{0,1,2,3\}^n \setminus \{0\}^n} r_i P_i \right). \tag{2.3}$$

The summation is over all n-qubit Pauli operators expect the identity because the coefficient of the identity is fixed such that the overall trace of the density matrix equals 1. Each multi-qubit state is therefore represented by a vector $r \in \mathbb{R}^{(4^n-1)}$. Unlike the single-qubit case, the positive-definite condition on ρ does not result in $|r| \leq 1$, and the state is no longer represented by a point in a unit hypersphere. Like in the single qubit case, the components of r correspond to $r_k = Tr[P_k \rho]$. Normalizing r with the size of the Hilbert space and including the identity leads to a vector commonly denoted as $|\rho) \in \mathbb{R}^{4^n}$ [33]. Its elements correspond to

$$|\rho)_k = \frac{1}{2^n} Tr[P_k \rho]. \tag{2.4}$$

From the trace condition on density operators follows that $|\rho)_0 = \frac{1}{2^n}$.

2.1.3. Classical shadow tomography

Although not directly connected to the Bloch sphere representation of states, this section introduces *classical shadow tomography* as it is used for the characterization protocol in section 4.1. Classical shadow tomography was initially introduced by Aaronson [48] and then advanced to a less hardware-demanding protocol by Huang *et al.* [49]. The central aspect of shadow tomography is that the number of measurements required to estimate the expectation values of m bounded norm operators only depends logarithmic on m and does not depend on the size of the Hilbert space of the state at all [49]. As the complexity of states grows exponentially with the system size, naively, one would expect that to determine a set of system properties, each property has to be measured on its own [50]. However, this is not true, as classical shadows prove that the length of a classical representation which suffices to predict m linear functions of the state only scales logarithmically in m [49].

As any quantum state tomography protocol, classical shadow tomography attempts to find a classical description of a quantum state that allows the prediction of different characteristics. In standard quantum state tomography, the classical description is created by measuring the state for a complete measurement set. Which then allows the reconstruction of the density matrix of the state. In classical shadow tomography, on the other hand, a classical shadow (or classical sketch) of the state is constructed by repeatedly evolving the state by a randomly sampled unitary and a measurement in the computational basis [49]. For each snapshot of the classical shadow, the information on the applied unitary and the measurement readout is stored. From multiple such pairs of information, many relevant properties of the state can be estimated without reconstructing the density matrix [49].

Classical shadows can, for example, be used for quantum fidelity estimation, entanglement verification, prediction of the expectation value of local observables, or also the estimation of non-linear properties of the system, such as the second-order Rényi entropy [49]. However, because the size of the required classical shadows to estimate general non-local properties grows exponentially with decreasing locality of the property, classical shadow tomography does not prove to be advantageous in the estimation of global observables, nor does it break the exponential scaling of full state tomography.

Classical shadows can also be used to efficiently predict all reduced density matrices of a specific size of a state. For this application, the number of necessary measurements T to ϵ-accurately estimate all reduced r-body density matrices of an n-qubit system is of order $\mathcal{O}\left(\text{const}^r \log n/\epsilon^2\right)$ [51]. This thesis uses classical shadows in the spirit of a quantum state tomography protocol to estimate reduced density matrices.

The protocol of estimating the density matrix of a state by classical shadow tomography contains the following steps [49]: first, draw a unitary from an ensemble of unitaries, then evolve a copy of the state ρ by the drawn unitary and measure the resulting state $U\rho U^\dagger$ in the computational basis. The resulting bitstring b is mapped to the computational pure state $|b\rangle\langle b|$, from which classically, the expression $U^\dagger |b\rangle\langle b| U$ is calculated. Averaging over all snapshots creates the quantum channel $\mathcal{M}(\rho) = \mathbb{E}\left[U^\dagger |b\rangle\langle b| U\right]$. For tomographically complete ensembles of unitaries, $\mathcal{M}$ is invertible, although not physical. In classical post-processing, $\hat{\rho} = \mathcal{M}^{-1}\left(U^\dagger |b\rangle\langle b| U\right)$ can be produced, which suffice to predict the

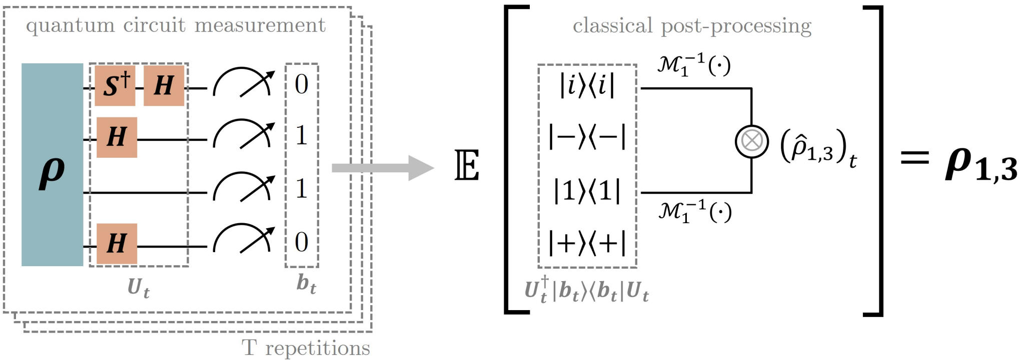


Figure 2.2.: *Classical shadow tomography* - Schematic of the classical shadow tomography protocol, illustrated at the example of estimating the reduced density matrix of qubits 1 and 3 for the 4-qubit state ρ. For each repetition t, a unitary evolution U_t is drawn. The state is then evolved by the unitary and measured in the computational basis. The bitstring b_t from the readout, together with the classical information on which unitary was applied, is called a single classical snapshot of the state. All T classical snapshots together correspond to the classical shadow of ρ. In classical post-processing, the reduced density matrix of qubits 1 and 3 is estimated for each snapshot. The expectation value of the average over all estimates $(\hat{\rho}_{1,3})_t$ corresponds to the exact reduced density matrix.

density matrix according to $\mathbb{E}[\hat{\rho}] = \rho$.

A particularly hardware-feasible version of this protocol is to select the unitaries from the set of tensor products of the single-qubit Clifford circuits $\{H, S^\dagger H, \mathbb{I}\}$, where H denoting the Hadamard-gate and S the Phase-gate. In the Bloch sphere picture, H and $S^\dagger H$ rotate the x-axis, respectively the y-axis, to the z-axis. Therefore, applying those unitaries in combination with a measurement in the computational basis corresponds to a Pauli measurement. In many existing hardware platforms, these unitaries can be implemented efficiently. As these unitaries are built from single-qubit operations, reduced r-body density matrices can be constructed by ignoring the data from all other sites. The random Pauli measurement leads to the inverse $\mathcal{M}_1^{-1}(X) = 3X + \mathbb{I}$ for a single qubit. The inverse for an r-body subsystem corresponds to the tensor product of the individual inverses $\mathcal{M}_r^{-1}(X) = \bigotimes_{i=1}^{r} \mathcal{M}_1^{-1}(X_i)$. Fig. 2.2 shows a schematic of the estimation protocol for the r-body density matrices.

2.2. Channels

A channel is a trace-preserving, completely positive, and linear map between operators of Hilbert spaces [7]. $\mathcal{E}_{B|A}$ denotes a channel that maps $\mathcal{L}(\mathcal{H}_A)$ to $\mathcal{L}(\mathcal{H}_B)$. From a quantum computational perspective, channels are, for example, the application of a gate, the coupling of a system to the environment, or the preparation of a state. Generally, the input and output of a channel are operators of different Hilbert spaces. For example, if a two-qubit state is prepared based on some instruction, a classical input is mapped on a quantum state in $\mathcal{L}\left(\mathbb{C}^4\right)$. However, this thesis focuses on the characterization of gate operations, and in this context, channels usually map the operators of a Hilbert space onto themselves $\mathcal{L}\left(\mathcal{H}_A\right) \to \mathcal{L}\left(\mathcal{H}_A\right)$.

There exist many different representations of channels. The *Kraus representation* is particularly useful in connection with the density matrix representation of states. For every channel, there exists a set of Kraus operators $\{K_i \in \mathcal{L}\left(\mathcal{H}_A, \mathcal{H}_B\right)\}_{i=1}^{m}$ with $\sum_i K_i^\dagger K_i = \mathbb{I}_A$, such that the action of the channel on a linear operator in $\mathcal{H}_A$ can be expressed [7] as

$$\mathcal{E}_{B|A}\left(\rho\right) = \sum_i K_i \rho K_i^\dagger. \tag{2.5}$$

Starting from the Bloch sphere representation of single-qubit states and the Kraus representation of channels, the following section discusses different allowed transformations of the Bloch sphere under the action of a channel and introduces the Pauli Transfer Matrix (PTM) formalism.

2.2.1. Geometric Interpretation of Single-Qubit Operations

By applying a single-qubit channel to all points within the Bloch sphere, it is possible to illustrate the action of the channel as a geometric transformation in $\mathbb{R}^3$. Fig. 2.3 shows the image of the Bloch sphere for different single-qubit operations. Since the Bloch sphere representation only shows the traceless part of the density matrix of a state, any Bloch sphere transformation is trace-preserving. Generally, the image of the Bloch sphere under the action of a channel is an ellipsoid within the initial Bloch sphere. However, due to the condition of complete positivity, not all such transformations are physical channels [52, 53].

Unitary channels have a unique Kraus representation with a single Kraus operator. As they correspond to a closed system evolution, they are not changing the purity of the state they are acting on. In the Bloch sphere picture, unitary operations are represented by a rotation of all points within the sphere around an axis through the origin (Fig. 2.3 c). Therefore, the image of the set of pure states is again the Bloch sphere. The Pauli operators are also unitary and correspond to a rotation of the Bloch sphere by π around the respective axis.

If a channel can be represented with a Kraus map where all Kraus operators are proportional to the Pauli operators, the channel is called a *Pauli channel*. Examples of Pauli channels are the depolarizing and the dephasing channel. Pauli channels are purity nonincreasing, meaning the distance from the sphere's origin to the representation of a state within the Bloch sphere is not increasing. In the Bloch sphere picture, the application of a Pauli channel corresponds to a scaling along the axes of the Pauli basis (Fig. 2.3 b). Each of the three axes can be scaled by a different scalar. The image of the

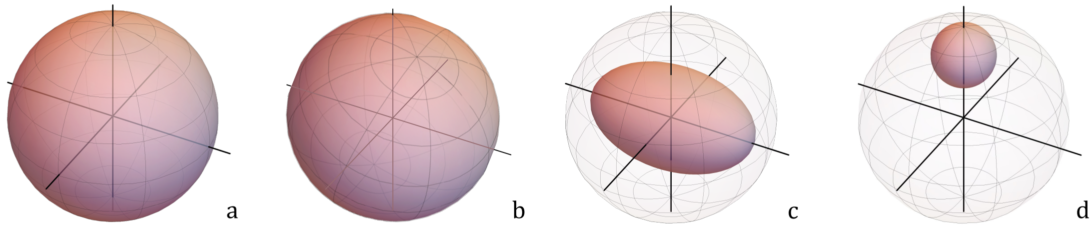


Figure 2.3.: *Bloch sphere transformations* - Image of the Bloch sphere under the action of different single-qubit channels: **a** identity, **b** unitary operation, **c** Pauli channel, **d** amplitude damping

pure states is an ellipsoid with principle axes parallel to the Pauli basis. Written as,

$$\mathcal{P}(\rho) = \sum_c p(c) P_c \rho P_c, \tag{2.6}$$

with $\sum_c p(c) = 1$, and $p(c) \geq 0$ (enforcing complete positivity [33]), it is evident that a Pauli channel is a probabilistic mixture of the application of Pauli operators. In the framework of noise characterization, the set of $p(c)$, which forms a probability distribution over all Pauli operators, are called *Pauli error rates*. The Pauli error rates $p(c)$ are not the same as the scaling factors of the Bloch sphere, called *Pauli channel eigenvalues* [33]. Combining unitary operations and Pauli channels makes it possible to construct a scaling of a Bloch sphere along arbitrary perpendicular axes.

As unitary and Pauli channels correspond to the rotation respectively, the scaling of states in the Bloch sphere picture, the fully mixed state, represented at the sphere's origin, is left unchanged. Such channels are called unital [25]. An example of a common non-unital channel is the amplitude damping channel (Fig. 2.3 d). In the Bloch sphere picture, this channel corresponds to the scaling of all points within the sphere around the point that corresponds to the state $|0\rangle$.

2.2.2. Pauli Transfer Matrix Representation

In the Bloch sphere picture, the action of a channel $\mathcal{E}$ on a single qubit is described by

$$\mathcal{E}\left(\frac{1}{2}\left(\mathbb{I} + r \cdot \sigma\right)\right) = \frac{1}{2}\left(\mathbb{I} + (t + Ar)\,\sigma\right). \tag{2.7}$$

The matrix A equals to the linear transformation of the Bloch sphere, and the vector t represents the translation of the Bloch sphere. For example, a Pauli channel is represented by a diagonal matrix A and a zero vector t since no translation is involved.

This approach can be generalized by using the notation introduced in section 2.1.2. A channel $\mathcal{E}$ acting on a state ρ can be denoted by $|\mathcal{E}(\rho)\rangle\rangle = T_{\mathcal{E}}\,|\rho\rangle\rangle$. $T_{\mathcal{E}}$ is called the *Pauli Transfer Matrix* (PTM) representation of $\mathcal{E}$ [25]. The matrix is of dimension

$4^n \times 4^n$ and consists of the elements $T_{\mathcal{E}}[i,j] = \frac{1}{2^n}\mathrm{Tr}\left[P_i \mathcal{E}(P_j)\right]$. The general structure of $T_{\mathcal{E}}$ corresponds to

$$T_{\mathcal{E}} = \begin{pmatrix} 1 & 0 \\ t & A \end{pmatrix}. \tag{2.8}$$

The first row of the matrix ensures that the channel is trace-preserving. The matrix A corresponds to the unital part of the channel, and t is associated with the nonutiality of the channel [30].

In the following, the PTMs of the channels discussed in section 2.2.1 are presented. Like the PTM of every Pauli channel, the PTM of the dephasing channel $\mathcal{Z}$ in Eq. (2.9) and the PTM of the depolarization channel Δ in Eq. (2.9) are diagonal,

$$T_{\mathcal{Z}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1-\lambda & 0 & 0 \\ 0 & 0 & 1-\lambda & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad T_{\Delta} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1-\lambda & 0 & 0 \\ 0 & 0 & 1-\lambda & 0 \\ 0 & 0 & 0 & 1-\lambda \end{pmatrix}. \tag{2.9}$$

In the context of Pauli channels, the diagonal elements of the unital part of the matrix are referred to as Pauli eigenvalues. They correspond to the scaling of the Pauli expectation values of a state under the action of the channel. The Pauli eigenvalues are related to the Pauli error rates of the channel by a Walsh-Hadamard transform [33]. Since all Pauli eigenvalues of the depolarizing channel are equal, the Bloch sphere contracts uniformly under its action. In contrast, the dephasing channel only changes the components of the state that relate to σ_x and σ_y, which correspond to the off-diagonal terms of the density matrix.

A single-qubit rotation $\mathcal{R}_y$ around the y-axis by angle θ corresponds to a unitary evolution with the operator $e^{i\frac{\theta}{2}\sigma_y}$. The PTM

$$T_{\mathcal{R}_y} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & 0 & \sin\theta \\ 0 & 0 & 1 & 0 \\ 0 & -\sin\theta & 0 & \cos\theta \end{pmatrix} \tag{2.10}$$

has the form of a rotation matrix. Lastly, the amplitude damping channel Γ with decay parameter γ is an example of a non-unital channel, as visible from the first column in the PTM,

$$T_{\Gamma} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \sqrt{1-\gamma} & 0 & 0 \\ 0 & 0 & \sqrt{1-\gamma} & 0 \\ \gamma & 0 & 0 & 1-\gamma \end{pmatrix}. \tag{2.11}$$

2.2.3. Pauli Twirling

Pauli Twirling is a technique to transform an arbitrary channel into a Pauli channel, (cf. Eq. (2.6)). In terms of the PTM formalism, the effect of twirling is the cancellation of all off-diagonal terms. Twirling is achieved by averaging over the set of Pauli operators, where the input and the output of the channel are conjugated by the same Pauli operator.

$$\tilde{\mathcal{E}}_T(\rho) = \mathbb{E}_i \left[P_i \mathcal{E}(P_i \rho P_i) P_i \right], \tag{2.12}$$

shows the Pauli twirled channel $\tilde{\mathcal{E}}_T$ that originated from channel $\mathcal{E}$. The PTM $T_{\tilde{\mathcal{E}}_T}$ of the twirled channel $\tilde{\mathcal{E}}_T$ is diagonal with the elements $T_{\tilde{\mathcal{E}}_T}[i,j] = \delta_{i,j} \operatorname{Tr}[P_i \mathcal{E}(P_j)]$. If applied to a noise channel, the transformation is helpful for the mitigation of quantum noise, as it reduces the noise to a channel with a clear structure and lower complexity [32, 54].

By twirling with the set of Clifford gates instead, the channel transforms into a depolarization channel with the same Pauli rates (cf. Eq. (2.9)). This further reduces the complexity of the channel to a total of 2^n parameters [5].

2.3. Pauli Algebra

The Pauli matrices of equation 2.1 are named after Wolfgang Pauli, who used them to describe spin systems [55]. In the Bloch sphere picture, the Pauli matrices can be interpreted as rotations of angle π around a specific axis. Hence it is fitting that the algebra generated by the Pauli matrices is called a geometric (Clifford) algebra. Clifford algebras extend the algebra of real numbers to vectors and are closed under addition and multiplication [56]. The simplest non-trivial Clifford algebra is isomorphic to the algebra of complex numbers, which creates an associative and commutative algebra for vectors in $\mathbb{R}^2$ [57]. Another well-known Clifford algebra, often used for the representation of 3D rotations, is generated by the quaternions. The Pauli algebra, generated by the Pauli matrices, is isomorphic to the Clifford algebra of $\mathbb{R}^3$ [56]. Since the Pauli matrices are extensively used in this thesis, this section introduces some of the properties of the underlying geometric algebra.

In Clifford algebras, it is fundamental that the product of a vector with itself is proportional to the identity element [56]. Vectors v created by linear combinations of Pauli matrices $v = a\sigma_1 + b\sigma_2 + c\sigma_3$ fulfill this property under the standard matrix multiplication as the Pauli matrices mutually anti-commute

$$\{\sigma_k, \sigma_l\} = \sigma_k\sigma_l + \sigma_l\sigma_k = 2\delta_{k,l}, \tag{2.13}$$

and the product with themselves corresponds to the identity $\sigma_k\sigma_k = \mathbb{I}$.

The 3D-rotation representation of the action of the Pauli matrices in the Bloch sphere picture intuitively reveals their non-commutativity

$$[\sigma_k, \sigma_l] = 2i \sum_m \epsilon_{klm}\sigma_m. \tag{2.14}$$

The products between the Pauli matrices and the identity are summarized in table 2.1.

	$\mathbb{I}$	σ_x	σ_y	σ_z
$\mathbb{I}$	$\mathbb{I}$	σ_x	σ_y	σ_z
σ_x	σ_x	$\mathbb{I}$	$i\sigma_z$	$-i\sigma_y$
σ_y	σ_y	$-i\sigma_z$	$\mathbb{I}$	$i\sigma_x$
σ_z	σ_z	$i\sigma_y$	$-i\sigma_x$	$\mathbb{I}$

Table 2.1.: Operation table of the Pauli matrices and the identity. The operators inside the table correspond to the products of the operators indicating the row with the operators indicating the column.

In the n-qubit Pauli basis $\{P_i\}_{i\in\{0,1,2,3\}^n}$, it is useful to introduce the symplectic inner Product $\langle k,l\rangle_{sp}$, which is 0 if the Pauli operators P_k and P_l commute and 1 otherwise [32]. This notation is extended to the elementwise product [39]

$$k\star l=\{\langle k_0,l_0\rangle_{sp},\langle k_1,l_1\rangle_{sp},...,\langle k_n,l_n\rangle_{sp}\in\{0,1\}^n. \tag{2.15}$$

Note that all n-qubit Pauli operators commute with exactly half of the n-qubit Pauli operators and anticommute with the other half, except for the identity, which commutes with all other Pauli operators. The product P_kP_l can be expressed as $-P_lP_k$ in the anticommuting case or as P_lP_k in the commuting case. In a more compact form $P_kP_l=(-1)^{\langle k,l\rangle_{sp}}P_lP_k$ is valid and more important in the course of this thesis

$$P_kP_lP_k=(-1)^{\langle k,l\rangle_{sp}}P_l. \tag{2.16}$$

With this property one can establishes the relation between the Pauli error rates $p(c)$ and the Pauli eigenvalues $f(b)$ of a Pauli Channel $\mathcal{P}$, as introduced in sections 2.2.1 and 2.2.2,

$$\begin{aligned} f(b)&=\frac{1}{2^n}\mathrm{Tr}\left[P_b\mathcal{P}\left(P_b\right)\right]=\frac{1}{2^n}\mathrm{Tr}\left[P_b\sum_c p(c)P_cP_bP_c\right]\\ &=\frac{1}{2^n}\sum_c p(c)(-1)^{\langle c,b\rangle_{sp}}\mathrm{Tr}\left[P_bP_b\right]=\sum_c p(c)(-1)^{\langle c,b\rangle_{sp}}. \end{aligned} \tag{2.17}$$

The transformation in Eq. (2.17) is called Walsh-Hadamard transform [33] and corresponds to a discrete n-dimensional Fourier transform [58].

Chapter 3

Models

This thesis aims to characterize the noise introduced by a quantum operation. The characterization protocol should not simply reveal a single summarizing noise rate that indicates how much the noise disturbs the ideal process but instead provide a physically insightful model of the active noise channels. For general unstructured noise channels, this task coincides with full quantum process tomography, and such protocols are not scalable to many qubits [39]. Therefore, we must make assumptions on the noise structure that lead to a model where the number of parameters does not scale exponentially with the number of qubits.

A quantum system is often exposed to various noise sources, such as interactions with the environment or between qubits, imperfect calibration, or leakage [59]. One general noise classification is the distinction between coherent and incoherent noise [60]. *Coherent noise* is the noise that corresponds to an undesired unitary evolution and preserves the purity of a state. Therefore, coherent noise does not lead to a loss of information in the qubit system. Typically, coherent noise is interpreted as small systematic over- or under-rotations [61] and can be dealt with by calibration. On the other hand, *Incoherent noise*, often arises from an interaction of the computing system with its environment. It does not preserve purity, and error correction is the standard way to suppress its effects. In the Bloch sphere picture, coherent noise is mirrored by a sphere rotation, while incoherent noise corresponds to a mixture of shrinking, rotation, and translation of the sphere. These two noise classes lead to very different error distributions [35]. For the same overall error rate, coherent noise can lead to a much bigger worst-case error (diamond distance of noisy process to ideal process) than incoherent noise [60]. This difference has, for example, practical implications in error correction, where the fault-tolerant threshold heavily depends on the type of noise [62].

We focus on Pauli Noise to treat incoherent noise and use low-angle qubit rotations to model the effect of coherent noise. This chapter further explains these different noise channels, introduces two main models on how to model noise around an ideal process, and discusses the assumptions behind these models.

3.1. Pauli Noise

Pauli noise can be represented as a Pauli channel (cf. Eq. (2.6)). This includes, for example, dephasing, which acts according to

$$\mathcal{Z}_\lambda(\rho) = (1-\lambda)\ \rho + \lambda\ \mathrm{diag}(\rho) = (1-\frac{\lambda}{2})\ \mathbb{I}\ \rho\ \mathbb{I} + \frac{\lambda}{2}\ \sigma_z\ \rho\ \sigma_z \tag{3.1}$$

for a single qubit and a dephasing parameter λ. The modeling of incoherent noise as Pauli noise is a common approach because depolarization and dephasing often dominate the noise introduced by the interaction of a qubit system with its environment [63, 64]. Furthermore, quantum error correction leads to noise that is better approximated by a Pauli channel [33]. It is also possible to enforce the Pauli noise assumption on general noise by applying twirling (cf. section 2.2.3), randomized compiling [65, 66], or Pauli frame randomization [67]. Lastly, protocols that estimate Pauli errors are realizable on existing quantum computing hardware with low-depth circuits, as the preparation of Pauli eigenstates and the measurement of Pauli expectation values often only require a few native gates. One possibility to increase the expressibility of a Pauli noise model is to extend Pauli channels by incorporating other Clifford gates, such as the Hadamard-gate (H) or the controlled-NOT-gate (CX), into the stochastic mixture [68].

For an n-qubit system, a general Pauli channel includes 4^n Pauli matrices and equally many Pauli error rates. This exponential scaling in the number of parameters necessitates further assumptions on the structure of the Pauli noise channel. To reach a model that is tractable for more than a few qubits, one option is to assume that the Pauli channel is sparse, meaning that most of the Pauli error rates are close to zero [39]. Another approach is to approximate the Pauli error rates with a Gibbs random field [5]. By choosing a factor graph that corresponds to the connectivity tree, shown, for example, in Fig. 5.5, of the computing hardware, this approach implies the assumption that the correlation of the noise is spatially local. A similar assumption, which we will take in this thesis, is to restrict the Pauli noise model to single and two qubit Pauli matrices [32]. This constraint limits the number of Pauli errors to $\mathcal{O}(n^2)$. Note that a model with such restrictions still includes cross-talk between all qubit pairs. A further reduction is possible by taking the topology of the computational hardware into account and considering only spatially local noise correlations.

3.1.1. Parallel Pauli Noise Model

A possible notion of a noisy channel $\mathcal{E}_{N_p}$ that incorporates incoherent Pauli noise is given by

$$\mathcal{E}_{N_p}(\rho) = (1-\eta)\ U_I \rho U_I^\dagger + \eta \sum_c p(c) P_c \rho P_c. \tag{3.2}$$

We refer to this model as the *parallel Pauli noise model*. In this model, U_I represents the ideal channel, and the noise is modeled by a Pauli channel acting on the initial state ρ. As illustrated in Fig. 3.1, the modeling channel equals a probabilistic mixture of the ideal channel and a Pauli channel acting on ρ. The parameter $\eta \in [0, 1]$ corresponds to the probability that the noise modeling Pauli channel is applied and can be associated with the process error rate. The case $\eta = 0$ corresponds to the ideal case without noise. As there are two types of parameters in this model, η and the Pauli error rates $p(c)$, the question arises on whether there are different parameter combinations that lead to the same model. The only unitary processes that are Pauli channels are single Pauli operations. Therefore, for a general U_I, each combination of η and Pauli error rates $p(c)$

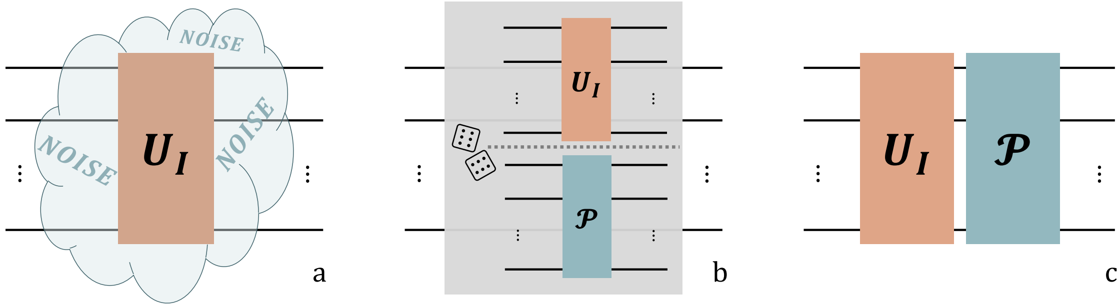


Figure 3.1.: *Pauli noise models* - **a** When implemented on real hardware, an ideal process U_I will always introduce noise. One approach is to model the action of the noise as a Pauli channel $\mathcal{P}$. **b** shows the parallel Pauli noise model, where the real process is modeled as a probabilistic mixture of the ideal process and a Pauli channel. In the sequential Pauli noise model **c**, the noise modeling Pauli channel is concatenated with the ideal process.

creates a unique noise model. For the special case where U_I is a Pauli operator P_c, we require the corresponding $p(c)$ to be 0. As mentioned in the last section, we will restrict the model to single- and two-qubit Pauli operators to eliminate the exponential scaling in the number of parameters in the model with an increasing number of qubits.

The motivation behind the parallel Pauli noise model is that the noise contributions are easy to interpret physically. This is favorable if the characterization results are used to improve the computing hardware. The central assumption of the model is that in a noisy process, either the ideal process happens or some other process that a Pauli channel can approximate. However, it is unclear how well such a model can approximate real processes on existing quantum computers, and there is no existing work for this model. Like any non-universal model, this model is certainly not useful for all cases of noise. Let us, for example, assume we want to characterize a 4-qubit GHZ preparation process, which starts from the state $|0000\rangle$ and prepares the state $\frac{1}{\sqrt{2}}(|0000\rangle + |1111\rangle)$. In the real process, however, with some probability, a bit-flip error (σ_x-error) occurs on the last qubit, and the prepared state ends up being a mixture between the GHZ-state and the state $\frac{1}{\sqrt{2}}(|0001\rangle + |1110\rangle)$. This noisy state preparation process already exceeds the expressibility of the parallel Pauli noise model, even more under the assumption that the model only includes single and two-qubit Pauli operators.

Decay is a type of incoherent noise that a Pauli channel fundamentally fails to express as it is nonunital [68]. In the chapter on characterization, we will also discuss the characterization of an extended model,

$$\mathcal{E}_{N_p^\gamma}(\rho) = \mathcal{F}_\gamma\left((1-\eta)\; U_I \rho U_I^\dagger + \eta \sum_c p(c) P_c \rho P_c\right), \tag{3.3}$$

where a decay channel $\mathcal{F}$ with decay rate γ acts in sequence with the noise model.

3.1.2. Sequential Pauli Noise Model

A widely used noise model [5, 9, 32] is the *sequential Pauli noise model.* In this model,

$$\mathcal{E}_{N_s}(\rho) = \sum_c p(c) P_c U_I \rho U_I^\dagger P_c, \tag{3.4}$$

the Pauli channel modeling the noise acts in sequence with the ideal channel (Fig. 3.1). For a noiseless operation, the Pauli error rate corresponding to the identity, $p(\{0\}^n)$, equals 1, while all other Pauli error rates are 0. For the characterization, we will again limit the model to single and two-qubit Pauli operators. The PTM formalism is advantageous for the treatment of the model, as the concatenation of channels corresponds to the product of PTMs. Furthermore, in the practically relevant case of U_I being a Clifford unitary, the PTM of U_I is a permutation matrix. This structure is preserved in the noise model, as the PTM of a Pauli channel is diagonal.

In the sequential Pauli noise model, the noisy GHZ preparation example of the last section can be modeled with single-qubit Pauli operators. The Pauli error rate corresponding to σ_x on the last qubit equals the error probability. The Pauli error rate connected to the identity corresponds to one minus the error probability.

3.2. Coherent Noise

Like quantum computing operations, coherent noise corresponds to a unitary evolution of the system. Thus coherent noise is associated with an incorrect knowledge of the system dynamic [69]. While Pauli noise is referred to as stochastic noise, coherent noise is also called systematic errors and is often connected to imperfect control of the quantum operation [61]. In contrast to the shrinking of the Bloch sphere being an insightful illustration of Pauli noise, coherent noise can be seen as small over- or under-rotations of the ideal process in the Bloch sphere picture. Therefore, the characterization of general coherent noise corresponds to the characterization of a general quantum computing process and is consequently of similar complexity as full QPT [33]. As coherent errors preserve the purity and do not lead to information loss, it is conceptually possible to undo the unwanted dynamics on the gate level. Common sources of coherent errors are crosstalk and other unitary dynamics during the idle time of a qubit. [32].

Unitary operators can be expressed as $U = e^{-iH}$, where H is a Hermitian matrix. The Pauli basis is a basis for the Hermitian matrices, and H can be written as $H = \sum \theta_k P_k$. For a general unitary operator, θ is a vector with 4^n elements. Under the assumption of small θ values, $U_\theta = \exp\left(\sum -i\theta_k P_k\right)$ is approximated to second order in θ by the expression $U_\theta \approx \prod \exp\left(-i\theta_k P_k\right)$. In the Bloch sphere picture, the terms $\exp\left(-i\theta_k P_k\right) = \cos\theta_k \, \mathbb{I} - i \sin\theta_k \, P_k$ represent rotations around the k-axis by angle θ. This interpretation relates to the earlier statement that coherent errors can be seen as over- or under-rotations. In order to reach a tractable model, similar to the approach chosen for the Pauli noise, we only consider one and two-qubit rotations. The integration of the coherent noise term

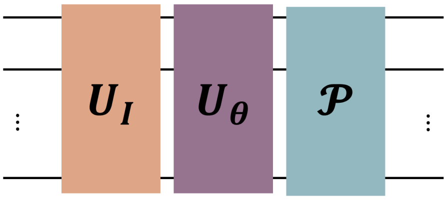


Figure 3.2.: *Coherent noise model* - Coherent noise model consisting of a concatenation of the ideal process U_I, a unitary evolution U_θ modeling the coherent part of the noise, and a Pauli channel $\mathcal{P}$ modeling the incoherent part of the noise.

into the sequential Pauli model is formalized by

$$\mathcal{E}_{N_c}(\rho) = \sum_c p(c) P_c U_\theta U_I \rho U_I^\dagger U_\theta^\dagger P_c. \tag{3.5}$$

In this thesis, we will refer to this model, illustrated in Fig. 3.2, as the *coherent noise model*.

An ongoing discussion takes place on the influence of coherent noise on error correcting schemes and the fault tolerance threshold. In this context, a similar noise model to Eq. (3.5) is used [62, 70, 71]. In Refs. [70] and [62], the model includes dephasing noise and coherent Z-rotations, as they consider them to be physically the most relevant. Their model $\mathcal{N}(\rho) = e^{-i\theta Z}\left[(1-p)\rho + pZ\rho Z\right] e^{i\theta Z}$ applies coherent noise and Pauli noise in a sequence similar to the approach of this thesis. In Ref. [71], on the other hand, where the model includes X-rotations and Pauli X errors, the coherent noise is applied in parallel to the coherent noise according to $\mathcal{N}(\rho) = \eta e^{-i\theta Z} \rho e^{i\theta X} + (1-\eta)\left[(1-p)\rho + pX\rho X\right]$. These works do not focus on the characterization of a noisy process but rather on analyzing the properties of a process with the respective noise model.

Chapter 4

Characterization

In the spirit of noise reconstruction (cf. section 1.1.3), we proceed to characterize processes based on the models introduced in chapter 3. We aim to derive a protocol for each model to estimate the parameters for which the model resembles the real process best. We presume that we can repeatedly apply the noisy process in an experiment and that the process can be integrated into a circuit. The starting point for the derivation of the protocols is the assumption that the model can express the real process, and the knowledge of the model included ideal unitary evolution U_I (cf. equations (3.2), (3.4), and (3.5)). The protocols should be scalable, both in the size of the executed circuit ensemble and the required post-processing steps. Furthermore, as we intend the execution of the protocols to be feasible on existing hardware, only low-depth state preparation and measurement circuits shall be used.

Notation

In the developed protocols, we only apply state preparation schemes that prepare from the ground state a separable state, where each qubit is in an eigenstate of a specific Pauli operator. Hence, each qubit is prepared in one of the states $|+\rangle$, $|-\rangle$, $|i\rangle$, $|-i\rangle$, $|0\rangle$, $|1\rangle$. Those states can be prepared by a combination of X-gate, phase-gate (S), and Hadamard-gate (H). According to the notation in Ref. [39], we denote such an n-qubit state as χ_b^a, where $a \in \{1, 2, 3\}^n$ labels the preparation basis for each qubit, 1 corresponding to an eigenstate of σ_x, 2 to an eigenstate of σ_y, and 3 to an eigenstate of σ_z. The bitstring $b \in \{0, 1\}^n$ indicates whether the state is an eigenstate of eigenvalue 1 (for $b_k = 0$) or -1 (for $b_k = 1$) of the corresponding Pauli operator. We refer to b as the excitation of the state. For example, $\chi_{\{0,0\}}^{\{1,2\}} = |+\rangle \otimes |i\rangle$ and $\chi_{\{1,0\}}^{\{1,2\}} = |-\rangle \otimes |i\rangle$. We will write $\chi_{\{0\}^n}^a$ as χ_0^a.

4.1. Parallel Pauli Noise

The parameters in the parallel Pauli noise model $\mathcal{E}_{N_p}$, from Eq. (3.2), are the process error rate η and the Pauli error rates $p(c)$. In this section, we first discuss the estimation problem of each part of the model separately and then move on to the full characterization protocol, including simulation results. Furthermore, we show how to extend the presented protocol to the decay extended parallel Pauli noise model of Eq. (3.3).

4.1.1. Process Error Rate Estimation

The main idea for estimating the process error rate η can be illustrated in the Bloch sphere picture. While a Pauli channel only squeezes the Bloch sphere, the ideal unitary will generally rotate the sphere. Therefore, if we prepare some state $|\chi_0^a\rangle$ in a Pauli basis a, then act on it with a Pauli operator P_c, the state is still in the basis of a and can be expressed as $\left|\chi_{\{a\star c\}}^a\right\rangle$, where $a \star c$ denotes the elementwise commutation product introduced in section 2.3. Consequently, measuring the expectation value of this state for some Pauli operator P_b with $a \star b \neq \{0\}^n$ will always be 0, as stated in

$$\mathrm{Tr}[P_b P_c \chi_0^a P_c] = (-1)^{\langle a,c\rangle_{sp}} \mathrm{Tr}[P_b \chi_0^a] = \begin{cases} (-1)^{\langle a,c\rangle_{sp}}, & \text{if } a \star b = \{0\}^n \\ 0, & \text{otherwise} \end{cases}. \tag{4.1}$$

For example, if we prepare the state $|+\rangle$ and act on it with Pauli operator σ_y, we reach, up to a phase, the state $|-\rangle$. The expectation value of $|-\rangle$ is 0 for σ_y and σ_z, 1 for the identity, and -1 for σ_x. This property is well reflected in the diagonal PTM structure of a Pauli channel and follows from Eq. (2.16).

Next, we consider the evolution of the state χ_0^a with a unitary operator U_I instead of a Pauli operator. If U_I does not correspond to a Pauli operator, the PTM of U_I contains off-diagonal terms. This implies the existence of a combination of state preparation $|\chi_0^a\rangle$ and Pauli P_b, with $a \star b \neq \{0\}^n$, such that

$$\xi_{a,b} := \mathrm{Tr}[P_b U_I \chi_0^a U_I^\dagger] \neq 0. \tag{4.2}$$

Combining this result with Eq. (4.1), still considering $a \star b \neq \{0\}^n$, the expectation value of the state χ_0^a after applying the noisy channel is equal to

$$\underbrace{\mathrm{Tr}[P_b \mathcal{E}_{N_p}(\chi_0^a)]}_{\langle P_b\rangle_{\mathcal{E}_{N_p}(\chi_0^a)}} = (1-\eta) \underbrace{\mathrm{Tr}[P_b U_I \chi_0^a U_I^\dagger]}_{\xi_{a,b} \text{ (Eq. (4.2))}} + \eta \sum_c p(c) \underbrace{\mathrm{Tr}[P_b P_c \chi_0^a P_c]}_{0 \text{ (Eq. (4.1))}}. \tag{4.3}$$

Rearranging, for nonzero $\xi_{a,b}$, leads to the expression

$$\eta = 1 - \frac{\langle P_b\rangle_{\mathcal{E}_{N_p}(\chi_0^a)}}{\xi_{a,b}}. \tag{4.4}$$

As the terms on the right-hand side depend on the ideal circuit $\xi_{a,b}$ or can be measured experimentally $\langle P_b\rangle_{\mathcal{E}_{N_p}(\chi_0^a)}$, this equation can be used to estimate η.

$\langle P_b\rangle_{\mathcal{E}_{N_p}(\chi_0^a)}$ is estimated by repeatedly measuring a circuit that prepares the state χ_0^a, applies the noisy channel, and does a Pauli measurement in the basis of b (shown in Fig. 4.1). The expectation value of P_b corresponds to $\mathbb{E}\left[(-1)^r\right]$, where r is the circuit readout.

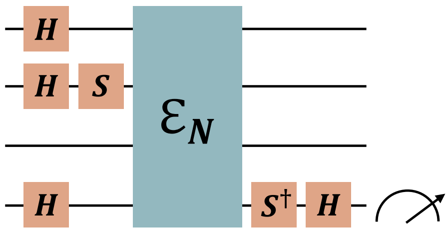


Figure 4.1.: *η-estimation circuit* - Example of a circuit to estimate η. The first gate-layer prepares the state $\chi_0^{a=\{1,2,3,1\}}$. Then the noisy implementation $\mathcal{E}_N$ of the ideal process is applied. Finally, qubit 4 is measured in the $y-$basis in order to estimate the expectation value of the operator $P_{b=\{0,0,0,2\}} = \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I} \otimes \sigma_y$.

By introducing the operator $\hat{\eta} := 1 - \hat{P}_b/\xi_{a,b}$, we can formalize the estimation of η as $\mathrm{Tr}\left[\hat{\eta}\chi_0^a\right]$. This allows evaluating the variance in the measurement process to

$$\sigma_\eta^2 = \frac{1}{\xi_{a,b}^2} - (1-\eta)^2. \tag{4.5}$$

The scaling of the expected estimation error with the number of qubits in the system, ξ, and the number of measurements is shown in Fig. A.2 and confirms Eq. (4.5).

In order to create a favorable measurement statistic, we are interested in choosing the state χ_0^a and operator P_b such that $\left|\xi_{a,b}^2\right|$ is maximal. The respective P_b will be of structure $\{\mathbb{I}, ..., \mathbb{I}, \sigma_i, \mathbb{I}, ..., \mathbb{I}\}$, as any additional non-identity Pauli matrix would decrease the expectation value $\mathrm{Tr}[P_b U_I \chi_0^a U_I^\dagger]$. We have not found an efficient way to optimize over a and b. Instead, we calculate $\xi_{a,b}$ for randomly pick a and b and select the best combination.

4.1.2. Pauli Channel Estimation

This section focuses on the estimation of the error rates of a Pauli channel, as shown in Eq. (2.6). The integration of this Pauli channel characterization protocol into the characterization protocol of the parallel Pauli noise model will be discussed in the next section. The presented protocol was developed by Flammia *et al.* [39], and we closely follow their derivation of the protocol.

For understanding the protocol, it is most helpful to interpret the Pauli channel as a black box with some hidden probability distribution $p(c)$. The goal of the protocol is to estimate this distribution. We can interact with the black box by inputting any state. The black box then chooses a Pauli operator P_c according to its probability distribution and evolves the state by this operator. Afterward, we can further process the evolved state. In this framework, the question arises of which state we should input into the black box and how to measure the output state to gain information on the applied Pauli operator. To answer this question, we regard the property

$$\mathcal{P}(\chi_0^a) = \sum_c p(c) P_c \chi_0^a P_c = \sum_c p(c) \chi_{a\star c}^a. \tag{4.6}$$

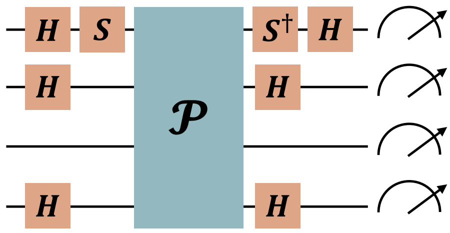


Figure 4.2.: *Pauli error estimation circuit* - One circuit instants for the estimation of the Pauli error rates $p(c)$ of a Pauli channel $\mathcal{P}$. The first set of applied single-qubit gates prepares the state $\chi_0^{a=\{2,1,0,1\}}$. After the application of the Pauli channel, a Pauli measurement in the basis a is conducted on the circuit.

The application of the Pauli channel does not change the basis of the input state χ_0^a. However, depending on whether the individual elements of P_a and P_c commute, the excitation is flipped. Therefore, the readout of a measurement in the Pauli basis a, after the black box applied the Pauli operator P_c on the input state χ_0^a, corresponds to the bitstring $r := a \star c$. Illustrated for one qubit in the following table:

	r	**c**			
	$a \star c$	0	1	2	3
	1	0	0	1	1
a	2	0	1	0	1
	3	0	1	1	0

(4.7)

Although for a given a, different P_c lead to the same r, r still reveals some information on which Pauli operator P_c was applied inside the black box. The next step is to analyze how we can estimate the Pauli error rates from a set of r and a. The experimental circuit corresponding to the described black box measurement is shown in Fig. 4.2.

We first discuss how to reconstruct the Pauli error rates for a single-qubit Pauli channel and then generalize this case to multi-qubit Pauli channels. Let us assume we want to find the Pauli error rate corresponding to the single-qubit Pauli operator P_b. For each experiment run, we can calculate $a \star b$, which corresponds to the result expected if the black box applied Pauli operator P_b. The actual result $r = a \star c$ is equivalent to $a \star b$, for a uniformly random drawn from $\{1, 2, 3\}$, with probability 1 if P_c equals P_b and with probability $\frac{1}{3}$ if P_c differs from P_b. Table 4.1 explains the respective probability distribution.

We introduce the binary variable $w := r +_2 a \star b$, with $+_2$ denoting the binary addition. Note that for one circuit measurement, we can calculate the quantity w for all guesses $b \in \{0, 1, 2, 3\}$. For an experiment with a uniformly random drawn a, the obtained value

	w	**c**			
	$b = 1$	0	1	2	3
	1	0	0	1	1
a	2	1	0	1	0
	3	1	0	0	1

Table 4.1.: The tabular shows the results for $w = a \star c +_2 a \star b$ with $b = 1$. For some $c \neq 1$ the probability for $w = 0$ assuming a uniformly random sampled a is $\frac{1}{3}$. If $c = 1$ the probability for $w = 0$ is 1.

of w is distributed for some guess b according to

$$w = r +_2 a \star b = \begin{cases} 0, & p(b) + \frac{1}{3}(1 - p(b)) \\ 1, & \frac{2}{3}(1 - p(b)). \end{cases} \tag{4.8}$$

With the observation that the expectation value for w corresponds to $\frac{2}{3}(1 - p(b))$, we can reconstruct the Pauli error rate $p(b)$ with the relation

$$p(b) = 1 - \frac{3}{2}\,\mathbb{E}[w]. \tag{4.9}$$

Another approach to restore $p(b)$ from Eq. (4.8) is introducing the variable $G := \left(-\frac{1}{2}\right)^w$. According to

$$\mathbb{E}[G|c] = \mathbb{E}\left[\left(-\frac{1}{2}\right)^w \middle| c\right] = \begin{cases} \left(-\frac{1}{2}\right)^0 = 1 & \text{if } b = c \\ \frac{1}{3}\left(-\frac{1}{2}\right)^0 + \frac{2}{3}\left(-\frac{1}{2}\right)^1 = 0 & \text{if } b \neq c, \end{cases} \tag{4.10}$$

the expectation value of G for a fixed P_c evaluates to 1 if $c = b$ and to 0 if $c \neq b$. This leads to the simple result

$$\mathbb{E}[G] = \sum_c p(c)\,\mathbb{E}[G|c] = p(b), \tag{4.11}$$

which can be evaluated for various b.

For n-qubit Pauli channels, $p(c)$ is a probability distribution over $c \in \{0,1,2,3\}^n$. a is an element in $\{1,2,3\}^n$ and b is element in $\{0,1,2,3\}^n$. Consequently, $w \in \{0,1\}^n$ is a bitstring of length n. $G = \prod_{t=0}^n \left(-\frac{1}{2}\right)^{w_t}$ is generalized as the product of all single-qubit G, where t denotes qubit index. The coordinates w_t are independent random variables. Therefore, $\mathbb{E}[G] = \prod_{t=0}^n \mathbb{E}\left[\left(-\frac{1}{2}\right)^{w_t}\right]$ is valid, and combined with Eq. (4.10) we reach

$$\mathbb{E}[G|c] = \prod_{t=1}^n \mathbb{E}\left[\left(-\frac{1}{2}\right)^{w_t} \middle| c\right] = \begin{cases} 1 & \text{if } b = c \\ 0 & \text{if } b \neq c. \end{cases} \tag{4.12}$$

This finally implies $\mathbb{E}[G] = p(b)$ and allows us to estimate all Pauli error rates. An example of the steps in the protocol for two qubits is illustrated in Fig. 4.3.

We can apply the above-stated protocol for all choices of b and estimate potentially any Pauli error rate from a set of measurements. However, as an n-qubit process contains 4^n potential Pauli errors, the question arises on how we efficiently find all Pauli error rates above some threshold ε.

The presented protocol not only allows to estimate $p(b)$ for any $b \in \{0,1,2,3\}^n$ but also for any shorter string, called marginal, $\beta \in \{0,1,2,3\}^l$. This is achieved by ignoring all data in positions $l+1, ..., n$. For example, for a 2-qubit Pauli channel, $p(\beta = \{1\})$ corresponds to $p(\{1,0\}) + p(\{1,1\}) + p(\{1,2\}) + p(\{1,3\})$. Therefore, if the estimation $\hat{p}(\beta = \{1\})$ is smaller than ε, all two-qubit Pauli error rates that emerge from this marginal are also lower than ε.

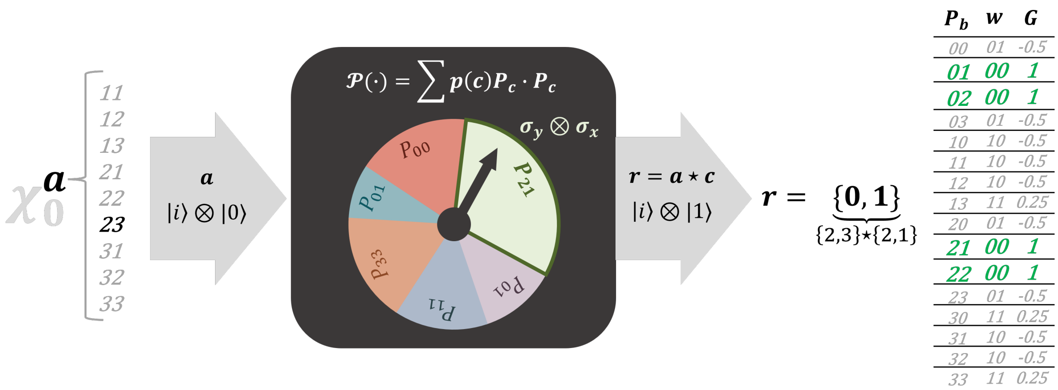


Figure 4.3.: *Pauli black box* - Illustration of the main concept of the Pauli channel estimation protocol by Flammia *et al.* [39]. A preparation basis a is chosen uniformly over $\{1,2,3\}^n$. In the example of the figure, the state $|i\rangle \otimes |0\rangle$ is prepared. The Pauli channel is interpreted as a black box that applies a Pauli operator P_c according to a hidden probability distribution $p(c)$ over $\{0,1,2,3\}^n$. The resulting state is $\chi^a_{a\star c}$. In the example, where the Pauli operator $\sigma_y \otimes \sigma_x$ is applied, the state $|i\rangle \otimes |1\rangle$ is reached. The quantity $r = a \star c$ can be accessed by a readout of the outputted state in the Pauli basis a. From the quantity $w = r +_2 a \star b$, $G = \prod_{t=0}^{n} \left(-\frac{1}{2}\right)^{w_t}$ can be calculated for all $b \in \{0,1,2,3\}^n$. The expectation value of G for a certain b equals $p(b)$. Therefore, by running the protocol repeatedly, it is possible to estimate all Pauli error rates of the Pauli channel $\mathcal{P}$.

With this concept, a cascade that starts with estimating all marginal Pauli error rates of the first qubit can be built. It proceeds with estimating all two-qubit marginals that arise from one-qubit marginals above the threshold. Then for all estimated two-qubit marginals above the threshold, the 3-qubit marginal error rates are calculated. This procedure is repeated until the Pauli error rates of the entire system are estimated.

4.1.3. Complete Protocol

The protocol to estimate the Pauli error rates introduced in the last subsection can not be applied directly to the parallel Pauli noise model of Eq. (3.2) because there is no experimental access to the Pauli channel itself but only to the probabilistic mixture of ideal and Pauli channel. Aiming to isolate the Pauli channel, we rearrange the terms in Eq. (3.2) to

$$\sum_c p(c) P_c \rho P_c = \frac{1}{\eta}\left(\mathcal{E}_{N_p}(\rho) - (1-\eta)\, U_I \rho U_I^\dagger\right). \tag{4.13}$$

The idea is to estimate the right-hand side for different states χ_0^a such that then the presented protocol of Flammia *et al.* [39] can be applied.

To perform the subtraction of the terms on the right-hand side of Eq. (4.13), a density matrix representation $\mathcal{E}_{N_p}(\chi_0^a)$ and $U_I\chi_0^aU_I^\dagger$ is required. The task of estimating $\mathcal{E}_{N_p}(\chi_0^a)$ corresponds to a quantum state tomography (QST) problem. The evolution $U_I\chi_0^aU_I^\dagger$ is calculated classically. The computational complexity for both parts, QST and the classical simulation of general unitary evolutions, scale exponentially with the number of involved qubits. Therefore, as we aim for a scalable protocol, we have to introduce further assumptions on U_I and the Pauli error rates.

As discussed in the introduction to the Pauli noise models in section 3.1, we restrict to two-qubit Pauli errors. First, this limits the number of different Pauli errors from $\mathcal{O}(4^n)$ to $\mathcal{O}(n^2)$. Secondly, this reduces the task from estimating Pauli errors of length n to estimating marginal Pauli error rates involving only two qubits. As described in the cascading estimation scheme in the last section, the data of all other than the involved qubits is not relevant for estimating marginal Pauli error rates. Mathematically speaking, this implies that for the estimation of marginal Pauli error rates, both sides of Eq. (4.4) can be traced over all except the involved two qubits (for qubits k and l denoted by $\mathrm{Tr}_{\neg(k,l)}[\cdot]$). This leads to $\mathcal{O}(n^2)$ equations of the form:

$$\sum_{c\in\{0,1,2,3\}^2} p(c)P_c\chi_0^{\{a_k,a_l\}}P_c = \frac{1}{\eta}\,\mathrm{Tr}_{\neg(k,l)}\left[\mathcal{E}_{N_p}(\chi_0^a)\right] - \frac{(1-\eta)}{\eta}\,\mathrm{Tr}_{\neg(k,l)}\left[U_I\chi_0^aU_I^\dagger\right]. \quad (4.14)$$

By considering only the marginal Pauli channels, we have changed the QST problem from estimating one density matrix of size $2^n \times 2^n$ to the estimation of $\mathcal{O}(n^2)$ reduced density matrices of size 4×4. For this estimation, we use the classical shadow tomography protocol introduced in section 2.1.3. The protocol efficiently estimates $\mathrm{Tr}_{\neg(k,l)}\left[\mathcal{E}_{N_p}(\chi_0^a)\right]$ for all reduced subsystems.

The complexity of the calculation $\mathrm{Tr}_{\neg(k,l)}\left[U_I\chi_0^aU_I^\dagger\right]$ scales exponentially with increasing system size. Therefore, we either assume that U_I can be written as a tensor product of multiple unitaries of tractable size or that U_I corresponds to a Clifford unitary [72]. In the case that we are characterizing gate-layers, both assumptions are usually valid.

Finally, all pieces are ready to characterize the complete model by the following procedure (illustrated in Fig. 4.4):

1. Estimate η according to Eq. (4.4).
2. Apply classical shadow tomography to estimate $\mathrm{Tr}_{\neg(k,l)}\left[\mathcal{E}_{N_p}(\chi_0^a)\right]$ for all two-qubit subsystems, consisting of qubits k and l, and different prepared states χ_0^a.
3. For the same prepared states, calculate for each subsystem $\mathrm{Tr}_{\neg(k,l)}\left[U_I\chi_0^aU_I^\dagger\right]$.
4. Estimate for each subsystem the Pauli channel evolution of all prepared states by applying Eq. (4.14).
5. Apply the Pauli error estimation protocol of section 4.1.2 for each subsystem.

While in the initially presented Pauli error rate estimation protocol, we apply Pauli measurements on the state $\mathcal{P}(\rho)$, we now can calculate the probabilities for all possible outcomes of this measurement in post-processing as we have already estimated the reduced density matrices of the evolved state in Eq. (4.14).

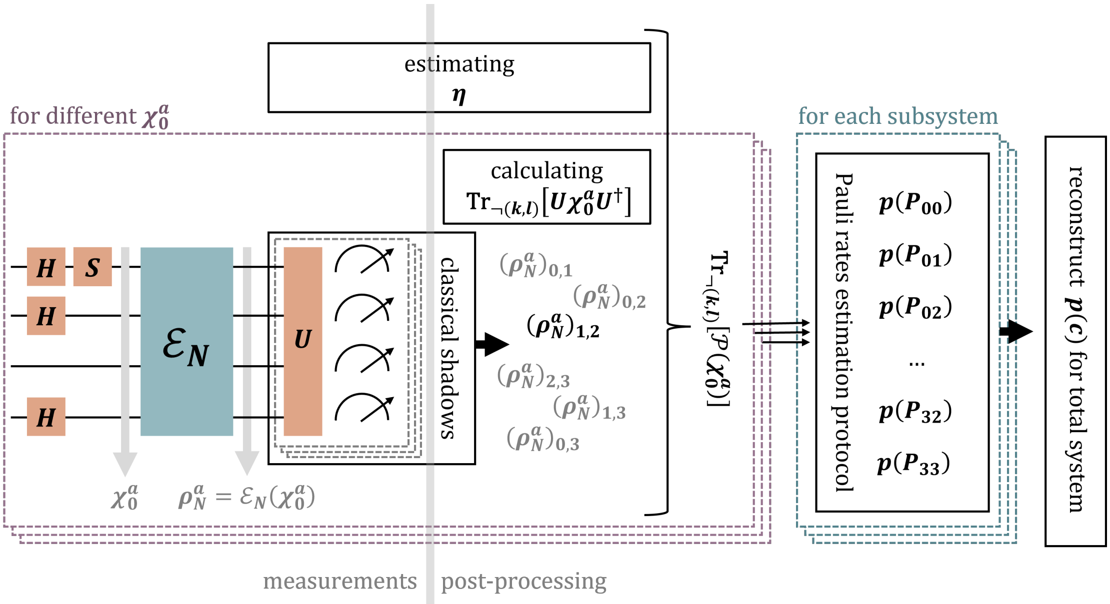


Figure 4.4.: *Parallel Pauli noise model characterization protocol* - Overview of the estimation protocol for the parallel Pauli noise model. The protocol starts by preparing a state χ_0^a and evolving it with the noisy process. The reduced density matrices of the output state are then estimated with classical shadow tomography. This, combined with the estimate for η and the calculated ideal evolution, allows estimating how the sole Pauli noise channel would evolve the initial state. After repeating this procedure for many initial states, the Pauli error rates can be estimated for each subsystem. Finally, under some assumptions regarding the correlation length of the Pauli errors, the full system Pauli errors are reconstructed.

Applying the Pauli error estimation protocol on all two-qubit subsystems returns 16 marginal Pauli error rates per subsystem. However, the marginal Pauli error rates are not independent of each other. For example, in the subsystem of qubits 0 and 1, a two-qubit Pauli error on qubits 1 and 3 will be observed as a single-qubit Pauli error on qubit 1. We assumed that only one- and two-qubit Pauli errors are present in the system. Therefore, whenever there is a two-qubit Pauli error in a two-qubit subsystem, we can erase its effect on single Pauli error rates in other subsystems. If we do that for all detected two-Pauli error rates, we restore the correct single-qubit error rates. Finally, we can find the Pauli error rate connected to the identity by subtracting the sum of all single- and two-qubit error rates from one, enforcing that all probabilities sum to one.

4.1.4. Extension to decay

A decay channel is not a Pauli channel. Consequently, the parallel Pauli noise model fails to model noise that originates from decay. In order to increase the expressibility of the model, we expand it to the decay extended model of Eq. (3.3). The multi-qubit decay channel corresponds to the application of the single-qubit decay channel to each qubit, where the decay parameter γ can vary between qubits. By introducing the Kraus operators

$$K_0 = \sqrt{\gamma}\sigma_- = \frac{\sqrt{\gamma}}{2}\left(\sigma_x + i\sigma_y\right) \text{ and } K_1 = \left(\frac{1}{2} + \frac{\sqrt{1-\gamma}}{2}\right)\mathbb{I} + \left(\frac{1}{2} - \frac{\sqrt{1-\gamma}}{2}\right)\sigma_z, \tag{4.15}$$

we express the single-qubit channel $\mathcal{F}_\gamma$ as $\mathcal{F}_\gamma(\rho) = K_0\rho K_0^\dagger + K_1\rho K_1^\dagger$.

The extended model of the noisy process, $\mathcal{E}_{N_p^\gamma}(\rho)$, consists of three independent sets of parameters, γ, η, and $p(c)$. We take the same assumptions for the Pauli error rates as in the protocol without decay, meaning we still restrict the model to two-qubit Pauli errors. γ consists of the individual decay rate of each qubit in the system, and η is a single parameter that determines the probabilistic mixture of the ideal process and the Pauli channel. In the following, we will introduce a procedure to characterize γ for all qubits and discuss for η and the Pauli error rates how the presented characterization protocol can be adjusted to the model with decay.

Decay Rate Estimation

In order to estimate the decay rates, we use the fact that, different from the Pauli channel and the ideal process, the decay channel is nonunital. This leads to $\mathcal{E}_{N_p^\gamma}(\mathbb{I}) = \mathcal{F}_\gamma(\mathbb{I}) \neq \mathbb{I}$. As established in

$$\begin{aligned}
\mathrm{Tr}\left[\sigma_z^{(k)}\mathcal{E}_{N_p^\gamma}\left(\frac{1}{2^n}\mathbb{I}\right)\right] &= \mathrm{Tr}\left[\sigma_z^{(k)}\mathcal{F}_\gamma\left(\frac{1}{2^n}\mathbb{I}\right)\right] \\
&= \frac{1}{2^n}\mathrm{Tr}\left[\left(K_0^{\dagger(k)}\sigma_z^{(k)}K_0^{(k)} + K_1^{\dagger(k)}\sigma_z^{(k)}K_1^{(k)}\right)\mathbb{I}\right] \\
&= \frac{2^{(n-1)}}{2^n}\mathrm{Tr}_k\left[\left(1-\gamma^{(k)}\right)\sigma_z + \gamma\mathbb{I}\right] \\
&= \gamma^{(k)},
\end{aligned} \tag{4.16}$$

the measurement of the expectation value of $\sigma_z^{(k)}$ of the evolved completely mixed state reveals the decay rate $\gamma^{(k)}$ of qubit k. $\sigma_z^{(k)}$ is the Pauli operator that acts as σ_z on qubit k and as the identity on all other qubits. However, instead of preparing the completely mixed state, which would require the use of usually undesired dissipation in the system, we exploit the linear property of channels. The completely mixed state corresponds to an even mixture of all states in the computational basis $\mathbb{I} = \sum_{a\in\{0,1\}^n}\chi_a^{\{3\}^n}$. Therefore, the protocol starts by preparing the state $\chi_a^{\{3\}^n}$ for a uniformly random sampled $a \in \{0,1\}^n$, which corresponds to applying a sparse layer of X-gates to the ground state. We then evolve the state by the noisy channel and measure it in the computational basis. From the measurement, we can reconstruct $\langle\sigma_z^{(k)}\rangle_{\chi_c^{\{3\}^n}}$ for all qubits. The average of $\langle\sigma_z^{(k)}\rangle$ over many preparations is finally an estimate for $\gamma^{(k)}$.

Estimation of η

After determining $\gamma^{(k)}$, we can proceed with estimating η. The protocol is similar to the presented protocol in the absence of decay in section 4.1.1. The idea is again to prepare a state χ_0^a and estimate the Pauli expectation value for Pauli operator P_b with $a \star b \neq \{0\}^n$ and $\mathrm{Tr}[P_b U_I \chi_0^a U_I^\dagger] \neq 0$. However, in the protocol with decay, we choose P_b to be a Pauli operator $\sigma_z^{(k)}$ and then choose a accordingly. The constraint $a \star b \neq \{0\}^n$ requires qubit k to be prepared either in the state $|+\rangle$ or $|i\rangle$. The expectation value of $\sigma_z^{(k)}$ of the state χ_0^a after the noisy channel action in terms of the decay rate, η, and the ideal expectation value $\xi_{a,b}$ is expressed as

$$\begin{aligned}
&\mathrm{Tr}\left[\sigma_z^{(k)}\mathcal{E}_{N_p^\gamma}\left(\chi_0^b\right)\right] \\
&= \mathrm{Tr}\left[\left(K_0^{\dagger(k)}\sigma_z^{(k)}K_0^{(k)} + K_1^{\dagger(k)}\sigma_z^{(k)}K_1^{(k)}\right)\mathcal{E}_{N_p}\left(\chi_0^b\right)\right] \\
&= (1-\gamma^{(k)})\mathrm{Tr}\left[\sigma_z^{(k)}\mathcal{E}_{N_p}\left(\chi_0^b\right)\right] + \gamma^{(k)}\mathrm{Tr}\left[\mathcal{E}_{N_p}\left(\chi_0^b\right)\right] \\
&= (1-\gamma^{(k)})(1-\eta)\underbrace{\mathrm{Tr}\left[\sigma_z^{(k)}U_I\chi_0^b U_I^\dagger\right]}_{\xi_{a,b}} + \gamma^{(k)}.
\end{aligned} \tag{4.17}$$

We can rearrange this equation to an expression for η

$$\eta = 1 - \frac{\mathrm{Tr}\left[\sigma_z^{(k)}\mathcal{E}_{N_p^\gamma}\left(\chi_0^b\right)\right] - \gamma^{(k)}}{(1-\gamma^{(k)})\mathrm{Tr}\left[\sigma_z^{(k)}U_I\chi_0^b U_I^\dagger\right]}. \tag{4.18}$$

The terms on the right-hand side of this equation are known $\gamma^{(k)}$, can be estimated by an experiment similar to the protocol shown in Fig. 4.1 $\mathrm{Tr}\left[\sigma_z^{(k)}\mathcal{E}_{N_p^\gamma}\left(\chi_0^b\right)\right]$, or are calculable in the knowledge of the ideal process $\mathrm{Tr}\left[\sigma_z^{(k)}U_I\chi_0^b U_I^\dagger\right]$.

Pauli Error Rates Estimation

With the estimation of γ and η, the missing piece is the estimation of the Pauli channels in the presence of decay. In the context of the whole protocol, this can be seen as the problem of estimating the Pauli error rates on the left-hand side of the expression

$$\mathcal{F}_\gamma\left(\sum_c p(c)P_c\rho P_c\right) = \frac{1}{\eta}\left(\mathcal{E}_{N_p^\gamma}(\rho) - (1-\eta)\,\mathcal{F}_\gamma\left(U_I\rho U_I^\dagger\right)\right). \tag{4.19}$$

We have access to the processes on the right-hand side of the equation and know the decay rates γ. The same methods achieve this as in the protocol without decay, namely the reduction to 2-qubit subsystems, the application of classical shadow tomography, and the calculation of the ideal unitary evolution with known decay. We now present our extension of the protocol of section 4.1.2 to the estimation of Pauli error rates in the presence of decay.

The main feature of Pauli channels that facilitates the protocol of Flammia *et al.* [39] is the symmetry of the action of Pauli channels on the eigenstates of each Pauli operator, as shown in table (4.7). This symmetry leads then to Eq. (4.8), from where the elegant recovery scheme of Eq. (4.12) is derived. Our attempt to solve the Pauli error estimation problem with decay is to recreate this symmetry and apply similar recovery techniques as in the protocol without decay.

The recreation of table (4.7) in the presence of decay leads to

$\mathbf{p}(\tilde{r})$				**c** 0	1	2	3
a	1	$\|+\rangle$	$r=0$	$0.5+\tilde{\gamma}$	$0.5+\tilde{\gamma}$	$0.5-\tilde{\gamma}$	$0.5-\tilde{\gamma}$
	2	$\|i\rangle$	$r=0$	$0.5+\tilde{\gamma}$	$0.5-\tilde{\gamma}$	$0.5+\tilde{\gamma}$	$0.5-\tilde{\gamma}$
	3	$\|0\rangle$	$r=0$	1	γ	γ	1
		$\|1\rangle$	$r=1$	$1-\gamma$	0	0	$1-\gamma$

(4.20)

The first difference to the case without decay is that the resulting state is no longer pure. Therefore, the readout $\tilde{r}$ of a Pauli measurement in the basis of the preparation can no longer be expressed as $a \star c$. Instead the outcome is probabilistic. The first three rows in table (4.20) show the probability of measuring 0, where $\tilde{\gamma}$ corresponds to $\sqrt{1-\gamma}/2$, which equals $1-\gamma/2$ in the first order of gamma. The symmetry in the preparation basis is broken, as the probabilities to measure 0 is different for the state $|0\rangle$ than for the states $|+\rangle$ and $|i\rangle$. The fourth line of the table shows the probabilities of having a readout of 1 if we prepare the state $|1\rangle$. While in the absence of decay, this would correspond to the case of preparing $|0\rangle$ and measuring 0, decay breaks this symmetry in excitation. Note that also considering the states $|-\rangle$ and $|-i\rangle$ would not add new aspects to the table, as those cases correspond to the cases of $|+\rangle$, respectively $|i\rangle$.

We can symmetrize table 4.20 and restore the symmetries of the case with no decay in two steps. First, whenever we draw to prepare a qubit in the z-basis, we prepare with an equal probability state $|0\rangle$ or $|1\rangle$. If we prepared state $|1\rangle$, the readout of the respective qubit is inverted in post-processing. Secondly, whenever we draw to prepare a qubit in the state $|+\rangle$ or $|i\rangle$, we invert in post-processing the obtained readout with a probability of $\frac{1}{2}+\frac{1-\gamma}{2\sqrt{1-\gamma}}$. With these two mitigation steps, we obtain the table:

$\mathbf{p}(\tilde{r}=0)$		**c** 0	1	2	3
a	1	$1-\gamma/2$	$1-\gamma/2$	$\gamma/2$	$\gamma/2$
	2	$1-\gamma/2$	$\gamma/2$	$1-\gamma/2$	$\gamma/2$
	3	$1-\gamma/2$	$\gamma/2$	$\gamma/2$	$1-\gamma/2$

(4.21)

This table resembles the structure of table 4.7. In order to create a similar recovery scheme as in Flammia *et al.* [39], we introduce the value $\tilde{w}=\tilde{r}+_2 a\star b$ and use the ansatz $\tilde{G}=f\cdot(-m)^{\tilde{w}}$. The values for f and m are chosen such that $\mathbb{E}\left[\tilde{G}\right]=p(b)$, which corresponds to $f=\frac{1-\gamma}{4(1-\gamma)}$ and $m=\frac{1+\gamma}{4-\gamma}$. We generalize this to multi-qubit Pauli error rates analogous to the case for no decay and arrive at equation

$$\mathbb{E}[\tilde{G}|c] = \prod_{t=1}^{n} \mathbb{E}\left[\frac{1-\gamma^{(t)}}{4(1-\gamma^{(t)})}\left(-\frac{1+\gamma^{(t)}}{4-\gamma^{(t)}}\right)^{\tilde{w}_t}\middle| c\right] = \begin{cases} 1 & \text{if } b = c \\ 0 & \text{if } b \neq c. \end{cases} \tag{4.22}$$

This subsection presented a protocol to estimate decay rates, η, and the Pauli errors in the framework of the parallel Pauli noise model with decay. Combining these concepts with the classical shadow application and the recovery scheme in the last section enables complete characterization of the introduced model. The protocol is scalable to many qubits under the same assumptions as the model without decay and for individual decay rates.

4.1.5. Simulation Results

We simulate the presented protocol using Qiskit [73]. Fig. 4.5 shows the characterized values for η, the decay rates, and the Pauli error rates for a four qubit system. For the simulation, we drew a random 4-qubit unitary and noise according to the parallel Pauli noise model. In total, 8 Pauli error rates were included in the model, 5 being two-qubit errors and two being single-qubit errors. For the simulation of the decay rates and the following estimation of η we run $400'000$ measurements each. For the Pauli error estimation, $1'440'000$ circuits are measured. The presented error bars are estimated by repeating the experiment multiple times.

We calculated the scaling of the Pauli error estimation error with the number of measurements by combining the reported scaling of classical shadow tomography [49] and the Pauli error estimation protocol [39], with the scaling factor due to the probabilistic mixture of Pauli and ideal channel. In the picture of the probabilistic mixture between noise and ideal process, the intuitive argument is that for smaller η, it is less likely that the state is evolved by the noise term and, therefore, less probable that a measurement reveals something about the noise. Finally, we reach a scaling of the error ε in the Pauli error rates with the number of measurements of $\mathcal{O}\left(\frac{\log n}{\epsilon^2\eta^2}\right)$. We confirmed this scaling for η and ε in a simulation (Fig. A.3).

The simulation results illustrate the capability of the protocol to estimate η, γ, and the 8 introduced Pauli error rates. The values of η and γ were estimated with high accuracy. We set η relatively high (compared to error rates of real processes), characterized a process perfectly described by the model, and executed an amount of measurement that would take several minutes to run on the provided hardware platform of IBM [74]. However, the estimation uncertainty is still in the order of 25% of the Pauli error rates. For lower η, an accurate estimation of the Pauli error rates will become challenging. Nevertheless, it is notable that all introduced Pauli errors were captured in the simulation. Demonstrating the protocol's potential to characterize leading errors on which one then can further elaborate. The simulation proves the protocol's applicability to processes that the parallel Pauli noise model can model.

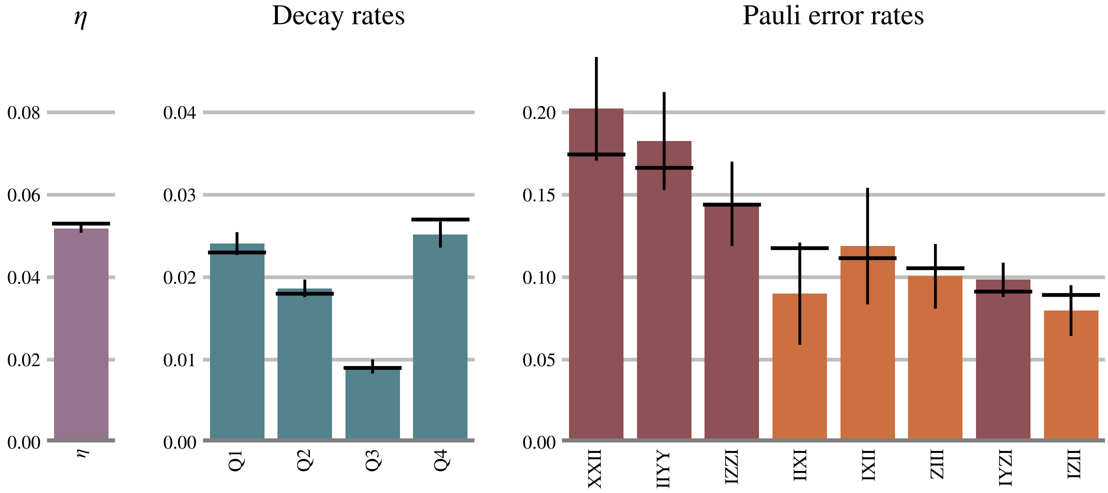


Figure 4.5.: *Parallel Pauli noise model characterization results* - Simulation result of the characterization of a 4-qubit process that is modeled by the parallel Pauli noise model. The ideal model parameters are indicated by the black line, while the characterized values are represented by the bars. The error bars are created by repeating the same characterization several times.

4.1.6. Analysis

We presented a scalable protocol to estimate all parameters in the parallel Pauli noise model. To achieve this, we combined different existing concepts, such as classical shadows [49] and an existing Pauli error estimation protocol [39], with new approaches. Furthermore, we managed to extend the characterization to a decay-including model. The protocol uses only simple state preparation and measurement processes. Its resource requirements are met by current existing hardware platforms [74].

However, the protocol does not resist state preparation and measurement errors (SPAM). The usual concept to deal with SPAM errors, for example, in randomized benchmarking, is to apply the noisy channel repeatedly. This procedure amplifies the noise introduced by the examined process. Within the assumption of a parallel Pauli noise model, such a repeated scheme will lead to a broad mixture of different structured channels. For example, already for two repetitions, the model creates the terms $U_I U_I \rho U_I^\dagger U_I^\dagger$ with probability $(1-\eta)^2$, $\mathcal{P}(U_I \rho U_I^\dagger)$ and $U_I \mathcal{P}(\rho) U_I^\dagger$ with probability $\eta(1-\eta)$, and $\mathcal{P}\left(\mathcal{P}(\rho)\right)$ with probability η^2. Although one can restrict the number of terms with a first-order approximation in η, the presented protocol fails to deal with these mixed terms, and we did not manage to create a protocol that facilitates the estimation of the parameters in the case of repeated application of the noisy channel.

Beyond the technical intricacies of the model, it is unclear whether the presented model adequately represents physical noise processes insightfully, as seen in section 3.1.1. Furthermore, coherent noise can not be modeled as a probabilistic mixture of an ideal and a noise channel because the mixture leads to an incoherent process. Therefore, an extension of the model to coherent noise in the spirit of the model is impossible.

4.2. Sequential Pauli Noise

In contrast to the parallel Pauli noise model, the sequential Pauli noise model is well-established in the literature [5, 32, 75]. In the spirit of randomized benchmarking, the protocols to characterize the Pauli errors in the sequential model usually contain multiple repetitions of the noise-inducing element. This repetition amplifies the noise one wants to characterize while the amount of SPAM stays unchanged as it only appears once per circuit. By looking at some quantity for different amounts of amplification, one can isolate the noise introduced by the probed element from the SPAM and reach a SPAM-resistant noise estimation protocol. This section will further discuss this concept with the example of characterizing a Pauli channel.

An open problem is estimating Pauli error rates for general unitary processes U_I in the sequential model given in Eq. (3.4). The approach in Ref. [5] assumes that the introduced noise is gate-independent. They then estimate the average Pauli noise channel over the set of Clifford gates. Gate-specific noise is characterized in Refs. [75] and [32]. However, those protocols are limited to the characterization of Clifford gates as the protocols use the property $U_c P_a = P_b U_c$, where U_c is a Clifford unitary and P_a and P_b are Pauli operators. The protocols presented in Refs. [5], [75] and [32] have in common that the circuits used for characterization correspond in the noiseless case to the identity. Such a circuit can be constructed by adding a circuit element at the end that inverts the action of all previous gates [5], analogous to randomized benchmarking. Refs. [32] and [75] focus on the characterization of CX-gate-based processes (CX, $SWAP$, and CZ). Their circuits contain an even number of those elements, creating an identity because each of those gates is its own inverse. This thesis does not further expand on the theme of characterizing Pauli noise in the presence of arbitrary unitaries. We will use the example of an identity gate consisting of two CX gates to illustrate the main concept for the SPAM-free estimation of a Pauli channel. We refer to the CX-gate pair as *noisy identity*.

4.2.1. Protocol

Starting-point for creating a SPAM-free Pauli error estimation of a noisy identity channel is the diagonal PTM of a Pauli channel. As mentioned previously, the concatenation of two channels corresponds to the product of PTMs. Therefore, the PTM of a repeated application of the same Pauli channel is still diagonal with exponentiated elements on the diagonal. As discussed in section 2.3, those diagonal elements, which relate to the scaling factors of the Bloch sphere, are called Pauli eigenvalues $f(b)$ and link to the Pauli error rates $p(c)$ by a Walsh-Hadamard transform.

The diagonal PTM elements of a channel $\mathcal{E}$ correspond to $\mathrm{Tr}[P_i \mathcal{E}(P_i)]$ (cf. section 2.2.2). However, this expression can not be measured explicitly, as a Pauli matrix is not a valid state, and therefore, it is impossible to experimentally apply a channel to it. Instead, by

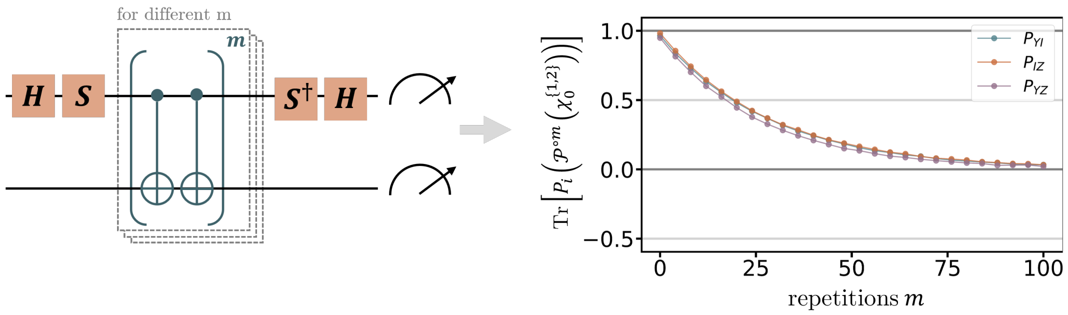


Figure 4.6.: *Repeated Pauli channel simulation* - Simulation data obtained with Qiksit [73]. The schematic of the executed circuits is shown on the left side of the figure. The expectation values of the measured Pauli operators, shown in the right half of the figure, decreases exponentially with the number of executed repetitions of the noisy identity. The noisy identity is implemented with two CX-gates. The noise model of Qiskit includes depolarization, measurement errors and decay.

using Eq. (2.16), we reach

$$\begin{aligned} \langle P_b \rangle_{\mathcal{P}(\rho)} &= \mathrm{Tr}\left[P_b \sum_c p(c) P_c \rho P_c\right] \\ &= \mathrm{Tr}\left[\sum_c p(c) P_c P_b P_c \rho\right] \\ &= \underbrace{\sum_c p(c)\,(-1)^{\langle c,b\rangle_{sp}}}_{f(b)}\ \mathrm{Tr}\left[P_b \rho\right]. \end{aligned} \tag{4.23}$$

If ρ is chosen to be $\chi_0^{b'}$, with $b' \star b = \{0\}^n$, $\chi_0^{b'}$ is an eigenstate of P_b. The equation simplifies further to $\langle P_b \rangle_{\mathcal{P}(\chi_0^{b'})} = f(b)$. The generalization of Eq. (4.24) for m repetitions of a Pauli channel, denoted as $\mathcal{P}^{\circ m}(\cdot)$ and state $\chi_0^{b'}$ is

$$\begin{aligned} \langle P_b \rangle_{\mathcal{P}^{\circ m}(\chi_0^{b'})} &= \mathrm{Tr}\left[P_b \sum_c p(c) P_c \left(\mathcal{P}^{\circ(m-1)}(\chi_0^{b'})\right) P_c\right] \\ &= f(b)\mathrm{Tr}\left[P_b \mathcal{P}^{\circ(m-1)}(\chi_0^{b'})\right] \\ &= f(b)\ \langle P_b \rangle_{\mathcal{P}^{\circ(m-1)}(\chi_0^{b'})} \\ &= \ldots = f(b)^m. \end{aligned} \tag{4.24}$$

We can illustrate the result of Eq. (4.24) by simulating the corresponding experiment. Fig. 4.6 shows the measurements of an experiment with two qubits that consists of preparing the state $\chi_0^{\{1,2\}} = |i0\rangle\langle i0|$, applying m pairs of the CX-gate and measuring

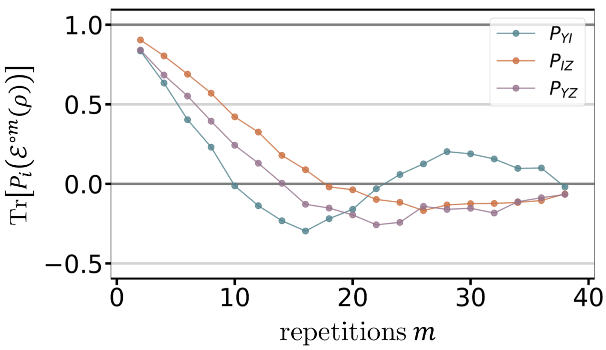


Figure 4.7.: *Repeated noisy identity experiment* - Experimental data from `ibm_perth`. We executed the same circuits as in Fig. 4.6. Different to the simulation, we observe oscillations of the measured expectation values with an increasing number of noisy identity repetitions.

the evolved state in the Pauli basis $\{1, 2\}$. From the readout, we can then restore the expectation value of the Pauli operators $P_{\{0,2\}}$, $P_{\{1,0\}}$, and $P_{\{1,2\}}$, which correspond to the Pauli eigenvalues of the m-times repeated channel. The simulation is executed on Qiskit [73]. It includes measurement errors, and gate errors are modeled as a combination of depolarizing errors and decay. As the depolarizing channel is a Pauli channel, we observe the expected exponential decay with the number of noisy identities. The fact that the curves do not exactly start at 1 for 0 repetitions relates to SPAM errors. The expectation values converge to a value slightly offset from 0 due to present decay. From this data, we can find the corresponding Pauli eigenvalues by fitting an exponential curve $g(x) = a \cdot b^x + c$, where the base b of the exponential term equals the eigenvalue. The presence of SPAM only influences the scaling factor a and not the estimation of the eigenvalue b [5]. Applying a Walsh-Hadamard transform on the estimated values reveals the Pauli error rates. To estimate all 16 Pauli eigenvalues of a two-qubit system, we must prepare and measure states in 9 different bases.

4.2.2. Experimental Results

Repeating the experiment in Fig. 4.6 on a superconducting quantum processor verifies whether the sequential Pauli noise model explains the noise induced by the noisy identity. We run this circuit with 1000 shots on `ibm_perth`, one of the IBM Quantum Falcon Processors[i] [74]. The resulting curves are displayed in Fig. 4.7.

While the expected exponential decay of the expectation values is visible, there is also an oscillatory behavior. The oscillations have different frequencies for different expectation

[i]We acknowledge the use of IBM Quantum services for this work. The views expressed are those of the authors, and do not reflect the official policy or position of IBM or the IBM Quantum team.

values. A purely diagonal PTM of the noise channel, as assumed in the sequential Pauli noise model, can not explain this behavior because, under repeated application of a diagonal PTM, the magnitudes of the expectation values are strictly decreasing and will never change sign. Therefore, the sequential Pauli noise model clearly fails to capture the main noise contributions for this process. From this finding, a usual way forward is to enforce the Pauli noise assumption by altering the noise with the application of twirling [5, 32, 75] (cf. section 2.2.3). This approach allows for the characterization of the modified channel according to the presented Pauli estimation protocol. The characterization information can then be used to mitigate the altered noise to restore the ideal behavior. However, the information that allows us to relate the characterized noise to physical sources is lost by twirling the noise. Consequently, the twirling approach is not heplfull if one wants to repeal the noise by hardware changes or device calibration. Furthermore, twirling creates some circuit overhead, and precise knowledge of the noise might facilitate more efficient gate-level noise corrections.

Instead of twirling, we propose a physically motivated model expansion that potentially allows explaining the observed oscillations. The experimental behavior reminds us of Rabi cycles [7] - a well-known oscillating process at the core of quantum computing. As Rabi cycles originate from Hamiltonian evolution, modeling the observed noise oscillations as a unitary evolution seems reasonable. This approach leads to the coherent noise model introduced in section 3.2.

4.3. Coherent Noise

The extension of the sequential Pauli model to the coherent noise model of Eq. (3.5) is motivated by the experimental results of the previous section. This section outlines a protocol to estimate the introduced noise modeling unitary U_θ in terms of single- and two-qubit rotation errors.

In Ref. [76], the systematic error between the intended rotation angle and the real rotation angle of a single-qubit x-rotation-gate is characterized ($R_x(\theta)$). They characterize this coherent error by applying the R_x-gate for various rotation angles to the ground state and subsequently measuring the probability of the qubit being in the excited state. They introduce an ansatz that includes the effect of the systematic error and find the systematic error by least-square fitting the measured data to the ansatz.

In this section, we follow a similar procedure. Our ansatz is the coherent noise model of Eq. (3.5). We measure circuits with various repetition lengths of the characterized process and apply a least-square fit to find the parameters of the model that resemble the real process best. As discussed in section 3.2, we consider the unitary that models the coherent error U_θ to be close to the identity. Furthermore, we only consider single-qubit coherent errors and correlated errors between neighboring qubits in the connectivity tree of the system. After the presentation of the characterization protocol, we demonstrate the method by characterizing the coherent errors introduced by a gate layer on a superconducting quantum processor.

4.3.1. Protocol

The protocol to estimate the rotation angles is first explained for an ideal process U_I corresponding to the identity. In the second step, we will generalize the result to other unitaries. The central concept of the characterization protocol is to fit the 15 rotation angles, corresponding to all single- and two-qubit rotations of each two-qubit subsystem, to the oscillations in measurements similar to those in Fig. 4.7.

According to the Pauli formalism introduced in chapter 2, the action of a noisy channel modeled by the coherent noise model $\mathcal{E}_{N_c}(\rho)$ is represented in terms of the vectorized density matrix of ρ and the PTM of the channel as $|\mathcal{E}_{N_c}(\rho)) = T_{\mathcal{E}_{N_c}}|\rho) = T_{\mathcal{P}}T_{U_\theta}T_{U_I}|\rho)$. As stated, we first discuss the case of T_{U_I} being the identity. Having access to an experiment, both $|\mathcal{E}_N(\rho))$ and $|\rho)$ can be estimated.

We reformulate the estimation problem of T_{U_θ} as a minimization problem for $\hat{T}_{U_\theta}$ of the term $\||\mathcal{E}_{N_c}(\rho)) - T_{\mathcal{P}}\hat{T}_{U_\theta}|\rho)\|^2$.

Statement. *With the assumption $T_{\mathcal{P}}$ and T_{U_θ} being close to the identity. The minimization of $\||\mathcal{E}_{N_c}(\rho)) - T_{\mathcal{P}}\hat{T}_{U_\theta}|\rho)\|^2$ for $\hat{T}_{U_\theta}$, corresponds to the minimization of $\||\mathcal{E}_{N_C}(\rho)) - \hat{T}_{U_\theta}|\rho)\|^2$ for $\hat{T}_{U_\theta}$.*

Proof. A noisy identity process described by the coherent noise model, $|\mathcal{E}_{N_c}(\rho))$ equals $T_{\mathcal{P}}T_{U_\theta}|\rho)$. We show that $\theta' = \theta$ minimizes the expression $\|T_{\mathcal{P}}T_{U_\theta}|\rho) - T_{U_{\theta'}}|\rho)\|$ independent of the choice of ρ.

$$\begin{aligned}
&\min_{\theta'} \||\mathcal{E}_N(\rho)) - T_{U_{\theta'}}|\rho)\|^2 \\
=&\min_{\theta'} \|T_{\mathcal{P}}T_{U_\theta}|\rho) - T_{U_{\theta'}}|\rho)\|^2 \\
=&\min_{\theta'} \left(|\rho)^\top \left(T_{U_\theta}^\top T_{\mathcal{P}}^\top - T_{U_{\theta'}}^\top\right)\left(T_{\mathcal{P}}T_{U_\theta} - T_{U_{\theta'}}\right)|\rho)\right) \\
=&\min_{\theta'} \left(|\rho)^\top \left(T_{U_\theta}^\top {T_{\mathcal{P}}}^2 T_{U_\theta} - T_{U_{\theta'}}^\top T_{\mathcal{P}}T_{U_\theta} - T_{U_\theta}^\top T_{\mathcal{P}}T_{U_{\theta'}} + T_{U_{\theta'}}^\top T_{U_{\theta'}}\right)|\rho)\right)
\end{aligned} \tag{4.25}$$

Where we used that $T_{\mathcal{P}}$ is diagonal. Given that the noise channel is close to the identity, we take the approximation that the diagonal matrix $T_{\mathcal{P}}$ and the matrix T_{U_θ} commute. Furthermore, because of the rotation matrix structure, $T_{U_\theta}^\top T_{U_\theta} = \mathbb{I}$ is valid.

$$\begin{aligned}
=&\min_{\theta'} \left(|\rho)^\top \left({T_{\mathcal{P}}}^2 - T_{\mathcal{P}}\left(T_{U_\theta}^\top T_{U_{\theta'}} + {T_{U_{\theta'}}}^\top T_{U_\theta}\right) + \mathbb{I}\right)|\rho)\right) \\
=&\min_{\theta'} \left(|\rho)^\top \left(\mathbb{I} + T_{\mathcal{P}}^2 - T_{\mathcal{P}}\left(T_{U_\theta}^\top T_{U_{\theta'}} + \left(T_{U_\theta}^\top T_{U_{\theta'}}\right)^\top\right)\right)|\rho)\right)
\end{aligned} \tag{4.26}$$

The off-diagonal part of $T_{U_\theta}^\top T_{U_{\theta'}}$ is anti-symmetric. Consequently, $T_{U_\theta}^\top T_{U_{\theta'}} + \left(T_{U_\theta}^\top T_{U_{\theta'}}\right)^\top$ is diagonal and $\leq 2\,\mathbb{I}$. Thereby, all involved matrices are diagonal. From $1 + T_{\mathcal{P}}^2 \geq 2\,T_{\mathcal{P}}$ follows that the examined expression is minimal if $T_{U_\theta}^\top T_{U_{\theta'}} + \left(T_{U_\theta}^\top T_{U_{\theta'}}\right)^\top = 2\,\mathbb{I}$. This corresponds to $T_{U_\theta} = T_{U_{\theta'}}$, which finally implies $\theta = \theta'$. □

This statement implies that we can optimize for the coherent noise term without considering the present Pauli noise, thus creating a relatively simple way to estimate T_{U_θ}. First, estimate experimentally $|\mathcal{E}_N(\rho))$ and $|\rho)$ for each two-qubit subsystem. Secondly, create a parameterized Pauli transfer matrix $\hat{T}_{U_{\theta'}}$, that consists of the 15 parameters of the single- and two-qubit rotation errors. Lastly, for each two-qubit subsystem, conduct a least square error optimization for $\||\mathcal{E}_N(\rho)) - \hat{T}_{U_\theta}|\rho)\|^2$.

While the PTM of the unitary evolution $U_\theta = \exp\left(\sum -i\theta_k P_k\right)$ can be calculated explicitly, simplifying the PTM to a structure that allows generating $\hat{T}_{U_\theta}$ quickly for any θ speeds up the optimization. Assuming small coherent errors, we use the approximation $U_\theta \approx \prod \exp\left(-i\theta_k P_k\right)$. This corresponds to the approximation that the single- and two-qubit rotations commute. The PTM $T_{U_{\theta_k}}$ that corresponds to each of the terms $\exp\left(-i\theta_k P_k\right)$ is constructed according to

$$
\begin{aligned}
\left[T_{U_{\theta_k}}\right]_{a,b} &= \frac{1}{2^n}\mathrm{Tr}\left[P_a U_{\theta_k} P_b U_{\theta_k}^\dagger\right] \\
&= \frac{1}{2^n}\mathrm{Tr}\left[P_a\left(\cos\theta_k\mathbb{I} - i\sin\theta_k P_k\right)P_b\left(\cos\theta_k\mathbb{I} + i\sin\theta_k P_k\right)\right] \\
&= \begin{cases} 0, & \langle b,k\rangle_{\mathrm{sp}} = 0 \wedge b \neq a \\ 1, & \langle b,k\rangle_{\mathrm{sp}} = 0 \wedge b = a \\ \frac{i}{2^n}\ \sin 2\theta_k\ \mathrm{Tr}\left[P_a P_b P_k\right], & \langle b,k\rangle_{\mathrm{sp}} = 1 \wedge b \neq a \\ \cos 2\theta_k, & \langle b,k\rangle_{\mathrm{sp}} = 1 \wedge b = a. \end{cases}
\end{aligned}
\tag{4.27}
$$

Finally, using again the assumption that the rotations $T_{U_{\theta_k}}$ are close to the identity, $T_{U_\theta} \approx \prod_k T_{U_{\theta_k}}$ is approximated by $T_{U_\theta} \approx \mathbb{I} + \sum_k \left(T_{U_{\theta_k}} - \mathbb{I}\right)$.

As we consider two-qubit subsystems, $T_{U_{\theta_k}}$ is parameterized with 15 parameters corresponding to the 6 single-qubit rotations related to the Pauli operators $P_k \in \{P_{\{1,0\}}, P_{\{2,0\}}, P_{\{3,0\}}, P_{\{0,1\}}, P_{\{0,2\}}, P_{\{0,3\}}\}$ and the 9 two-qubit rotations $P_k \in \{P_{\{i,j\}} \mid i,j \in \{1,2,3\}\}$. The vectorized density matrices $|\mathcal{E}_N(\rho))$ and $|\rho)$ also have 15 non-trivial entries, corresponding to the expectation values of the above-stated single- and two-qubit Pauli operators. Therefore, Pauli measurements in all bases of a general state ρ and the state $\mathcal{E}_N(\rho)$ suffice to predict $\hat{T}_{U_{\theta_k}}$. However, in practice, the prediction accuracy increased by adding redundancy through the preparation of many different states. Although not yet analyzed, the improved results from adding redundancy could originate from a higher SPAM error resistivity. In the implemented version of the protocol used for simulations and real experiments, we prepare states in each of the 9 Pauli bases and apply multiple repetitions of the noisy identity. We repeat the channel application as this allows us to use some of the estimated states for two optimizations, which lowers the complexity. For example, by estimating $|\rho)$, $|\mathcal{E}_N(\rho))$, and $|\mathcal{E}_N(\mathcal{E}_N(\rho)))$ we can apply the optimization between the states ρ and $\mathcal{E}_N(\rho)$, and between $\mathcal{E}_N(\rho)$ and $\mathcal{E}_N(\mathcal{E}_N(\rho))$. An example of all the data collected to estimate the coherent noise and the corresponding circuit schematic is shown in Fig. B.3.

Furthermore, when looking at each qubit subsystem separately, one has to address the correlated two-qubit noise between a qubit in the regarded subsystem and a qubit outside the subsystem. We avoid the influence of this correlated noise by preparing the subsystem-surrounding-qubits in multiple different bases for each preparation basis of the subsystem. This preparation prevents any correlation between the preparation of the subsystem with the preparation of the neighboring qubits and thus randomizes the effect of the noise introduced from outside the regarded subsystem. Therefore, such a preparation scheme mitigates the bias of the outside noise in the optimization process. . If the ideal process does not correspond to the identity, an additional step after estimating $|\mathcal{E}_N(\rho)\rangle$ and $|\rho\rangle$ is necessary to characterize the coherent errors. In order to create the same optimization problem as for the identity case, we classically apply in post-processing the channel U_I on the estimated state $|\rho\rangle$. However, in order to do this for general U_I, it is no longer possible to treat each subsystem separately, and instead, full state tomography is necessary. To preserve scalability, we restrict the characterization to unitaries that are separable into multiple single- and two-qubit unitaries $U_I = U_1 \otimes U_2 \otimes ... \otimes U_l$. This requirement is fulfilled, for example, in the characterization of gate layers consisting of single- and two-qubit gates. For estimating the coherent two-qubit errors of qubits a and b, $|\rho\rangle$ must be estimated for the reduced subsystem consisting of the qubits a, b, and the qubits entangled by U_I with either of the two qubits. Therefore, maximally 4 qubit subsystems must be considered, which corresponds to the case that both qubits a and b are involved in a different two-qubit process. After estimating the vectorized partial density matrix $|\rho\rangle$, the partial unitary is applied. The vectorized density matrix of the enlarged subspace is then reduced to the subsystem that only involves qubits a and b. Now the same steps are applied as in the case where U_I corresponds to the identity. Note that, like the restriction to correlated errors of 2 qubits, the restriction to one- and two-qubit unitaries is arbitrary and different choices are possible.

4.3.2. Experimental Results

We characterize the coherent noise induced by a gate-layer executed on a real hardware platform using the presented protocol. The experiment is conducted on the IBM Quantum Falcon Processor `ibm_lagos` [74]. The processor specifications from the experiment time are displayed in table 4.2, and the connectivity tree is shown in Fig. 4.8. We aim to characterize the layer shown in the top right panel of Fig. 4.8, consisting of 2 CX-gates, one SX-gate, one X-gate, and one Hadamard-gate (H). Note that the Hadamard-gate is not a native gate of the used processor and is transpiled to multiple native single-qubit gates.

In the characterization procedure, we prepared the 7-qubits in a total of 216 different initial states. Each qubit is prepared in one of the states $|+\rangle, |-\rangle, |i\rangle, |-i\rangle, |0\rangle$, or $|1\rangle$. The large amount of different combinations, as well as the expansion of the usual basis $|+\rangle, |i\rangle, |0\rangle$ to the basis that includes preparations with an additional X-gate at the start, aims to lower the influence of noise from neighboring qubits in the two-qubit subsystem optimization. For each state, we apply 0 to 3 repetitions of the layer. The correction of the effect of the unitary in the estimation process necessitates estimating the state of 3-qubit subsystems. This leads to a total of 27 Pauli measurements per circuit. Overall we run $216 \cdot 4 \cdot 27 = 23'328$ different circuits with 128 shots each. The choice of the number

Qubits	Q0	Q1	Q2	Q3	Q4	Q5	Q6
readout-error [10^{-2}]	1.3	0.7	0.5	1.9	2.1	2.5	1.1
T1 [µs]	68	103	186	167	103	116	108
T2 [µs]	53	87	101	41	33	89	141
sx-error [10^{-4}]	2.8	2.3	1.9	1.9	2.7	1.9	3.0
x-error [10^{-4}]	2.8	2.3	1.9	1.9	2.7	1.9	3.0
CX-gate	0-1	5-6					
CX-error [10^{-3}]	6.8	7.2					
gate time [ns]	305	291					

Table 4.2.: Specification of the IBM Quantum Falcon Processor `ibm_lagos` at the time of the characterization. The numbering of the qubits refers to the numbers in Fig. 4.8.

of preparation bases and repetitions is based on hardware simulations of the platform executed with Qiskit [73]. It took approximately 30 minutes of system time to run all the circuits. The post-processing time of the optimization is irrelevant compared to the experiment time. We successfully tested the characterization protocol in simulations with processes described by the coherent noise model (not shown here).

The result of the coherent characterization is displayed in Fig. 4.8. The plot shows the absolute value of the angles of the characterized single- and two-qubit coherent rotation errors. The most significant two-qubit correlated coherent errors occur between qubits 1 and 2 and between qubits 5 and 6. This is expected, as those are the pairs of qubits on which the CX-gate is acting. The biggest two-qubit coherent error induced by both CX gates is associated with the Pauli $P_{\{2,3\}} = \sigma_y \otimes \sigma_z$. Overall predominantly single-qubit coherent errors are present. There seems to be no strong bias towards a specific single-qubit rotation error.

We conducted the exact characterization of Fig. 4.8 again after 19 days. The second time we run the characterization twice, waiting 6 hours between the two executions. The shifts in the characterization results between those three experiments are shown in Fig. A.1. First, we observe that the leading coherent errors in the system did not change over the 19 days. In the total period, the most significant single-qubit coherent errors have changed by maximally 10% (corresponding to about 0.03 radians). The two-qubit coherent rotation errors have changed by maximally 0.018 radians. As expected, the changes observed for the two experiments within 6 hours are significantly smaller than the differences over the complete time span. The results suggest that the platform is affected by systematic coherent errors which are not influenced by the daily re-calibration of the system.

The fact that the protocol manages in simulation to correctly characterize the noise of a known coherent noise model does not imply the correctness of the presented characterization, as the noise of the real process might not be expressible in the assumed coherent noise model. The next chapter aims to verify the assumed noise model by showing that the characterization results can be used to mitigate errors in a real system.

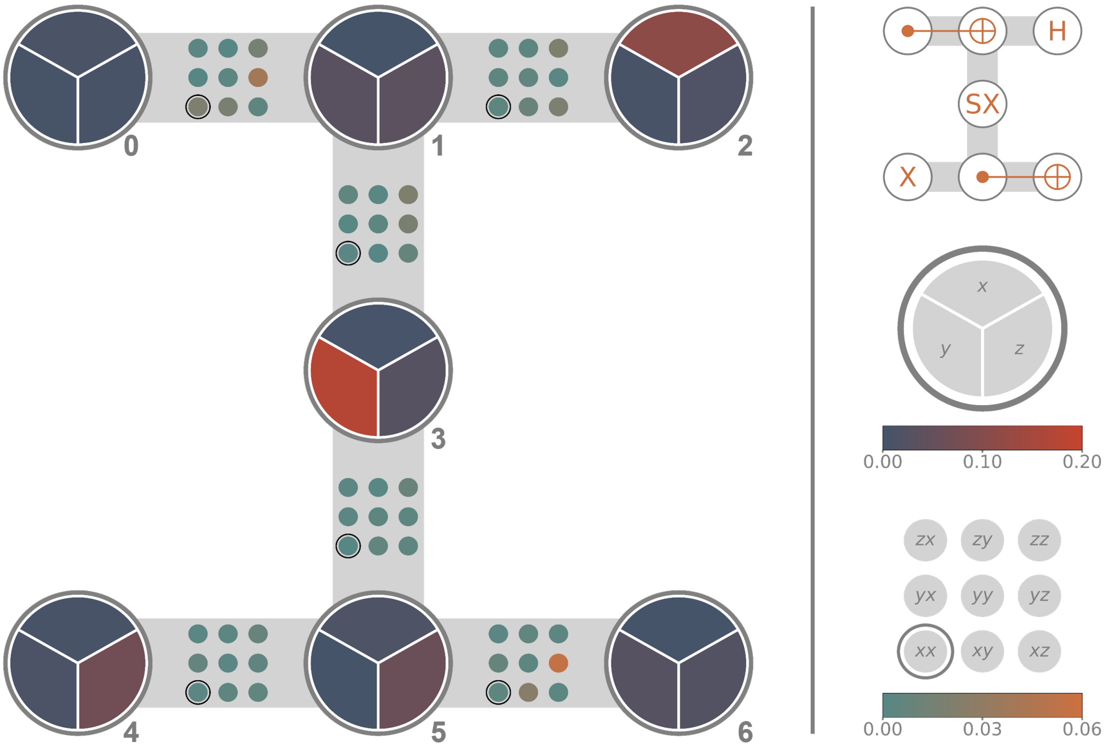


Figure 4.8.: *Coherent characterization of a gate-layer* - Coherent error characterization of the noise induced by the layer shown in the top right panel on the IBM Quantum Processor `ibm_lagos`. In the left plot, the characterized coherent errors overlay the connectivity tree. The big circles picture the single-qubit errors and the 9 small circles between two qubits display the two-qubit coherent errors.

Chapter 5

Mitigation

The fact that the protocols of chapter 4 demonstrate how to determine the corresponding model parameters of a process that behaves according to that model, does not imply that the characterized physical noise processes are truly represented by these models. In order to prove that the proposed characterization procedure may serve as a tool in the NISQ era for gate calibration, hardware development, or error mitigation, we must show that it can resemble aspects of real processes executed on existing hardware platforms. We aim to undertake this validation for the coherent-noise model by attempting to cancel the noise of an experimental process with a mitigation scheme that is heavily dependent on the characterized model parameters. This chapter introduces mitigation techniques related to Pauli noise, coherent noise, and readout errors. Those schemes are then applied in a two-step experiment for the coherent noise model. We first characterize a noisy process and then try to reclaim the noiseless measurement statistic by mitigating the process based on the extracted parameters.

5.1. Mitigation Schemes

At some point, quantum error correction will hopefully enable quantum fault tolerance. However, due to the gate overhead of quantum error correction schemes and the requirement of adaptive control, error correction is hardly applicable for quantum circuits on hardware platforms in the NISQ era [77]. In this phase of quantum computing, quantum error mitigation schemes enable simulations of simple quantum systems on noisy hardware without error correction [78]. While error correction aims to restore the noise-free output states in a noisy circuit, quantum error mitigation generally settles on finding bias-free measurement statistic [37]. The main differences between error mitigation protocols and error correction schemes are that error mitigation does not redundantly encode the information of a single qubit over many additional qubits and does not use adaptive operations that depend on syndrome measurements [77]. Quantum error mitigation protocols often execute multiple slightly different versions of a circuit, followed by classical post-processing that combines those results into a noise-free estimation of some observable. A prominent example that nicely illustrates the main idea of quantum error mitigation is quantum error extrapolation [37]. The idea is to estimate the expectation value of some observables for multiple instances of a circuit, where the circuits only differ in the amount of noise they contain. The amount of noise in the circuit can, for example, be changed by adding additional gates to the circuit. The expectation value for the circuit without noise is then extrapolated by fitting a curve through the expectation values at different amounts of noise. Since the protocol of this example depends very little on the actual noise model, it would not be feasible to validate our characterization results. Instead, we will discuss probabilistic error cancellation for mitigating Pauli errors and introduce a correction scheme for coherent errors. Lastly, we examine a readout error mitigation scheme used in hardware experiments to limit the influence of measurement errors on our characterization protocols.

5.1.1. Pauli Error Mitigation

When thinking about the contracting action of the Pauli channel on the Bloch sphere, at first glance, it seems that one can undo the action of the Pauli channel by simply applying a process that rescales the Bloch sphere to its initial size. However, such a process does not physically exist because it would require restoring lost information. For example, after applying the depolarizing channel to the states $|0\rangle$ and $|1\rangle$, the two states are no longer entirely distinguishable by a single measurement. As an additional physical channel can not increase the distinguishability of two states, it can not cancel the effect of the depolarizing channel. An alternative, chosen in this thesis, is to proceed with error mitigation. More specifically, we apply probabilistic error cancellation (PEC) [6] to mitigate the effect of the Pauli noise, closely following the approach of Berg *et al.* [32]. The central idea of PEC is to express the nonphysical inverse of some noise channel by a linear combination of physical channels. As each channel of the linear combination is physical, each can be applied to the noisy gate in a different circuit instance. The estimation results for some observable of all those circuits are then combined according to the linear combination weights. This procedure provides an estimation for the noise-free expectation value of the observable. In the following, we discuss the outlined procedure in the example of mitigating a Pauli noise channel.

The PTM of a Pauli channel $T_{\mathcal{P}}$ is diagonal. As discussed in section 2.3, the Pauli eigenvalues $f(b)$ on the diagonal of $T_{\mathcal{P}}$ are linked to the Pauli error rates $p(c)$ by a Walsh-Hadamard transform. The inverse of $T_{\mathcal{P}}$ corresponds to $T_{\mathcal{P}}^{\mathrm{inv}} = \mathrm{diag}\left(\frac{1}{f(b)}\right)$. By applying the Walsh-Hadamard transform on the inverted Pauli eigenvalues, we find the nonphysical Pauli error rates $q^{\mathrm{inv}}(c)$ that would correspond to the noise canceling process [32]. This transformation reads as

$$q^{\mathrm{inv}}(c) = \frac{1}{2^n} \sum_b (-1)^{\langle c,b\rangle_{sp}} \frac{1}{f(b)}. \tag{5.1}$$

In the example of a dephasing channel $\mathcal{Z}(\rho) = 0.8\ \rho + 0.2\ \mathrm{diag}(\rho)$, this would look like:

$$\begin{array}{ccc} \mathcal{Z}(\rho) = 0.9\ \mathbb{I}\rho\mathbb{I} + 0.1\ \sigma_z\rho\sigma_z & \Rightarrow & T_{\mathcal{Z}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.8 & 0 & 0 \\ 0 & 0 & 0.8 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ & & \Downarrow \\ \mathcal{Z}^{\mathrm{inv}}(\rho) = 1.125\ \mathbb{I}\rho\mathbb{I} - 0.125\ \sigma_z\rho\sigma_z & \Leftarrow & T_{\mathcal{Z}}^{\mathrm{inv}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1.25 & 0 & 0 \\ 0 & 0 & 1.25 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array} \tag{5.2}$$

As visible in this example of the dephasing channel, the weights of the linear combination do not resemble a probability distribution and contain negative values. However, by introducing the normalization factor $\gamma = \sum_c \left|q^{\mathrm{inv}}(c)\right|$, we can rescale $q^{\mathrm{inv}}(c)$ to $p^{\mathrm{inv}}(c) = \frac{1}{\gamma}\left|q^{\mathrm{inv}}(c)\right|$, which then represents a probability distribution. This allows us to rewrite the nonphysical Pauli channel in terms of $p^{\mathrm{inv}}(c)$ as

$$\mathcal{P}^{\mathrm{inv}}(\rho) = \sum_c q^{\mathrm{inv}}(c) P_c \rho P_c = \gamma \sum_c \mathrm{sgn}\left(q^{\mathrm{inv}}(c)\right) p^{\mathrm{inv}}(c) P_c \rho P_c. \tag{5.3}$$

The PEC error scheme applied by Berg *et al.* [32] contains the following steps in order to estimate the expectation value of some operator $\hat{O}$ of a state $U_I \rho U_I^\dagger$ when one has only access to the noisy version of the state $\mathcal{P}\left(U_I \rho U_I^\dagger\right)$. First, draw a Pauli operator P_c according to the probability distribution $p^{\mathrm{inv}}(c)$ and store the related sign $\mathrm{sgn}\left(q^{\mathrm{inv}}(c)\right)$. Add to the estimation circuit a gate sequence that implements P_c next to the noise channel. Estimate the expectation value of $\hat{O}$ by measuring this altered circuit, obtaining an estimate for

$$\mathrm{Tr}\left[\hat{O} P_c \left(\mathcal{P}\left(U_I \rho U_I^\dagger\right)\right) P_c\right]. \tag{5.4}$$

To rescale this expectation value according to Eq. (5.3), multiply it by the corresponding sign and scale it with the normalization factor γ, leading to

$$\gamma\ \mathrm{sgn}\left(q^{\mathrm{inv}}(c)\right)\ \mathrm{Tr}\left[\hat{O} P_c \left(\mathcal{P}\left(U_I \rho U_I^\dagger\right)\right) P_c\right]. \tag{5.5}$$

These steps are repeated for multiple samples of P_c. Finally, the average over all those rescaled expectation values estimates the unbiased expectation value of the operator. Fig. 5.1 shows the corresponding circuits for the dephasing noise channel example of Eq. (5.2).

The rescaling of the expectation values increases the variance of the expectation value by a factor of $\mathcal{O}\left(\gamma^2\right)$ compared to the noiseless case [32]. The protocol, therefore, creates a circuit overhead of factor γ^2. In order to mitigate multiple layers, for each layer, the mitigating Pauli is sampled according to the layer-specific probability distribution. The

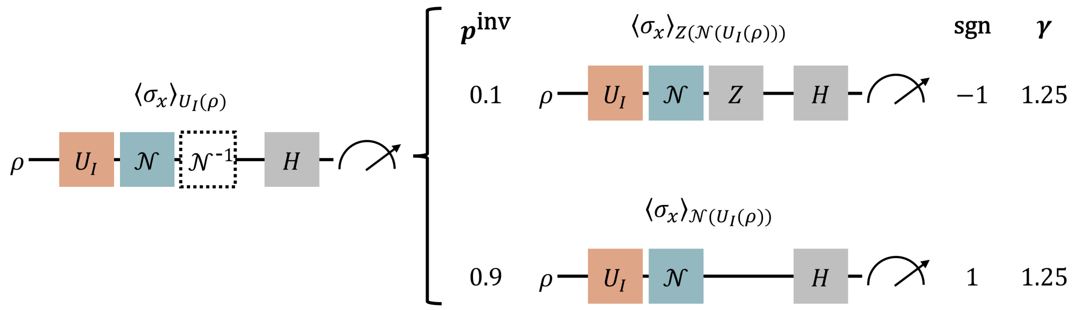


Figure 5.1.: *PEC dephasing example* - Schematic of the PEC procedure to estimate the expectation value of the operator σ_x for a state ρ evolved by the ideal unitary U_I. Experimentally we only have access to the noisy channel $\mathcal{N}(U_I \rho U_I^\dagger)$. For the example of this figure, the noise channel $\mathcal{N}$ corresponds to the channel $\mathcal{Z}$ of Eq. (5.2). As showed in Eq. (5.2), the nonphysical inverse process consists of a mixture of two Pauli operators. This figure shows the two corresponding mitigation circuits with their probabilities p^{inv}. By estimating the expectation value of σ_x for both circuits one can estimate the bias-free expectation value by scaling the obtained results with the sign of the circuit and the normalization factor γ.

weight of each circuit instance corresponds to the product of all layer-specific normalization factors γ and signs. This causes growth of the circuit overhead with an increasing circuit depth.

The PEC implementation in this thesis slightly differs from the described protocol. Instead of sampling P_c, we select all P_c corresponding to the r most significant $p^{\text{inv}}(c)$ and denote this set of c as $\overline{c}$. For all $P_{\overline{c}}$, we then run the mitigated circuit and restore the expectation value of the operator. We must rescale the expectation values differently because we have not sampled probabilistically. We weight each expectation value by its weight in the linear combination $q^{\text{inv}}(\overline{c})$ and scale by the normalization factor $\overline{\gamma} = \sum_{\overline{c}} q^{\text{inv}}(\overline{c})$. The noise channel is close to the identity channel for assumed error rates in the order of a few percent or lower. Consequently, $p^{\text{inv}}(\{0\}^n)$ is at least about an order of magnitude bigger than all other $p^{\text{inv}}(c)$. Therefore, when mitigating multiple layers in our truncation approach, we will only add a mitigating Pauli operator at one layer per circuit instance. The parameters of each layer are then multiplied analogous to the above protocol. An advantage of this truncation method is that it prevents the same circuit from being created and transpiled multiple times. We have not studied the implication of the truncation on the circuit overhead or the variance scaling. We expect slightly improved behavior due to reduced sampling noise since, in our method, we explicitly multiply the expectation values by their probability $p^{\text{inv}}(c)$ instead of sampling over the distribution $p^{\text{inv}}(c)$.

5.1.2. Coherent Error Mitigation

Unlike the Kraus representation of an incoherent noise channel, the Kraus representation of a coherent noise channel only consists of a single unitary operator. Rather than by a probabilistic mixture of different unitary evolutions, coherent noise is described by unitary dynamics. Therefore, the inverse of a unitary channel is also a valid and physically implementable channel. In the framework of the coherent noise model introduced in this thesis, this implies that the effect of the noise can be canceled by introducing single- and two-qubit rotations. Where the latter act opposite to the characterized noise rotations.

The ideal implementation of such a noise-correcting circuit element heavily depends on the native gate-set and the quantum hardware platform's error rates. It is desired to keep the circuit depth of the correcting scheme as low as possible to limit the additional noise introduced by the correcting circuit element. Implementing single-qubit rotations is straightforward on most hardware platforms and often only involves one or a few single-qubit gates. On the other hand, a two-qubit rotation requires at least one two-qubit entangling gate, which usually is of lower fidelity than the single-qubit gates. In this section, we come up with a correcting circuit element that allows correcting for all 15 single and two-qubit rotations. The element is based on the CX-gate, the native two-qubit gate on the used IBM Quantum platforms. We still operate under the assumption that the rotation angles are small and the rotations commute with each other.

The designed circuit element is parameterized by the 6 single- and 9 two-qubit rotation errors obtained from the characterization. We aim for a design where each of the 15 parameters is related to a different single-qubit rotation. The single-qubit rotations are interleaved in a structure of CX-gates and fixed rotations, in order to create two-qubit rotations. The sole structure of the fixed rotations and the CX-gates should correspond to the identity such that no correction is applied if no error has been detected.

A two-qubit rotation is created by enclosing a single-qubit rotation with two CX-gates, which is the entangling source in the circuit. The CX-gate commutes with the single-qubit rotations R_{zi} and R_{ix}, as visible from its unitary representation

$$U_{CX} = e^{-i\frac{\pi}{4}(\mathbb{I}-\sigma_z)_1 \otimes (\mathbb{I}-\sigma_x)_2}. \tag{5.6}$$

Therefore, out of the 6 possibilities of interleaving a single-qubit rotation between two CX-gates, only 4 lead to a two-qubit rotation. Those 4 created two-qubit rotations can then be altered to different two-qubit rotations by adding additional single-qubit rotations. While different rotations can be used to achieve this, when using IBM Quantum platforms [74], it is advantageous to use the Z-rotation gate as this corresponds to a virtual gate that does not introduce any errors [76].

From the circuit identities of Fig. B.1, we finally construct the mitigation gate shown in Fig. 5.2. While this is not a unique solution to the posed problem, we have linked each rotation error to one single-qubit rotation in the correcting gate. This allows us to connect this circuit element with the characterization results easily. Note that the circuit schematic in figure 5.2 still has to be transpiled to the native gate-set of the computing hardware, which likely increases the number of gates. For the circuit diagrams in the

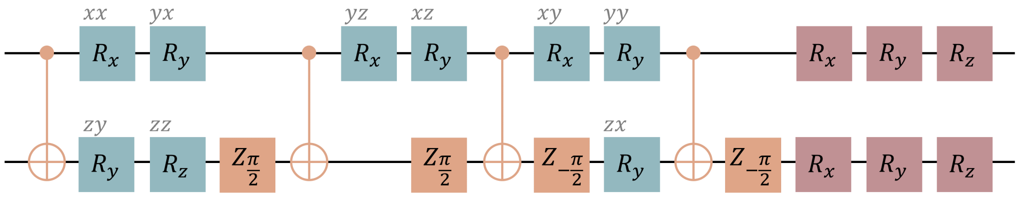


Figure 5.2.: *Coherent error mitigation element* - Circuit element to mitigate single- and two-qubit coherent errors. The orange circuit elements are fixed elements. Each blue single-qubit rotation corresponds to a different two-qubit rotation (indexed by the gray superscript) after it is propagated through the fixed structure. The red single-qubit rotations correct for single-qubit rotation errors. The angle of each single-qubit rotation depends on the characterization result.

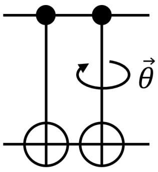


Figure 5.3.: *Mitigation element symbol* - Circuit symbol used to represent the circuit element of Fig. 5.2, where θ contains the parameters for the rotation. Neglecting the first CX-gate of the element, leads to a mitigated version of the CX-gate.

upcoming sections, we will use the symbol introduced in Fig. 5.3.

5.1.3. Readout Mitigation

The used IBM Quantum platforms report readout errors of the order of 1%, caused by relaxation, imperfect coupling of the readout resonator, and signal amplification [76]. To counter the influence of these errors, we apply a post-processing readout mitigation scheme in all of our experiments. The typical readout noise of superconducting platforms can be described well with a purely classical noise model [78]. According to

$$p_{\text{noisy}} = A\ p_{\text{ideal}}, \tag{5.7}$$

the readout model links the readout probabilities of the noisy measurement p_{noisy} to the distribution of an ideal measurement p_{ideal} with the matrix A [79]. Each vector contains the probabilities for all 2^n possible readout results of an n-qubit system. Consequently, A is a matrix of size $2^n \times 2^n$.

The readout noise of the used IBMq platform is mostly uncorrelated between qubits [79]. Therefore, we can express the full-system matrix A as the tensor product of the individual error matrices $A = A^{(n-1)} \otimes ... \otimes A^{(1)} \otimes A^{(0)}$. IBM Quantum lists the readout characterization terms $P^{(k)}(i|j)$ for each qubit. They correspond to the probability that the measurement of the computational basis state $|j\rangle$ leads to the readout value i. The matrix $A^{(k)}$ can then be expressed as

$$A^{(k)} = \begin{pmatrix} P^{(k)}(0|0) & P^{(k)}(0|1) \\ P^{(k)}(1|0) & P^{(k)}(1|1) \end{pmatrix}. \tag{5.8}$$

In the presented protocols, we assume a limited correlation length of the noise. This approach leads to post-processing procedures that never simultaneously analyze the full-system readout but only the readout of subsystems of size r. Therefore, Eq. (5.7) can be solved for each subsystem individually. It is tractable for subsystems of a few qubits to invert the matrix A, and we can find the ideal readout distribution p_{ideal} by applying A^{-1} to the measure noisy distribution p_{noisy}.

For the same circuit, a noisy measurement in combination with the readout mitigation leads to a higher measurement uncertainty than an ideal measurement [79]. Therefore, to limit the error in estimating an observable, a circuit with a noisy readout must be sampled more than a circuit with an ideal readout.

5.2. Experimental Results

This section intends to verify the utility of the proposed characterization method for a process running on a near-term device. We examine whether the model, output by the coherent noise characterization protocol, facilitates the mitigation of the coherent noise introduced by a process run on an IBM Quantum device. After mitigating the coherent errors, we proceed to characterize the Pauli errors and apply PEC to mitigate them. This procedure not only puts our characterization method to the test but also primarily the noise model itself. We will only be able to significantly suppress the effect of the noise on the measurement statistic if the coherent noise model of Eq. (3.5) is expressible enough to capture the dominant noise contributions of the probed process.

The introduced circuit element to mitigate coherent noise consists of 4 CX-gates and multiple single-qubit gates. Due to this, for NISQ devices considerable circuit depth, the element substantially introduces noise itself. Therefore, we need to incorporate the mitigation element with mitigation angles 0 into the characterization circuit, to advance to mitigate the coherent noise without adding a new noise source to the circuit.

To illustrate the effect of the different mitigation steps, we characterize and mitigate the sole coherent mitigation circuit element, where the angles of all parameterized rotations are set to 0. Ideally, this process corresponds to the identity. The procedure contains preparing a state in a Pauli basis, applying multiple repetitions of the process, and estimating all two-qubit Pauli expectation values. For the ideal process, we would measure an expectation value of 1 for all Pauli operators P_b of which the prepared state χ_0^a is an eigenstate ($a \star b = \{0\}^2$) and an expectation value of 0 for all Pauli operator P_c of which the prepared state χ_0^a is not an eigenstate ($a \star b \neq \{0\}^2$). Those expectation values are independent of the number of repetitions of the identity. Fig. 5.4 shows this ideal behavior for the prepared state $|+i\rangle$.

5.2.1. Hardware Platform

The characterization and mitigation are run on `ibm_lagos`, one of the IBM Quantum Falcon Processors [74]. We choose to characterize the coherent error mitigation element on qubits 1 and 2 as they report a low readout error. The specifications of the involved

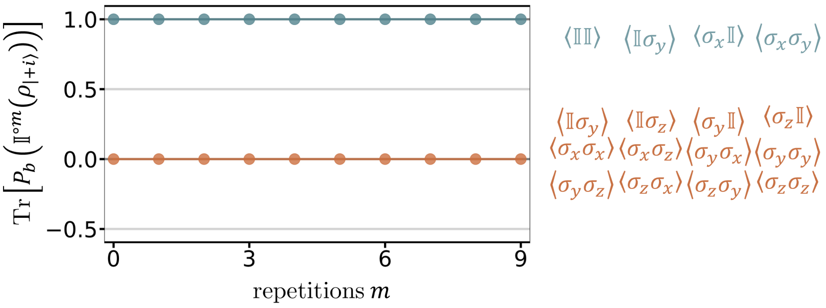
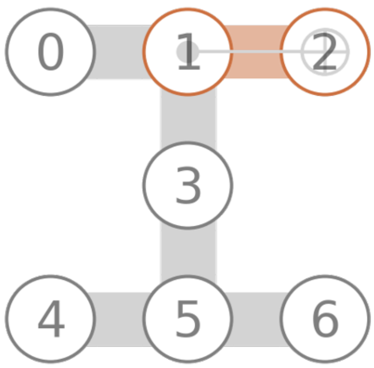


Figure 5.4.: *Ideal Pauli expectation values* - A plot of all 16 two-qubit Pauli expectation values of the state $|{+i}\rangle$ after it was evolved m times by the identity. By the nature of the identity, the expectation values are independent of the number of applied repetitions. The measured expectation value is 1 if $|{+i}\rangle$ is an eigenstate of the respective Pauli (shown in blue) and 0 otherwise (shown in orange).

qubits from the day of the experiment (provided by IBM) and the connectivity tree of the processor are shown in Fig. 5.5. The native gate-set of `ibm_lagos` contains the two-qubit CX-gate and the single-qubit gates X, SX, and RZ. The RZ-gate is only executed virtually and does not introduce any noise [76]. Transpiling the coherent noise mitigating circuit element to this native gate-set leads to the circuit of Fig. B.2. We apply readout error mitigation based on the reported readout errors in all the experiments according to the protocol of section 5.1.3. As the application of 9 instances of the noisy circuit element corresponds to 36 CX-gates and 126 single-qubit SX gates, the readout statistic of such a circuit on a NISQ device with no error correction or mitigation is highly influenced by errors. By simply multiplying up the reported error fidelities for these gates and not considering SPAM errors, the total error rate is estimated to be about 38%.



Qubits	Q1	Q2
T1	124.49 µs	122.45 µs
T2	121.82 µs	152.94 µs
sx-error	$1.89\ 10^{-4}$	$1.62\ 10^{-4}$
x-error	$1.89\ 10^{-4}$	$1.62\ 10^{-4}$
readout-error $p(1\|0)$	$6.60\ 10^{-3}$	$2.80\ 10^{-3}$
readout-error $p(0\|1)$	$1.02\ 10^{-2}$	$6.40\ 10^{-3}$
CX-gate		
CX-error	$6.40\ 10^{-3}$	
gate time	327.11 ns	

Figure 5.5.: *`ibm_lagos`* - Connectivity tree and specifications of the `ibm_lagos` Falcon Processor reported by IBM Quantum at the day of the experiment.

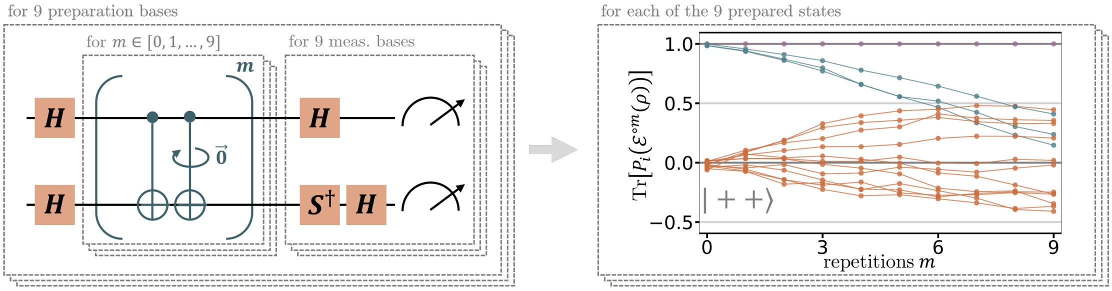


Figure 5.6.: *Coherent noise characterization experiment* - Circuit diagram used for characterization of the coherent noise mitigating circuit element. The right plot shows the Pauli expectation values for different numbers of repetitions of the element and the prepared state $|{+}{+}\rangle$. Expectation values that in the ideal case are 0 are plotted in orange, the blue curves correspond to the expectation values of the operators $\mathbb{I}X$, $X\mathbb{I}$, and XX, which are 1 in the noise-free-case, and the purple curve represents the identity.

5.2.2. Coherent Noise Characterization and Mitigation

We execute the noise characterization protocol of section 4.3 on the element of Fig. B.2. A schematic of the corresponding circuits is shown in the left half of Fig. 5.6. In total, 810 different circuits are executed, corresponding to all combinations of the 9 possible preparation bases, 0 to 9 repetitions of the noisy process, and the 9 Pauli measurements for two qubits. Per circuit, we measure 512 shots, leading to a total execution time in the system of approximately 2.5 minutes. We obtain for each prepared state a chart similar to the right chart in Fig. 5.6. Fig. B.3 shows the complete set of these plots for the unmitigated circuit.

Applying the post-processing procedure to estimate the single- and two-qubit rotations that resemble the coherent noise unitary leads to the following values:

single-qubit rotations								
$\theta_{X\mathbb{I}}$	$\theta_{Y\mathbb{I}}$	$\theta_{Z\mathbb{I}}$	$\theta_{\mathbb{I}X}$	$\theta_{\mathbb{I}Y}$	$\theta_{\mathbb{I}Z}$			
0.09	0.01	−0.15	−0.03	0.01	−0.03			
two-qubit rotations								
θ_{XX}	θ_{YX}	θ_{ZX}	θ_{XY}	θ_{YY}	θ_{ZY}	θ_{XZ}	θ_{YZ}	θ_{ZZ}
0.07	−0.00	0.03	−0.09	−0.03	−0.01	−0.02	−0.08	−0.01

The characterization shows expected symmetries in the characterized angles that arise from the propagation of the coherent error introduced by each of the 4 CX-gates to the end of the circuit element. For example, when propagated through a CX-gate, an x-rotation error on qubit 1 becomes a correlated two-qubit xx-rotation error by the same angle and vice versa. Therefore, assuming the CX-gate is the dominating noise source, $\theta_{X\mathbb{I}}$ is expected to be similar to θ_{XX}. Other such pairs are $\theta_{Y\mathbb{I}} \leftrightarrow \theta_{YX}$, $\theta_{\mathbb{I}Y} \leftrightarrow \theta_{ZY}$,

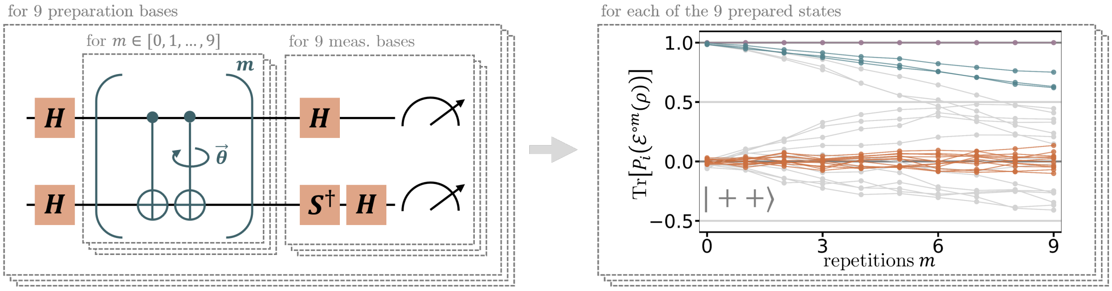


Figure 5.7.: *Coherent noise mitigation result* - Same plot as Fig. 5.6 for the mitigated circuit. In the right plot, the traces of the unmitigated case are plotted in light gray.

$\theta_{\mathbb{I}Z} \leftrightarrow \theta_{ZZ}$, $\theta_{XY} \leftrightarrow \theta_{YZ}$ and $\theta_{YY} \leftrightarrow \theta_{XZ}$.

To further improve the characterization, we rerun the characterization scheme, including the mitigation of the circuit element with the angles of the first characterization run. The number of shots for each of the 810 circuits is doubled to 1024. This iterative approach is expected to improve the characterization because it compensates for violating the assumed commutation of the single- and two-qubit rotation errors. We conduct a third iteration with 2048 shots based on the angles obtained from the second characterization. Table B.1 shows the rotation angles of all iterations. The obtained angles between the first and the last iteration do not change drastically. This observation supports the validity of the characterization results in section 4.3, where we only run one iteration of characterization.

Fig. 5.7 shows the coherent error mitigated Pauli measurement statistic for the prepared state $|{+}{+}\rangle$, where the circuit is mitigated by the angles obtained in the 3rd iteration of the characterization. Comparing Fig. 5.7 with Fig. 5.6, the effect of the coherent mitigation is clearly visible. The mitigation largely suppresses the diverging behavior of the expectation values that ideally are of value 0, although not for all preparation bases equally well, as visible in Fig. B.4. The expectation values of the Pauli operators of which the prepared state is an eigenstate of are raised in the mitigated case compared to the unmitigated estimation. However, the coherent noise mitigation does not retrieve their ideal-case expectation value of 1, as according to our model, the statistic is still affected by Pauli noise. Note that the trace corresponding to the expectation value of the identity does not correspond to a physical measurement and consequently appears shot-noise-free in the plots.

Beyond the described qualitative changes in the measurement statistic, we can quantitatively show the effect of coherent error mitigation by analyzing the fidelity of the ideal and the real output state of the circuit. As the ideal circuit equals the identity, the ideal output state corresponds to the ideal input state. We estimate the density matrix of the real output state by applying classical shadow tomography to the measured data. The standard state fidelity is used $F(\sigma, \rho) = \mathrm{Tr}\left[\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right]^2$. As visible in Fig. 5.8, the fidelities corresponding to the unmitigated circuit are significantly lower than the fidelity of the mitigated circuits. The mitigation leads to a decrease in the average infidelity by about a factor of two.

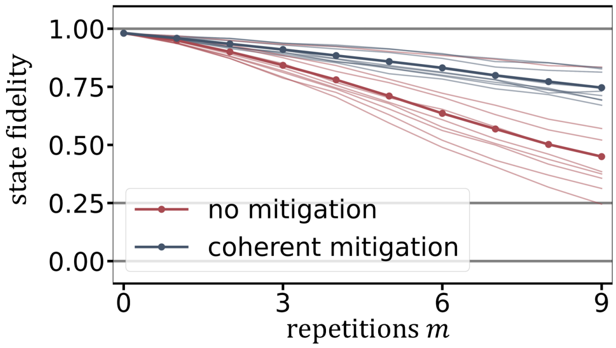


Figure 5.8.: *Fidelity measurements* - The fidelity of the ideal input state and the state after multiple applications of the noisy identity. Red and blue curves correspond to the evolution by the unmitigated, respectively, the coherent-error mitigated circuit element. The thin lines represent the fidelity of each of the 9 prepared states. The bold line is the average over all state preparation.

5.2.3. Pauli Error Characterization and PEC

After the mitigation of the coherent noise in the system, by the assumption of the introduced model, all remaining noise can be described by a Pauli channel. The Pauli eigenvalues are estimated from the Pauli expectation values of the mitigated circuits by applying the protocol of section 4.2. This is achieved by fitting an exponential to the blue curves in Fig. 5.7 and reading out the base of the exponent. By converting the Pauli expectation to Pauli error rates, the following values are obtained:

single-qubit Pauli error rates								
$\mathbb{I}\mathbb{I}$	$X\mathbb{I}$	$Y\mathbb{I}$	$Z\mathbb{I}$	$\mathbb{I}X$	$\mathbb{I}Y$	$\mathbb{I}Z$		
0.968	0.002	0.000	0.006	0.001	0.001	0.006		
two-qubit Pauli error rates								
XX	YX	ZX	XY	YY	ZY	XZ	YZ	ZZ
0.002	0.001	0.002	0.001	0.001	0.003	0.000	0.001	0.005

In accordance with the findings in Ref. [32] for another IBM Quantum hardware platform, the most significant Pauli-errors correspond to Z-errors. We apply the presented truncated version of the PEC protocol with a truncation depth of 6. This leads to a total of 2520 circuits, where 512 shots each are taken. The total time in the system was approximately 10 minutes. Due to the high complexity in the number of circuit measurements, we restricted the measurement to the Pauli expectation values for a single prepared state ($|{++}\rangle$), shown in Fig. 5.9.

Since the data seems to contain more shot-noise, a clear statement on whether the

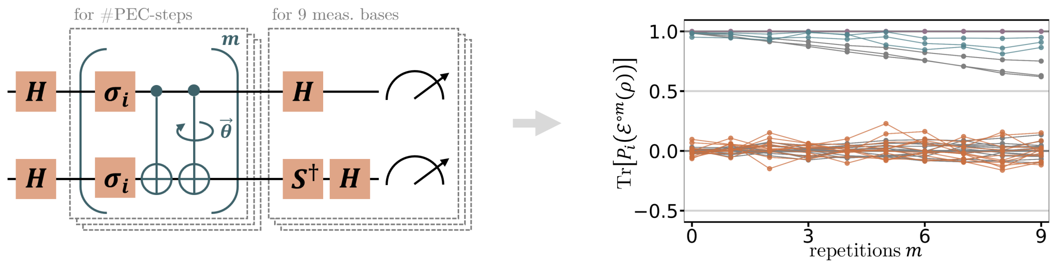


Figure 5.9.: *PEC results* - Schematics of the circuit, including PEC and coherent error mitigation. The plot shows the estimated Pauli expectation values. The gray traces are the expectation values of the circuit with coherent error mitigation but no PEC.

PEC improves the measurement statistic compared to the sole coherent noise mitigation is impossible. However, qualitatively for this preparation basis, it appears that the PEC leads to slightly higher expectation values for the Pauli operators of which the prepared state is an eigenstate. At the same time, PEC did not greatly change the spreading of the expectation values of the other operators.

5.2.4. Analysis

We managed to suppress the expectation of all operators of which the prepared state is not an eigenstate (the orange curves in figures 5.7, 5.6, and 5.9). This validates that the coherent noise characterization protocol is feasible to execute on real hardware and provides physically insightful information about the noisy process. Furthermore, the suppression implies that the Pauli transfer matrix that describes the noise after the coherent mitigation steps is predominantly diagonal. Otherwise, the initially non-zero expectation values would lead, at least for one preparation basis, to the increase of an initially zero expectation value. As a diagonal transfer matrix represents a Pauli channel, the model of a sequential action of a coherent term and a Pauli channel on the ideal unitary manages to resemble the main aspects of the real structure of the noise. This is also supported by the fact that PEC further improves the recovery of the ideal measurement statistic. We run the complete protocol also on `ibm_jakarta` and obtained similar results, shown in Fig. B.5.

Chapter 6

Conclusion

Motivated by the importance of physical insights into noise processes for improving quantum computing platforms, this thesis aimed to expand existing characterization protocols to include additional types of noise.

In the thesis's first part, we proposed the parallel Pauli noise model as a new approach to model the incoherent noise originating from a unitary process. By combining concepts of shadow tomography, existing Pauli estimation protocols, and newly developed estimation schemes, we managed to derive a characterization protocol for the model. We then extended the protocol beyond the characterization of Pauli channels and included decay. In a simulation, we confirmed that the protocol is feasible to execute on near-term devices. The biggest open question regarding this model is whether it manages to resemble the noisy behavior of processes executed on real devices. It would be interesting to compare the model to the more common sequential Pauli noise model. However, to do this, one must first address the susceptibility of the proposed protocol to SPAM errors.

In the thesis's second part, we turned to more established noise reconstruction protocols. We observed, as expected, that Pauli channels on their own can not explain the noise of a process executed on a superconducting platform of IBM. We then expanded a standard noise model to include coherent noise and confirmed that this model resembles the main noise contributions of the examined process. We developed a coherent noise characterization protocol to estimate the coherent noise introduced by a process. The protocol characterizes the coherent noise in terms of single- and two-qubit rotations and is thereby physically insightful. Under the assumption of limited noise correlation lengths, the protocol is scalable. To the best of our knowledge, this is the first protocol that characterizes correlated coherent noise of gate-layers on a NISQ device.

We created a circuit element that allows correcting the characterized coherent errors. Together with established techniques to mitigate readout errors and Pauli noise, this element allowed us to significantly suppress the effects of the noise on the analyzed process. While we illustrated the utility of the characterization for a gate-level correction of the coherent noise, one could also use the characterized over- and under-rotation as a starting point for calibration or as guidance for hardware improvements.

As a possible model expansion, adding decay would be interesting. We have not analyzed the resistivity of the protocol to SPAM errors. Furthermore, we believe it should be possible to decouple the estimation process of the single- and two-qubit rotations from one another. This methodology could also be necessary to create a SPAM-resistant protocol. It would also be interesting to run the protocol on different quantum computing hardware platforms. Finally, further investigation into how to characterize gate-specific Pauli noise for more general unitaries would allow applying the presented mitigation scheme to a broader range of processes.

Appendix A

Characterization

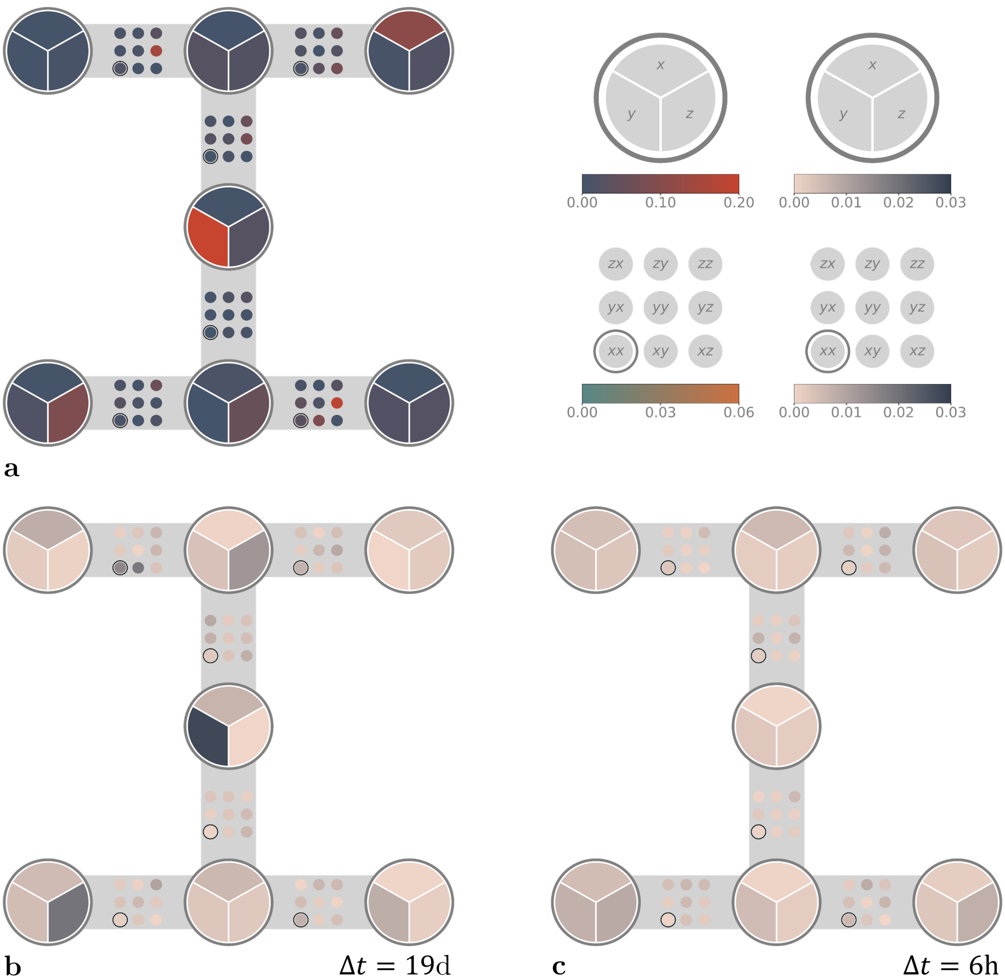


Figure A.1.: *Coherent noise drift* - Coherent error characterization analogous to the experiment of section 4.3.2, conducted at 3 different times. **a** shows the experimental results from October 11, 2022. **b** shows the difference of the characterization in (a) to the results of Fig. 4.8 from September 22, 2022. **c** shows the difference of the characterization in (a) to an identical characterization conducted 6 hours earlier.

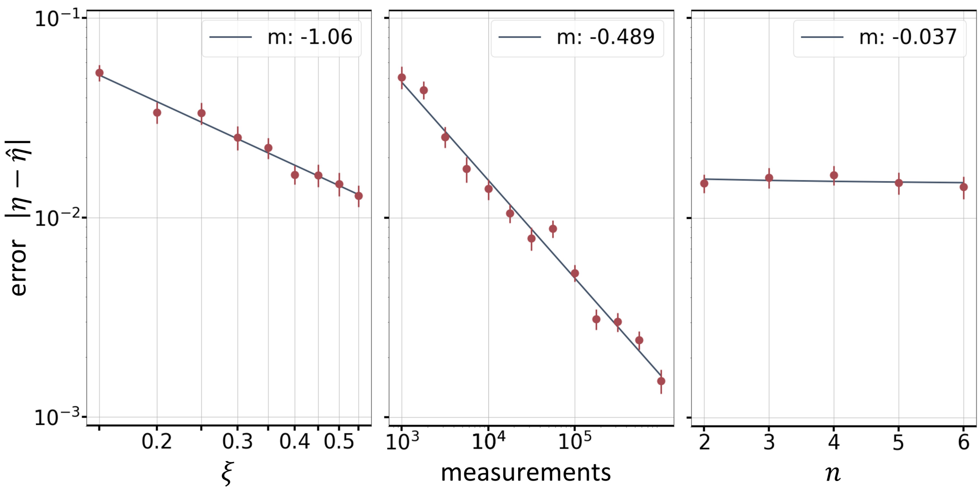


Figure A.2.: *η estimation error* - Plots of the expected error in the estimation of η. For each of the plot we fix two out of the three parameters ξ, number of measurements and number of qubits. We then estimate the error for different values of the third parameter. m represents the slope of the line obtained by linear regression in the logarithmic base.

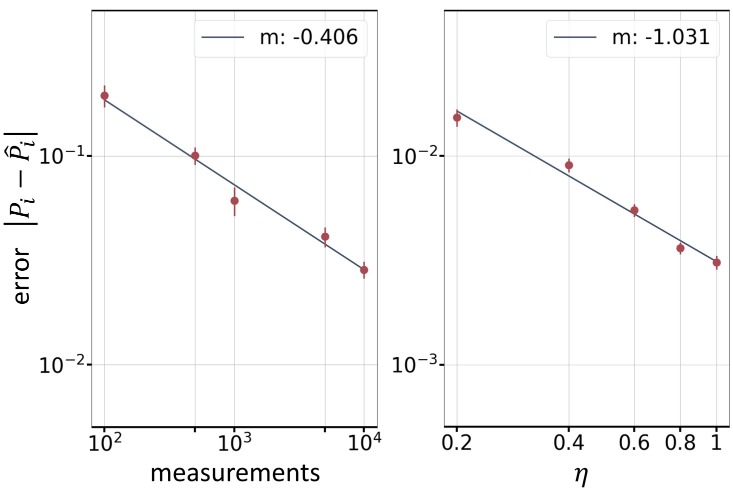


Figure A.3.: *Pauli error rate estimation error* - Plots of the expected estimation error in the estimation of the Pauli error rates. The plots are created in the absence of decay. The average estimation error is plotted against the number of circuit measurement with fixed η in the left plot. In the right plot, η is varied while the number of measurements is kept constant. m represents the slope of the line obtained by linear regression in the logarithmic base.

Appendix B

Mitigation

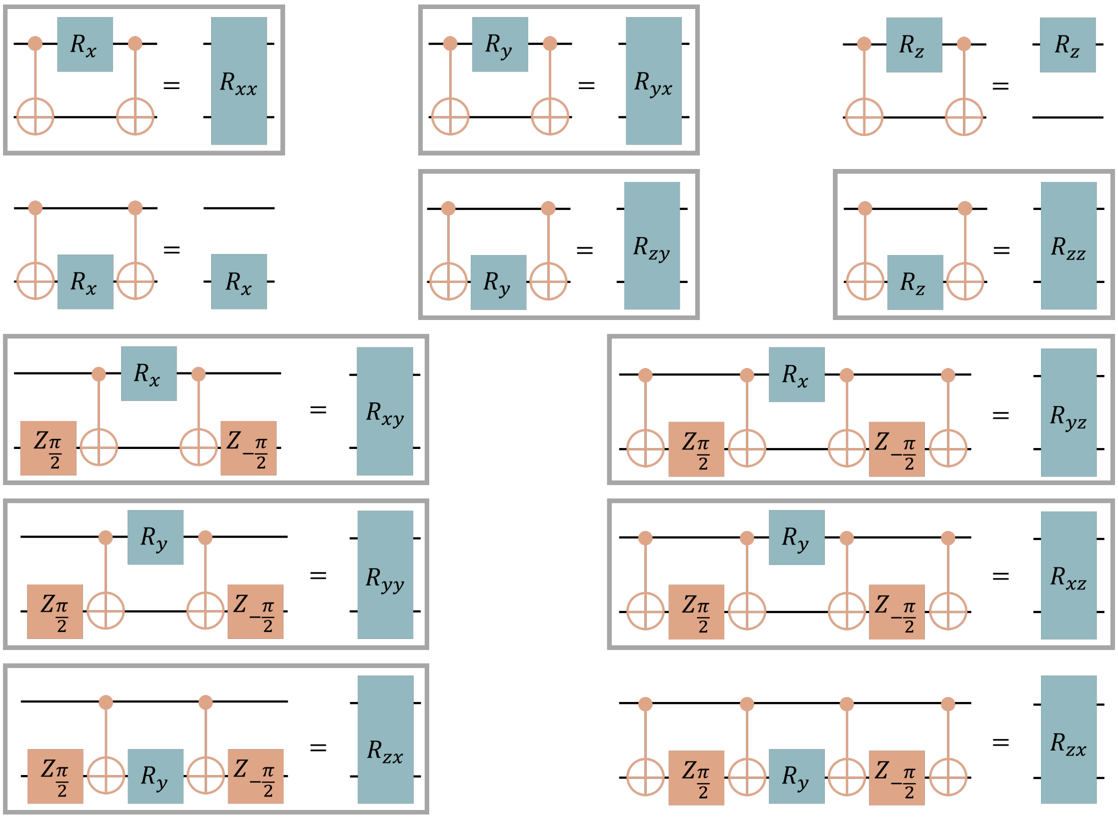


Figure B.1.: *Qubit rotation circuit identities* - Circuit identities to construct two-qubit rotation from single-qubit rotations and the CX-gate. The gate $Z_{\frac{\pi}{2}}$, corresponds to a single-qubit z-rotation by $\frac{\pi}{2}$. The framed identities are used in the construction of the circuit element of Fig. 5.2.

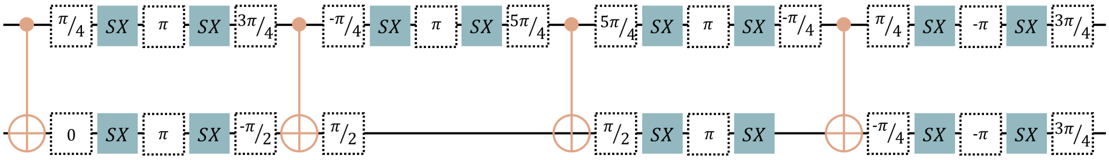


Figure B.2.: *Transpiled circuit element* - The coherent error mitigating circuit element of Fig. 5.2, with all correction angles set to 0, transpiled to the native gate-set of `ibm_lagos`. The dashed gates correspond to virtualy executed RZ-gates by the specified angle. We manually prevent simplifications of the circuit during the transpilation by inserting circuit barriers.

single-qubit rotations

	$\theta_{X\mathbb{I}}$	$\theta_{Y\mathbb{I}}$	$\theta_{Z\mathbb{I}}$	$\theta_{\mathbb{I}X}$	$\theta_{\mathbb{I}Y}$	$\theta_{\mathbb{I}Z}$
1st	0.09	0.01	−0.15	−0.03	0.01	−0.03
2nd	0.06	−0.01	−0.13	−0.00	0.00	−0.02
3rd	0.08	−0.01	−0.13	−0.01	0.01	−0.03

two-qubit rotations

	θ_{XX}	θ_{YX}	θ_{ZX}	θ_{XY}	θ_{YY}	θ_{ZY}	θ_{XZ}	θ_{YZ}	θ_{ZZ}
1st	0.07	−0.00	0.03	−0.09	−0.03	−0.01	−0.02	−0.08	−0.01
2nd	0.06	−0.01	0.01	−0.06	−0.03	0.02	0.03	−0.10	−0.01
3rd	0.06	−0.01	0.02	−0.06	−0.02	−0.01	0.03	−0.07	−0.01

Table B.1.: *Characterized coherent noise angles* - Characterized single- and two-qubit rotation angles, for the three iterations of the coherent noise characterization described in section 5.2.2. The circuits for the second and third round of characterization are already mitigated based on the results of the previous characterization run.

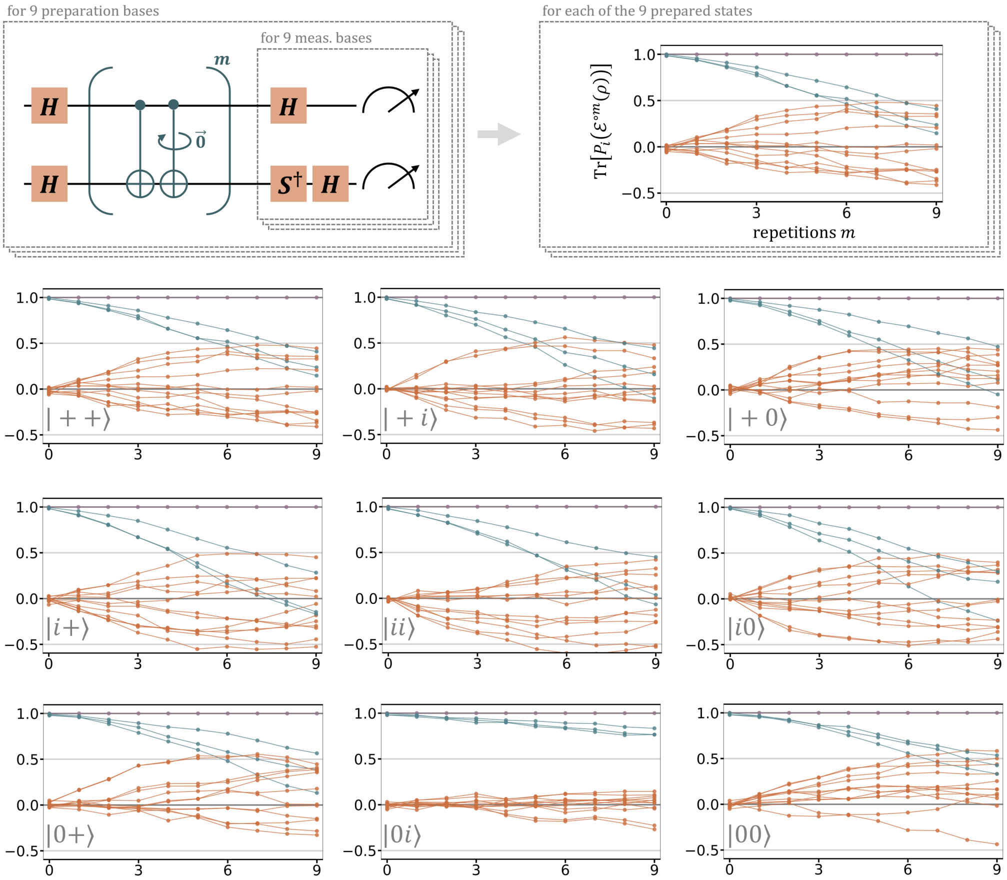


Figure B.3.: *Complete data set from characterization* - Collected data for the first coherent error characterization iteration in section 5.2.2. Each of the 9 plots below the circuit scheme at the top of the figure corresponds to a different prepared state. The state are labeled at the left bottom of each plot. The expectation values of each single- and two-qubit Pauli operator is plotted against the number of repetitions of the noisy identity. The purple trace corresponds to the identity, the blue curves to Pauli operators of which the prepared state is an eigenstate of, and the orange curves to all other Pauli operators. In the noiseless case the blue curves are expected to be 1 and the orange curves to be 0.

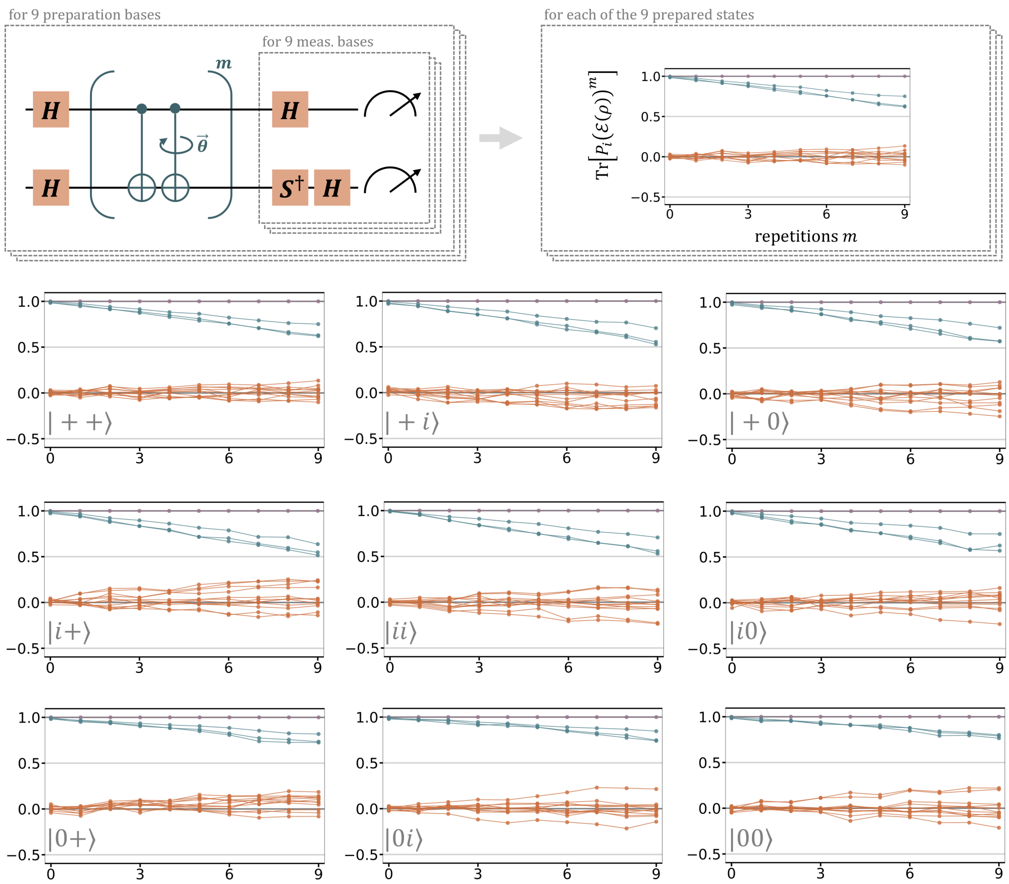
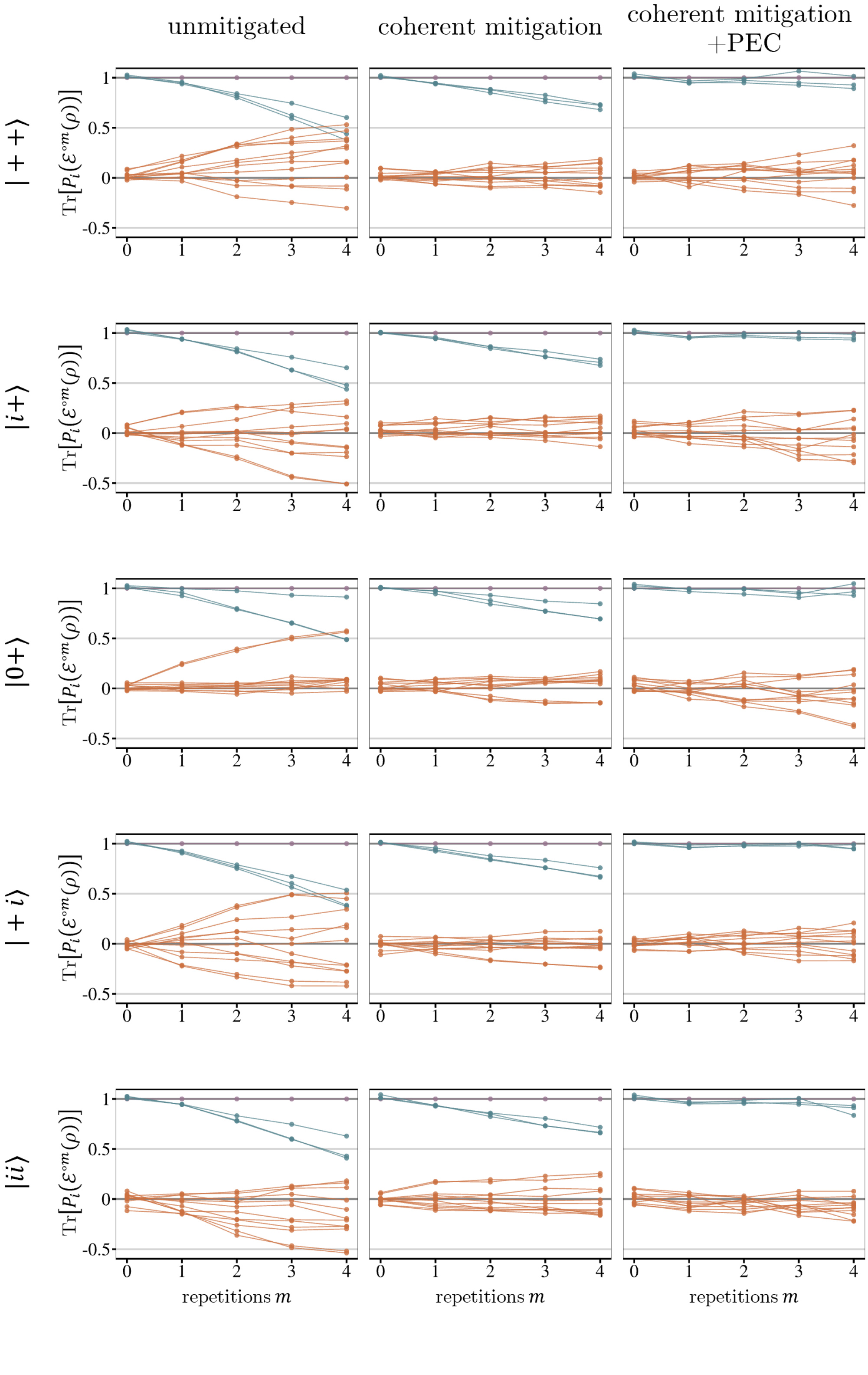


Figure B.4.: *Complete data set from coherent error mitigation* - Collected data of the coherent error mitigated experiment in section 5.2.2. Each of the 9 plots below the circuit scheme at the top of the figure corresponds to a different prepared state. The state are labeled at the left bottom of each plot. The expectation values of each single- and two-qubit Pauli operator is plotted against the number of repetitions of the noisy identity. The purple trace corresponds to the identity, the blue curves to Pauli operators of which the prepared state is an eigenstate of, and the orange curves to all other Pauli operators. In the noiseless case the blue curves are expected to be 1 and the orange curves to be 0. The Pauli eigenvalues are estimated from those plots by fitting an exponential to the blue curves.


unmitigated
coherent mitigation
coherent mitigation +PEC
$|++\rangle$
$|i+\rangle$
$|0+\rangle$
$|+i\rangle$
$|ii\rangle$
$\mathrm{Tr}[P_i(\mathcal{E}^{\circ m}(\rho))]$
1
0.5
0
-0.5
0
1
2
3
4
repetitions m

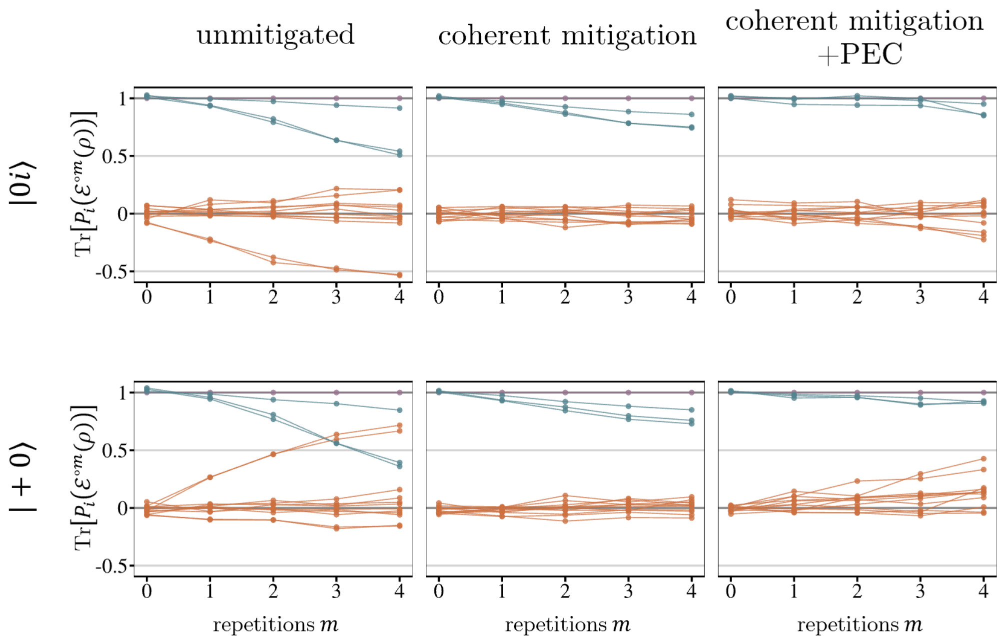


Figure B.5.: *Experimental data of* `ibmq_jakarta` - Data from the characterization, coherent error mitigation and probabilistic error cancellation protocols executed on `ibmq_jakarta`. The experiment is conducted analogous to the presented procedure on `ibmq_lagos` in section 5.2. We only applied the probabilistic error cancellation protocol to 7 states due to run-time constraints on the hardware platform.

Bibliography

1. Preskill, J. Quantum computing in the NISQ era and beyond. *Quantum* **2,** 79 (2018).
2. Martinis, J. M. Qubit metrology for building a fault-tolerant quantum computer. *npj Quantum Inf.* **1,** 1–3 (2015).
3. Nielsen, E. *et al.* Probing quantum processor performance with pyGSTi. *Quantum science and technology* **5,** 044002 (2020).
4. Jiang, W., Xiong, J. & Shi, Y. A co-design framework of neural networks and quantum circuits towards quantum advantage. *Nature communications* **12,** 1–13 (2021).
5. Harper, R., Flammia, S. T. & Wallman, J. J. Efficient learning of quantum noise. *Nat. Phys.* **16,** 1184–1188 (2020).
6. Temme, K., Bravyi, S. & Gambetta, J. M. Error mitigation for short-depth quantum circuits. *Physical review letters* **119,** 180509 (2017).
7. Nielsen, M. A. & Chuang, I. *Quantum computation and quantum information* 2002.
8. Knill, E. *et al.* Randomized benchmarking of quantum gates. *Phys. Rev. A* **77,** 012307 (2008).
9. Dahlhauser, M. L. & Humble, T. S. Benchmarking characterization methods for noisy quantum circuits. *arXiv preprint arXiv:2201.02243* (2022).
10. Flammia, S. T., Gross, D., Liu, Y.-K. & Eisert, J. Quantum tomography via compressed sensing: error bounds, sample complexity and efficient estimators. *New J. Phys.* **14,** 095022 (2012).
11. Holzäpfel, M., Baumgratz, T., Cramer, M. & Plenio, M. B. Scalable reconstruction of unitary processes and Hamiltonians. *Phys. Rev. A* **91,** 042129 (2015).
12. Magesan, E., Gambetta, J. M. & Emerson, J. Scalable and robust randomized benchmarking of quantum processes. *Physical review letters* **106,** 180504 (2011).
13. Flammia, S. T. & Liu, Y.-K. Direct fidelity estimation from few Pauli measurements. *Physical review letters* **106,** 230501 (2011).
14. Chuang, I. & Nielsen, M. A. Prescription for experimental determination of the dynamics of a quantum black box. *J. Mod. Opt.* **44,** 2455–2467 (1997).
15. Poyatos, J., Cirac, J. I. & Zoller, P. Complete characterization of a quantum process: the two-bit quantum gate. *Phys. Rev. Lett.* **78,** 390 (1997).
16. Mohseni, M., Rezakhani, A. T. & Lidar, D. A. Quantum-process tomography: Resource analysis of different strategies. *Phys. Rev. A* **77,** 032322 (2008).
17. Choi, M.-D. Completely positive linear maps on complex matrices. *Linear algebra and its applications* **10,** 285–290 (1975).

18. Jamiołkowski, A. Linear transformations which preserve trace and positive semidefiniteness of operators. *Rep. Math. Phys.* **3,** 275–278 (1972).

19. Govia, L., Ribeill, G., Ristè, D., Ware, M. & Krovi, H. Bootstrapping quantum process tomography via a perturbative ansatz. *Nature communications* **11,** 1–9 (2020).

20. Levy, R., Luo, D. & Clark, B. K. Classical shadows for quantum process tomography on near-term quantum computers. *arXiv preprint arXiv:2110.02965* (2021).

21. Kunjummen, J., Tran, M. C., Carney, D. & Taylor, J. M. Shadow process tomography of quantum channels. *arXiv preprint arXiv:2110.03629* (2021).

22. Merkel, S. T. *et al.* Self-consistent quantum process tomography. *Phys. Rev. A* **87,** 062119 (2013).

23. Blume-Kohout, R. *et al.* Robust, self-consistent, closed-form tomography of quantum logic gates on a trapped ion qubit. *arXiv preprint arXiv:1310.4492* (2013).

24. Nielsen, E. *et al.* Gate set tomography. *Quantum* **5,** 557 (2021).

25. Greenbaum, D. Introduction to quantum gate set tomography. *arXiv preprint arXiv:1509.02921* (2015).

26. Emerson, J., Alicki, R. & Życzkowski, K. Scalable noise estimation with random unitary operators. *J. Opt. B: Quantum Semiclassical Opt.* **7,** S347 (2005).

27. Cross, A. W., Magesan, E., Bishop, L. S., Smolin, J. A. & Gambetta, J. M. Scalable randomised benchmarking of non-Clifford gates. *npj Quantum Inf.* **2,** 1–5 (2016).

28. Carignan-Dugas, A., Wallman, J. J. & Emerson, J. Characterizing universal gate sets via dihedral benchmarking. *Phys. Rev. A* **92,** 060302 (2015).

29. Magesan, E. *et al.* Efficient measurement of quantum gate error by interleaved randomized benchmarking. *Physical review letters* **109,** 080505 (2012).

30. Wallman, J., Granade, C., Harper, R. & Flammia, S. T. Estimating the coherence of noise. *New J. Phys.* **17,** 113020 (2015).

31. Erhard, A. *et al.* Characterizing large-scale quantum computers via cycle benchmarking. *Nature communications* **10,** 1–7 (2019).

32. Berg, E. v. d., Minev, Z. K., Kandala, A. & Temme, K. Probabilistic error cancellation with sparse Pauli-Lindblad models on noisy quantum processors. *arXiv preprint arXiv:2201.09866* (2022).

33. Flammia, S. T. & Wallman, J. J. Efficient estimation of Pauli channels. *ACM Transactions on Quantum Computing* **1,** 1–32 (2020).

34. Geller, M. R. & Zhou, Z. Efficient error models for fault-tolerant architectures and the Pauli twirling approximation. *Phys. Rev. A* **88,** 012314 (2013).

35. Barnes, J. P., Trout, C. J., Lucarelli, D. & Clader, B. Quantum error-correction failure distributions: Comparison of coherent and stochastic error models. *Phys. Rev. A* **95,** 062338 (2017).

36. Cai, Z. *et al.* Quantum Error Mitigation. *arXiv preprint arXiv:2210.00921* (2022).
37. Endo, S., Cai, Z., Benjamin, S. C. & Yuan, X. Hybrid quantum-classical algorithms and quantum error mitigation. *J. Phys. Soc. Jpn.* **90,** 032001 (2021).
38. Greene, A. *et al.* Error mitigation via stabilizer measurement emulation. *arXiv preprint arXiv:2102.05767* (2021).
39. Flammia, S. T. & O'Donnell, R. Pauli error estimation via Population Recovery. *Quantum* **5,** 549 (2021).
40. Bloch, F. Nuclear Induction. *Phys. Rev.* **70,** 460–474. `https://link.aps.org/doi/10.1103/PhysRev.70.460` (7-8 1946).
41. Wie, C.-R. Bloch sphere model for two-qubit pure states. *arXiv preprint arXiv:1403.8069* (2014).
42. Gamel, O. Entangled Bloch spheres: Bloch matrix and two-qubit state space. *Phys. Rev. A* **93,** 062320 (2016).
43. Jakóbczyk, L. & Siennicki, M. Geometry of Bloch vectors in two-qubit system. *Phys. Lett. A* **286,** 383–390 (2001).
44. Mosseri, R. & Dandoloff, R. Geometry of entangled states, Bloch spheres and Hopf fibrations. *J. Phys. A: Math. Gen.* **34,** 10243 (2001).
45. Jevtic, S., Pusey, M., Jennings, D. & Rudolph, T. Quantum steering ellipsoids. *Physical review letters* **113,** 020402 (2014).
46. Milne, A., Jevtic, S., Jennings, D., Wiseman, H. & Rudolph, T. Quantum steering ellipsoids, extremal physical states and monogamy. *New J. Phys.* **16,** 083017 (2014).
47. Bravyi, S. & Kitaev, A. Universal quantum computation with ideal Clifford gates and noisy ancillas. *Phys. Rev. A* **71,** 022316 (2005).
48. Aaronson, S. *Shadow tomography of quantum states* in *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing* (2018), 325–338.
49. Huang, H.-Y., Kueng, R. & Preskill, J. Predicting many properties of a quantum system from very few measurements. *Nat. Phys.* **16,** 1050–1057 (2020).
50. Huang, H.-Y. Learning quantum states from their classical shadows. *Nature Reviews Physics* **4,** 81–81 (2022).
51. Huang, H.-Y., Kueng, R., Torlai, G., Albert, V. V. & Preskill, J. Provably efficient machine learning for quantum many-body problems. *arXiv preprint arXiv:2106.12627* (2021).
52. King, C. & Ruskai, M. B. Minimal entropy of states emerging from noisy quantum channels. *IEEE Transactions on information theory* **47,** 192–209 (2001).
53. Ruskai, M. B., Szarek, S. & Werner, E. An analysis of completely-positive trace-preserving maps on M2. *Linear algebra and its applications* **347,** 159–187 (2002).
54. Tomita, Y. & Svore, K. M. Low-distance surface codes under realistic quantum noise. *Phys. Rev. A* **90,** 062320 (2014).

55. Pauli, W. in *Wolfgang Pauli* 282–305 (Springer, 1988).
56. Baylis, W. E. *Clifford (Geometric) Algebras: with applications to physics, mathematics, and engineering* (Springer Science & Business Media, 2012).
57. Delanghe, R., Sommen, F. & Soucek, V. *Clifford algebra and spinor-valued functions: a function theory for the Dirac operator* (Springer Science & Business Media, 2012).
58. Kunz, H. O. On the equivalence between one-dimensional discrete Walsh-Hadamard and multidimensional discrete Fourier transforms. *IEEE Trans. Comput.* **28,** 267–268 (1979).
59. Resch, S. & Karpuzcu, U. R. Benchmarking quantum computers and the impact of quantum noise. *ACM Computing Surveys (CSUR)* **54,** 1–35 (2021).
60. Wallman, J., Granade, C., Harper, R. & Flammia, S. T. Estimating the coherence of noise. *New J. Phys.* **17,** 113020 (2015).
61. Bravyi, S., Englbrecht, M., König, R. & Peard, N. Correcting coherent errors with surface codes. *npj Quantum Inf.* **4,** 1–6 (2018).
62. Huang, E., Doherty, A. C. & Flammia, S. Performance of quantum error correction with coherent errors. *Phys. Rev. A* **99,** 022313 (2019).
63. Magesan, E., Puzzuoli, D., Granade, C. E. & Cory, D. G. Modeling quantum noise for efficient testing of fault-tolerant circuits. *Phys. Rev. A* **87,** 012324 (2013).
64. Sanders, Y. R., Wallman, J. J. & Sanders, B. C. Bounding quantum gate error rate based on reported average fidelity. *New J. Phys.* **18,** 012002 (2015).
65. Knill, E. Quantum computing with realistically noisy devices. *Nature* **434,** 39–44 (2005).
66. Wallman, J. J. & Emerson, J. Noise tailoring for scalable quantum computation via randomized compiling. *Phys. Rev. A* **94,** 052325 (2016).
67. Ware, M. *et al.* Experimental Pauli-frame randomization on a superconducting qubit. *Phys. Rev. A* **103,** 042604 (2021).
68. Gutiérrez, M., Svec, L., Vargo, A. & Brown, K. R. Approximation of realistic errors by Clifford channels and Pauli measurements. *Phys. Rev. A* **87,** 030302 (2013).
69. Devitt, S. J., Munro, W. J. & Nemoto, K. Quantum error correction for beginners. *Rep. Prog. Phys.* **76,** 076001 (2013).
70. Kueng, R., Long, D. M., Doherty, A. C. & Flammia, S. T. Comparing experiments to the fault-tolerance threshold. *Physical review letters* **117,** 170502 (2016).
71. Suzuki, Y., Fujii, K. & Koashi, M. Efficient simulation of quantum error correction under coherent error based on the nonunitary free-fermionic formalism. *Physical review letters* **119,** 190503 (2017).
72. Gottesman, D. The Heisenberg representation of quantum computers. *arXiv preprint quant-ph/9807006* (1998).
73. ANIS, M. S. *et al. Qiskit: An Open-source Framework for Quantum Computing* 2021.

74. IBM Quantum. `https://quantum-computing.ibm.com/` (2021).
75. Chen, S. *et al.* The learnability of Pauli noise. *arXiv preprint arXiv:2206.06362* (2022).
76. Bultrini, D., Gordon, M. H., López, E. & Sierra, G. Simple mitigation strategy for a systematic gate error in IBMQ. *arXiv preprint arXiv:2012.00831* (2020).
77. Takagi, R., Endo, S., Minagawa, S. & Gu, M. Fundamental limits of quantum error mitigation. *arXiv preprint arXiv:2109.04457* (2021).
78. Bravyi, S., Sheldon, S., Kandala, A., Mckay, D. C. & Gambetta, J. M. Mitigating measurement errors in multiqubit experiments. *Phys. Rev. A* **103,** 042605 (2021).
79. Nation, P. D., Kang, H., Sundaresan, N. & Gambetta, J. M. Scalable mitigation of measurement errors on quantum computers. *PRX Quantum* **2,** 040326 (2021).

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Declaration of originality

The signed declaration of originality is a component of every semester paper, Bachelor's thesis, Master's thesis and any other degree paper undertaken during the course of studies, including the respective electronic versions.

Lecturers may also require a declaration of originality for other written papers compiled for their courses.

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by the supervisor.

Title of work (in block letters):

Noise Characterization of Near-Term Quantum Devices

Authored by (in block letters):
For papers written by groups the names of all authors are required.

Name(s):	**First name(s):**
Kaufmann	Noah

With my signature I confirm that
- I have committed none of the forms of plagiarism described in the 'Citation etiquette' information sheet.
- I have documented all methods, data and processes truthfully.
- I have not manipulated any data.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for plagiarism.

Place, date
Zürich, 14.10.2022

Signature(s)
N. K

For papers written by groups the names of all authors are required. Their signatures collectively guarantee the entire content of the written paper.