Dissertation

submitted to the

Combined Faculty of Mathematics, Engineering and Natural Sciences

of

Heidelberg University, Germany

for the degree of

Doctor of Natural Sciences

Put forward by

Adrian Skasberg Aasen

born in: Kongsvinger, Norway

Oral examination: 30.01.2026

EXTRACTING INFORMATION FROM NOISY QUANTUM SYSTEMS

Readout-error mitigation and adaptive measurement strategies

Referees: Prof. Dr. Martin Gärttner
Prof. Dr. Razvan Gurau

ABSTRACT

The field of quantum technologies has experienced explosive growth in the past decade. As hardware capabilities have advanced, quantum computing has exhibited a scaling trend analogous to Moore's law, often referred to as Rose's law, where the number of coherently controlled qubits roughly doubles every two years. Associated with such rapid growth is an equally growing need for benchmarking tools and error handling. While advancements in quantum hardware and software have mainly concentrated on enhancing gate fidelity and qubit connectivity, the progress in addressing readout errors has been relatively limited during the same period. This thesis tries to close this gap by advancing and developing new tools for information extraction in quantum systems. We do this in two different ways:

In Part II, we focus on mitigating the effect of general readout errors in quantum systems. We developed a comprehensive protocol for readout-error-mitigated state tomography (REMST) and implemented it on a superconducting qubit system. The REMST protocol was designed to mitigate not only classical readout errors but also more challenging coherent and correlated readout errors. The protocol uses only a deterministic application of single-qubit readout, making it one of the most experimentally friendly approaches available. We developed a scalable version of the REMST protocol and showed numerically that it is suitable for systems with more than 100 qubits, matching the size of current state-of-the-art systems. The REMST protocol was applied to superconducting qubit experiments, where we saw an improvement in reconstruction accuracy by a factor of up to 30.

In Part III, we move our focus to optimal measurement strategies. Adaptive measurement strategies promise a quadratic improvement in the statistical fluctuation of pure and close-to-pure state reconstruction by dynamically optimizing measurement settings during sampling. We investigated whether these adaptive measurement strategies were viable in realistic experimental settings. We demonstrated both analytically and in numerical simulations that going from ideal readout to noisy readout disproportionally deteriorates the quality of the adaptive measurement strategy, to a point where no asymptotic benefit exists for adaptive measurements.

These results not only deepen our understanding of readout processes in superconducting qubits but also provide concrete implementations of more robust quantum measurement and tomography protocols, offering clearer insight into the interplay between readout noise and information extraction.

ZUSAMMENFASSUNG

Das Feld der Quantentechnologien hat im letzten Jahrzehnt eine Phase von explosionsartigem Wachstum durchlaufen. Mit dem Fortschreiten der Hardwarefähigkeiten folgten Quantencomputer einem Skalierungsgesetz analog zum Mooreschen Gesetz, oft auch Rosesches Gesetz genannt, laut dem sich die Anzahl der kohärent kontrollierten Qubits ungefähr alle zwei Jahre verdoppelt. Ein solch schneller Anstieg geht einher mit einem ebenfalls ansteigenden Bedarf an Benchmark-Werkzeugen und Fehlerkorrektur-Fähigkeiten. Während sich die Software- und Hardwarefortschritte hauptsächlich auf die Verbesserung von Gatter-Fidelitäten und Qubit-Konnektivität konzentriert haben, waren die Entwicklungen bei der Behandlung von Auslesefehlern im gleichen Zeitraum vergleichsweise eingeschränkt. Diese Doktorarbeit versucht diese Lücke zu schließen, indem Werkzeuge zur Extraktion von Informationen aus Quantensystemen neu- und weiterentwickelt werden. Wir wählen hierfür zwei verschiedene Ansätze:

In Part II fokussieren wir uns auf die Abmilderung des Effekts allgemeiner Auslesefehler in Quantensystemen. Wir haben ein umfangreiches Protokoll für Auslesefehler-abgemilderte Zustandstomografie (REMST) entwickelt und auf ein System von supraleitenden Qubits angewandt. Das REMST-Protokoll wurde entwickelt, um nicht nur klassische Auslesefehler, sondern auch stärker herausfordernde kohärente und korrelierte Auslesefehler abmildern zu können. Das Protokoll basiert auf einer rein deterministischen Anwendung von Einzelqubit-Messungen, was es zu einem der anwendungsfreundlichsten verfügbaren Ansätze macht. Wir haben auch eine skalierbare Version des REMST-Protokolls entwickelt und numerisch gezeigt, dass es geeignet ist für Systeme mit mehr als 100 Qubits, was Systemgrößen auf dem neuesten Stand der Technik entspricht. Das Protokoll wurde auch bei einem Experiment mit supraleitenden Qubits angewandt, wodurch wir eine Verbesserung der Rekonstruktions-Genauigkeit um einen Faktor von bis zu 30 erreichen konnten.

In Part III verlagern wir unseren Fokus auf optimale Messstrategien. Adaptive Messstrategien versprechen eine quadratische Reduktion der statistischen Fluktuationen bei der Rekonstruktion von reinen und beinahe-reinen Quantenzuständen durch die dynamische Optimierung der Messeinstellungen während des Messprozesses. Wir haben untersucht, ob diese adaptiven Strategien unter realistischen experimentellen Bedingungen umsetzbar sind. Wir konnten sowohl analytisch als auch anhand von numerischen Simulationen zeigen, dass der Übergang von idealen zu fehlerbehafteten Messungen die Qualität der adaptiven Messstrategie unverhältnismäßig stark abnehmen lässt, sodass kein asymptotischer Vorteil für adaptive Messmethoden übrig bleibt.

Diese Ergebnisse sorgen nicht nur für ein tieferes Verständnis des Messprozesses bei supraleitenden Qubits, sondern liefern auch eine konkrete Implementierung eines robusteren Quantenmessungs- und Tomografie-Protokolls, welches klarere Einsicht in das Zwischenspiel von Auslesefehlern und Informationsextraktion bietet.

CONTENTS

PUBLICATIONS

This thesis is based on the following manuscripts and publications:

- [A] **A. S. Aasen***, A. Di Giovanni*, H. Rotzinger, A. V. Ustinov, and M. Gärttner, 'Readout error mitigated quantum state tomography tested on superconducting qubits,' Communications Physics **7**, 301 (2024).
- [B] **A. S. Aasen**, A. Di Giovanni, H. Rotzinger, A. V. Ustinov, and M. Gärttner, 'Mitigation of correlated readout errors without randomized measurements,' Physical Review A **112**, 022601 (2025).
- [C] A. Di Giovanni, **A. S. Aasen**, J. Lisenfeld, M. Gärttner, H. Rotzinger, and A. V. Ustinov, 'Benchmarking the quality of multiplexed qubit readout beyond assignment fidelity,' Physical Review Applied **24**, 044043 (2025).
- [D] **A. S. Aasen** and M. Gärttner, 'Adaptive quantum state tomography in the presence of readout noise,' In preparation (2025).

The author also contributed to the following publications:

- [E] **A. S. Aasen**, S. Floerchinger, G. Giacalone, and D. Guenduez, 'Thermal fluctuations on the freeze-out surface of heavy-ion collisions and their impact on particle correlations,' Physical Review C **108**, 014904 (2023).
- [F] A. D. Giovanni, **A. S. Aasen**, M. Gärttner, H. Rotzinger, and A. V. Ustinov, 'Correlation-conscious optimization of multiplexed qubit readout,' In preparation (2025).

*These authors contributed equally.

ACKNOWLEDGMENTS

This work would not have been possible without the help of a great ensemble of characters. I have had the privilege to speak with and learn from some of the most unique and diverse set of people, academically, administratively, culturally, and socially. I have truly stood on the shoulders of giants and I want to extend my warmest *Tusen takk, Danke schön, Grazie mille,* 谢谢 *(Xièxiè)* to everyone! However, there are some people who deserve special mention.

First and foremost, I would like to express my greatest gratitude to my supervisor Martin. Being a good supervisor is challenging, and being a great person even more. With confidence, I can say that you managed to be both. You have given me the freedom to go my own way while sharing your intuition and guidance with me.

I also want to extend my thanks to my co-supervisors, Thomas and Manfred, the members of my committee, Razvan, Jörg and Lauriane, and finally the team at HGSFP.

I want to thank Andras, my closest collaborator, professional screwdriver finder and operator, experimentalist, former theorist, current Riggeti glazer, and friend, for everything.

I want to thank the MBQD (Adrian, Moritz, Oliver, Tobias, Niklas, Javad, Max, Etienne, Max, Mirco, Mohammed, Johannes, Sebastian, Luca, Tim, Leo) and $\langle \mathcal{Q}_i Q_s \rangle$ (Elena, Mauro, Martina, Maike, Ege, Siavash, Ben, Susanna, Robin, Alexander, Verena, Vlad) team for making my time in Heidelberg and Jena feel like it was spent with family. From the legendary bread debate to long island ice tea not containing ice tea to the Aasen-induced Italian-exclusion principle to the third (and last) pirates of the Caribbean being a cinematographic masterpiece to the ethics of fist-fighting emperor penguins, it has all been wonderful and I would not have traded away any of it.

Much appreciation and thanks to Elena and Mauro for taking the time to give feedback on parts of this thesis.

I want to thank Davide for showing me that olive oil, raw fish, and al dente pasta are all you need to be happy. See you in France!

Ich möchte Niklas, meinem langjährigen Kollegen und Freund, danken. Gemeinsam haben wir die wichtigsten Erinnerungen meiner Promotion geschaffen, an die ich noch viele Jahre zurückdenken werde.

我还要感谢我的伴侣一丹，如果没有你的爱和鼓励，这将会更加艰难。

Jeg vil til slutt, men ikke minst, takke mamma og pappa, Kari og Håvard, for alt. Selv om vi har vært langt fra hverandre mer eller mindre hele doktorgraden, har dere gitt meg tryggehet, på alle mulige måter. Sikkerheten til å ta mine egene valg, og vite at disse valgene aldri var feil. Jeg er glad i dere ♡.

"*Fuck you all and see you on Monday.*"

ACRONYMS

BME	Bayesian Mean Estimation
DoF	Degree of Freedom
FIM	Fisher Information Matrix
IC	Informationally Complete
IMLE	Iterative Maximum Likelihood Estimation
LI	Linear Inversion
MLE	Maximum Likelihood Estimation
MSE	Mean Squared Error
MUB	Mutually Unbiased Bases
POVM	Positive Operator-Valued Measure
PVM	Projective-Valued Measure
QDT	Quantum Detector Tomograph
QEC	Quantum Error Correction
QEM	Quantum Error Mitigation
QST	Quantum State Tomography
RDM	Reduced Density Matrix
REMST	Readout-Error-Mitigated State Tomography
SIC	Symmetric Informationally Complete
SPAM	State Preparation And Measurement

Part I

Introduction

1

MOTIVATION

> "*'Begin at the beginning,' the King said, very gravely, 'and go on till you come to the end: then stop.'*"
>
> – Lewis Carroll, Alice in Wonderland

The ultimate goal of a physical theory is to be an effective description of the natural world. Such a description must not only account for observed phenomena, but also predict new phenomena, verifiable through experiments. Few frameworks have achieved this as successfully as the theory of quantum mechanics. UNESCO has dedicated 2025 to the 100th anniversary of quantum mechanics, but quantum mechanics started already in the last days of the 19th century as a patchwork solution to the "Ultraviolet Catastrophe". The classical theory of radiation formulated by the Rayleigh-Jeans law predicts that as the wavelength of the radiation of a black body goes towards zero, the total energy emitted diverges. This not only differed sharply from the spectral data seen in stars but also went against common sense. Max Planck postulated a simple fix to the issue [1], namely to assume that light could only be emitted in discrete chunks, called *quanta*, defined through the wave frequency ν and an unknown parameter h, later to be known as the Planck constant,

$$E = h\nu. \tag{1.1}$$

Today this constant has become synonymous with "quantumness" to a point where to recover classical physics, the limit $h \to 0$ is taken.

Although the physicists at the time were skeptical of Planck's ad hoc solution to the problem, similar ideas were quickly employed to quantize other aspects of contemporary physics, such as the orbital angular momentum of the electrons around atoms, to great effect. This marked the beginning of a dynamic era of rapid theoretical and experimental advancements, beginning with Einstein's photoelectric effect in 1905 [2], followed by the Bohr model of the atom in 1913 [3], the Stern-Gerlach experiment in 1922 demonstrating the quantization of the electron spin [4], and Arthur Compton's theory of particle scattering in 1923 [5]. These contributions culminated in what is considered the establishment of quantum mechanics in 1925, marked by the development of wave [6] and matrix mechanics [7]. This mathematical framework, pioneered by leading figures such as Erwin Schrödinger, Paul Dirac, Wolfgang Pauli, Lucy Mensing, and Werner Heisenberg, continues to define our understanding of the quantum world today.

Fast forward to the 1980s, and the presence of quantum mechanics has been firmly and undisputedly felt by the experimental demonstration of two of the strangest properties of quantum physics, *entanglement* [8–10] and *wave-particle duality* [11].

> "*Since this is so bizarre, since this is so revolutionary, can't something useful be done with it?*"
>
> — Richard Feynman [12]

These discoveries prompted new questions: Could this quantum strangeness be harnessed for practical use, much like how semiconductors[1] revolutionized computing and the discovery of the double helix transformed genomics decades earlier? One of the first practical technologies grounded in the principles of quantum mechanics, the atomic clock, had already been developed and is based on the discrete transition frequencies between different energy levels of the cesium atom. The idea of quantum simulators, using quantum systems to simulate other systems, both classical and quantum, was put forward by Yuri Manin and Paul Benioff in 1980 [13, 14] and later famously by Richard Feynman in 1982 [15] and David Deutsch in 1985 [16].

> "*...nature isn't classical, dammit, and if you want to make a simulation of nature, you'd better make it quantum mechanical, and by golly it's a wonderful problem, because it doesn't look so easy.*"
>
> — Richard Feynman [15]

It had already been understood that it would never be possible to fully simulate quantum systems, and therefore nature, with only classical computers, which at that point had just become commercially available.

Jumping to the present, two different approaches to harness "quantumness" to do our bidding emerged, general-purpose quantum computers and special-purpose quantum simulators. Constructed in distinctly different ways, quantum computers are programmable systems capable of executing any sequence of quantum operations, analogous to classical computers, while quantum simulators are highly controllable physical systems that exploit their natural quantum dynamics to emulate other complex systems. Quantum computers use quantum algorithms that take advantage of the vast possibilities of quantum superposition and entanglement to perform computation, whereas quantum simulators exploit our ability to precisely manipulate physical systems, such as trapped ions confined in electromagnetic fields [17], ultracold atoms in optical lattices [18], and arrays of Rydberg atoms [19–21], to emulate systems that are otherwise difficult to control or probe, such as how amino acid chains are folded to become proteins and discover the ground-state energies of chemicals [22, 23].

Both quantum simulators and quantum computers are still in their early stages of technological advancement, plagued by problems quite unique to quantum platforms. Unlike contemporary classical computers, where each physical bit can be made up of more than 50 000 electrons, a quantum bit (qubit) can be a single electron[2]. This means that quantum systems are uniquely prone to being disturbed by external factors. On top of this, parts of the information in quantum systems are stored in what are known as coherences, which can be completely destroyed even by a single other electron. Together, both of these aspects create an extraordinarily challenging task to be able to operate these quantum systems.

---

[1] It can be argued that the transistor was the first quantum technology, due to the use of semiconductors, which can only be properly described by quantum mechanics [12].

[2] There also exist macroscopic quantum systems such as quantum dots that are also made up of thousands of electrons [24, 25]. Nevertheless, they are plagued by many of the same problems as single electron qubits.

Perhaps the most profound departure from classical computing lies in the way information is encoded, processed, stored, and extracted. Measurements of quantum systems introduce challenges unlike anything encountered in the classical world. If you perform a series of computations on a classical computer, the result of the final measurement is always the same, the bit is either a 0 or a 1. If the computation is repeated, the answer does not change. In quantum mechanics, this is not the case. A qubit can exist in a state known as a *superposition*, where it is in a combination of 0 and 1 at the same time. If you measure the qubit, the outcome will be either 0 or 1, just like a classical bit, but if you were to do the same computation and measure it again, you could get a different outcome than before. In classical computing, the information is encoded directly in the binary outcome, 0 or 1. In contrast, a qubit encodes information in the probabilities of these outcomes, and in the coherences between them, which are inaccessible from a single measurement. On top of this, the measurement process itself is destructive for quantum systems, and once a qubit is measured, the original quantum state vanishes. We are therefore only able to extract a fraction of the information in a qubit per measurement we perform. Compounding the difficulties, we need to transport fragile quantum information to classical systems, typically computers, with a quantum-classical coupling. This means that, following our previous example, we need to have a probe, usually made up of hundreds if not thousands of electrons, to interact with a single electron in a very precise way to not destroy the little information we are able to get. Needless to say, this can create even more problems for the already delicate system.

> "*Quantum physics is extraordinary. After a century of effort, we know of no scenario in which the theory appears to have reached its limits.*"
> — Alain Aspect [12]

We are currently in the era of noisy intermediate-scale quantum computation [26], characterized by quantum devices with more than 100 noisy qubits. While progress in controlling and scaling these systems has been remarkable [27–30], we remain in a phase where maintaining qubit coherence long enough to extract useful computational results is challenging [28, 31–33]. An integral part of these efforts is *quantum error correction*, where quantum information is redundantly encoded in many physical qubits, and a sophisticated error correction protocol dynamically removes physical errors acquired during computation. Although the long-term goal is fault-tolerant quantum computation via fully error-corrected qubits, hardware demands remain beyond current capabilities [34]. In the meantime, statistical tools and methods are being developed to reduce the impact of noise on current generations of quantum devices. These methods, dubbed *error mitigation*, aim to extract clean quantum information from noisy quantum devices [33, 35]. Among the various sources of error in quantum systems, readout errors, which arise during the final measurement of the qubits, become increasingly dominant as hardware operations continue to improve. Despite their growing impact, existing readout-error-mitigation techniques are often narrow in scope and few are capable of addressing the full spectrum of readout errors, including correlated and coherent errors found in real devices [A, 36–39]. This thesis aims to address this gap through two primary contributions:

1. Development and experimental validation of a practical readout-error-mitigation protocol capable of mitigating a broad class of quantum measurement errors, including coherent and correlated readout errors between arbitrary qubits, using only single-qubit measurements.

2. Investigation of how readout errors influence the performance of optimal adaptive measurement strategies, with implications for the viability of adaptive strategies for state reconstruction tasks.

Together, these investigations aim to deepen our understanding of how noise and readout imperfections affect quantum information processing and to offer practical tools for mitigating such effects in real-world quantum devices.

The thesis is structured in the following way: In Chapter 2, we introduce the foundational theoretical concepts of quantum mechanics necessary to describe both noisy quantum systems and noisy measurement processes. Additionally, this chapter presents the statistical tools used to assess the quality of the information extracted through measurement. The results of this thesis can be separated into two main parts. In Part II, we concentrate on readout errors in quantum systems. We begin in Chapter 3 with a brief overview of the existing approaches to error management in quantum systems. Building on this, we develop an algorithm for mitigating the effects of readout errors in quantum state reconstruction tasks. This method is experimentally validated on a superconducting qubit in Chapter 4. The techniques are then extended in Chapter 5 to handle correlated readout errors, making them applicable to quantum systems with more than 100 qubits. Using these tools, we propose an alternative readout quality metric for superconducting qubits in Chapter 6.

In Part III, our focus shifts towards adaptive measurement strategies. These strategies have a strong potential to improve the quality of the extracted information by dynamically selecting future measurements. However, it is crucial to understand their performance in the presence of noise. In Chapter 7, we start by reviewing the field of optimal measurement design and demonstrate that adaptive strategies are essential to achieve a practical advantage in state reconstruction. We contribute to this field by showing that introducing readout noise disproportionally deteriorates the potential advantage of adaptive strategies. Finally, in Chapter 8, we summarize and reflect on the main contributions of the thesis.

2

INTRODUCTION TO QUANTUM THEORY

> "*Anyone not shocked by quantum theory has not understood it.*"
>
> – Niels Bohr [40]

Quantum physics is the foundational framework used to describe systems on a microscopic scale. Much of the unintuitive quantum strangeness becomes manageable with the appropriate mathematical tools. This chapter is not intended as a full review of quantum theory, but rather as a compact toolbox focused on the concepts most relevant to quantum information extraction and estimation. There exist many widely available comprehensive introductions to quantum information and computation online [40–42].

We begin by defining quantum states (Section 2.1), how they evolve (Section 2.2) and how to extract information from them (Section 2.3 and Section 2.4), followed by a treatment of open quantum systems and noise modeling (Section 2.5 and Section 2.6). The chapter concludes with estimation theory (Section 2.7 and Section 2.8), quality metrics relevant to state tomography (Section 2.9) and a brief introduction to the main quantum architecture used in this thesis, the superconducting qubit (Section 2.10). This review takes inspiration from many well-established works, such as Nielsen and Chuang [40] covering quantum technology, Sakurai and Napolitano [43] for a formal mathematical treatment, Jordan and Siddiq [44] for a comprehensive treatment of quantum measurements, and seminal work by Blume-Kohout [45] on robust tomography of quantum states.

2.1 QUANTUM STATES AND QUBITS

The core object in quantum theory is the quantum state, mathematically described by a vector $|\psi\rangle$ in a complex Hilbert space $\mathcal{H}$. The mathematical framework does not stem from any intuitive progression from classical physics, but rather from a necessity to describe the phenomena we see. The number of parameters needed to specify the system's state, known as its *degrees of freedom* (DoF), not only increases exponentially with the system size but also reflects its ability to store information. This concept is central to the idea of a quantum advantage for quantum computers. Although there is no fundamental principle for determining what DoFs are relevant for any specific system, in practice we have devised effective methods to recognize them in relevant systems.

In Hilbert space, the simplest object is the two-level quantum system known as a *qubit*, because of its analogy to the classical two-level system, the bit. A qubit is described by the state vector

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle , \tag{2.1}$$

where $|0\rangle = (1,0)^T$ and $|1\rangle = (0,1)^T$ are the *computational basis states*. The complex basis coefficients α and β are bound by the normalization condition on state vectors, which is related to conservation of probabilities,

$$|\alpha|^2 + |\beta|^2 = 1. \tag{2.2}$$

The largest departure from classical bits is that the qubit does not have to be in a binary state, $|0\rangle$ or $|1\rangle$, but can be any linear combination of the two, known as a *superposition*.

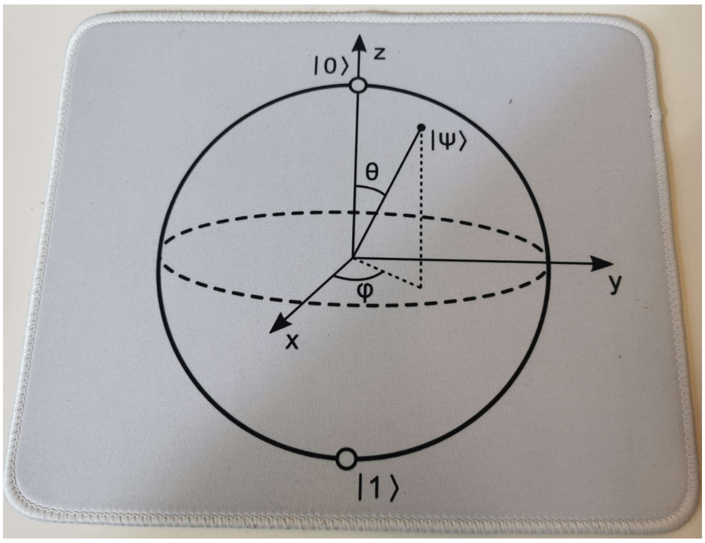


Figure 2.1: A picture of my office mouse pad, and also a diagram of a Bloch sphere.

An intuitive geometric visualization tool for a single qubit is the *Bloch sphere*. The state vector originally possesses four degrees of freedom, one is eliminated because of normalization (Equation 2.2) and another due to a global unitary gauge freedom, leaving only two. These two remaining degrees of freedom are then mapped onto the surface of a 2-sphere, providing us with the Bloch representation[1]

$$|\psi\rangle = \cos(\theta/2)\,|0\rangle + e^{i\phi}\sin(\theta/2)\,|1\rangle\,, \tag{2.3}$$

where the angles θ and ϕ correspond to the spherical coordinates. Thus, every state vector can be associated with a specific point on the 2-sphere, as shown in Figure 2.1.

A particularly important family of qubit states is the eigenstates of the Pauli matrices, visualized on the Bloch sphere in Figure 2.2. The Pauli matrices are defined as

$$\sigma_x = X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \text{and} \quad \sigma_z = Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{2.4}$$

---

[1] The reason this is possible is because the quantum state belongs to the symmetry group $SU(2)$, which is isomorphic to the symmetry group $SO(3)$, which is the symmetry group of the 2-sphere.

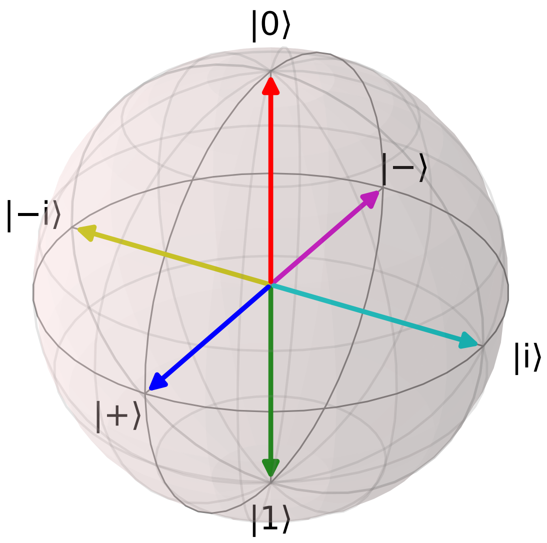


Figure 2.2: The eigenstates of the Pauli matrices, Equation 2.5, plotted on the Bloch sphere.

where each matrix has two eigenstates,

$$\begin{gathered}|+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, \qquad |-\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}, \qquad |+i\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix} \qquad |-i\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-i\end{pmatrix}, \\ |0\rangle = \begin{pmatrix}1\\0\end{pmatrix} \quad \text{and} \quad |1\rangle = \begin{pmatrix}0\\1\end{pmatrix},\end{gathered} \tag{2.5}$$

for σ_x, σ_y and σ_z respectively. In the literature, you may find other notations for the eigenstate, such as $\{|0_i\rangle, |1_i\rangle\}$ for $i \in \{x, y, z\}$. In the Bloch sphere representation, these eigenstates form a unit cube, with each vector pointing towards the center of a face of the cube.

2.1.1 *Composite systems*

Using the single qubit as a fundamental building block, the extension to multi-qubit systems is straightforward. A two-qubit system can be written as

$$|\psi\rangle = \psi_{00}|00\rangle + \psi_{01}|01\rangle + \psi_{10}|10\rangle + \psi_{11}|11\rangle, \tag{2.6}$$

where $|i\rangle \otimes |j\rangle = |ij\rangle$ with $i, j \in \{0, 1\}$. The number of complex coefficients has doubled from the single-qubit case. This pattern of exponential growth of DoFs continues in the n-qubit states, which are described by 2^n complex numbers. If the global phase and normalization conditions are included, a n-qubit state has a total of $2(2^n - 1)$ real degrees of freedom, exponentially more than the individual subsystems. What is the source of this added complexity? It appears that as larger systems are created, the corresponding Hilbert space expands exponentially. To gain a clearer understanding, it is helpful to classify the states in this composite Hilbert space as either *separable* or *entangled*.

We say a system is separable if it is possible to write a quantum state purely in terms of the tensor product between two subsystems A and B, $|\psi\rangle = |\theta_\mathrm{A}\rangle \otimes |\phi_\mathrm{B}\rangle$. If we consider a separable two-qubit system and compare it to the example in Equation 2.6,

$$\begin{aligned} |\psi\rangle &= |\theta_\mathrm{A}\rangle \otimes |\phi_\mathrm{B}\rangle = (\theta_0 |0\rangle + \theta_1 |1\rangle) \otimes (\phi_0 |0\rangle + \phi_1 |1\rangle) \\ &\overset{\text{Eq. (2.6)}}{=} \psi_{00} |00\rangle + \psi_{01} |01\rangle + \psi_{10} |10\rangle + \psi_{11} |11\rangle \,. \end{aligned} \tag{2.7}$$

The DoFs are drastically reduced. Since $\psi_{ij} = \theta_i \phi_j$ for $i, j \in 0, 1$, each coefficient is normalized independently and remains unaffected by a global phase change on its subsystem. Consequently, the combined DoFs for this system are four real parameters, reflecting the straightforward addition of the DoFs of both subsystems. If it is not possible to express a system as a tensor product of two subsystems A and B, the system is said to be inseparable or bipartite entangled between A and B.

The additional DoFs available in entangled states enable genuinely non-classical phenomena[2]. One well-known example is superdense coding, where more than one bit of classical information is transmitted using a single qubit [48]. Another is quantum key distribution, which guaranties the detection of eavesdroppers through the fundamental disturbance caused by measurement [49, 50]. At the same time, the exponential scaling of the DoFs makes quantum systems extremely difficult to simulate classically. With the structure of quantum states now established, we turn to how these systems evolve over time.

2.2 QUANTUM EVOLUTION

An isolated quantum system evolves according to the Schrödinger equation,

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle \,, \tag{2.8}$$

where $\hbar = \frac{h}{2\pi}$ is the reduced Planck constant and $\hat{H}$ is the *Hamiltonian* of the system. The hat indicates that H is an operator acting on the Hilbert space. For simplicity, we will drop this notation going forward.

If the Hamiltonian is time-independent, the solution to Equation 2.8 is

$$|\psi(t)\rangle = e^{-iHt/\hbar} |\psi(0)\rangle = U(t) |\psi(0)\rangle \,, \tag{2.9}$$

where the unitary matrix $U(t) = e^{-iHt/\hbar}$ fully describes the evolution of the system over time t. Thus, time evolution in quantum mechanics corresponds to applying a unitary matrix transformation to the initial state.

2.2.1 *Quantum gates*

In quantum computing, state evolution is discretized into a series of unitary operators, called *gates*. Unitary gates are the quantum analogs to classical logical gates. Any finite n-qubit unitary

---

[2] The precise meaning of “non-classical” remains nuanced. Typically, it refers to correlations between observables that cannot be accounted for by classical probability theory, such as those that violate Bell inequalities. Entanglement is a requirement for general non-classicality, but is not sufficient. Recently, the notion of “quantum magic” has been proposed as an additional resource for quantumness in the context of computational advantage [46, 47].

evolution can be approximated efficiently by a small set of one- and two-qubit gates, called a universal gate set [40, 51–53][3]. One common example of a universal gate set is the single-qubit gates

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}, \tag{2.10}$$

and the two-qubit gate

$$\text{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{2.11}$$

The choice of gate sets is influenced by the operation that is easy to execute on different quantum platforms. For example, the Egal chip for IBM quantum uses the gate set RZ, SX, X and the echoed cross resonance (ECR) gate[4] [54]. Research is also being conducted on the incorporation of higher-qubit gates into the universal gate set to achieve a better overall approximation accuracy [55–57].

2.3 EXPECTATION VALUES OF OBSERVABLES

In quantum mechanics, any measurable quantity can be represented by a Hermitian operator O, sometimes called an *observable*. The expected outcome of measuring O on a quantum state $|\psi\rangle$ is given by the *expectation value*,

$$\langle O \rangle = \langle \psi | O | \psi \rangle . \tag{2.12}$$

Since O is a Hermitian operator, the expectation value will always be a real number. Individual measurements of the operator do not yield the expectation value itself, but one of its eigenvalues. As an example, consider the measurement of the spin of a single qubit, given by the Z operator. It has two eigenvalues, ± 1, for the eigenstate $|0\rangle$ and $|1\rangle$, respectively. If we measure the $|+\rangle$ state, the expectation value, using Equation 2.12, is

$$\langle Z \rangle = \langle + | Z | + \rangle = \frac{1}{2}\left(\langle 0 | Z | 0 \rangle + \langle 1 | Z | 1 \rangle \right) = 0. \tag{2.13}$$

Since each individual outcome is limited to either ±1, the probabilities of receiving either outcome must be equal. However, if we measure the X operator instead, there would be a 100% probability of measuring the outcome +1. How are the probabilities of the different outcomes determined?

---

[3] The Solovay-Kitaev theorem [40, 51] and it's generalization [52], guaranties that an n-qubit unitary can be approximated efficiently, requiring only a polylogarithmic number of gates for a given target accuracy.

[4]

$$RZ(\theta) = \begin{pmatrix} e^{i\theta/2} & 0 \\ 0 & e^{-i\theta/2} \end{pmatrix}, \quad SX = \frac{1}{2}\begin{pmatrix} 1+i & 1-i \\ 1-i & 1+i \end{pmatrix}, \quad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad ECR = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 0 & 1 & i \\ 0 & 0 & i & 1 \\ 1 & -i & 0 & 0 \\ -i & 1 & 0 & 0 \end{pmatrix}.$$

Using the spectral decomposition of the observable, one can systematically derive the probabilities of receiving the different outcomes,

$$O = \sum_i o_i P_i, \tag{2.14}$$

where $P_i = |o_i\rangle \langle o_i|$ are the *projectors* onto the eigenstate of the observable, $O |o_i\rangle = o_i |o_i\rangle$. Substituting Equation 2.14 into Equation 2.12, we obtain

$$\langle O \rangle = \langle \psi | O | \psi \rangle = \sum_i o_i \langle \psi | P_i | \psi \rangle = \sum_i p_i o_i, \tag{2.15}$$

where in the last equality, the definition of an expectation value was inserted. This leads to the interpretation that when measuring the operator O on the quantum state $|\psi\rangle$, the result o_i is obtained with the probability $p_i = \langle \psi | P_i | \psi \rangle$. This indicates that the projectors are intimately connected with the measurement process.

It is important to emphasize that the probabilities arising from projecting a quantum state onto an eigenbasis are fundamentally different from those encountered in statistical physics. In statistical physics, probabilities reflect our inability to track all relevant degrees of freedom in a system. The statistical description arises from our ignorance about the exact state of the system. If we had complete knowledge of the microstate of the system, we could, in principle, determine the precise value of any observable rather than relying on ensemble averages. In contrast, the probabilistic nature of quantum mechanical systems is intrinsic. The quantum state is already fully specified, and measurement outcomes occur with probabilities determined by the state's overlap with the eigenstates of the observable[5]. Building on our revised understanding that expectation values of projectors represent probabilities, we will now examine their relationship to quantum measurements.

2.4 IDEAL MEASUREMENTS AND POST-MEASUREMENT STATES

> " *... measurement is no longer simply "looking" ; what you choose to learn changes the system being measured ...*"
>
> – Andrew Jordan and Irfan Siddiqi [44]

We saw in Section 2.3 that projectors seem to describe probabilities quite naturally. We will show that, in fact, projectors describe a type of ideal measurement in quantum mechanics. The mathematical formulation of ideal measurements, known as von Neumann measurements [58], is made using projection-valued measures (PVM). A PVM is a set of Hermitian operators $\mathcal{P} = \{P_i\}$, where each element P_i has rank 1 and satisfies $P_i P_j = \delta_{ij} P_j$, where δ_{ij} is the Kronecker-delta.

To motivate the use of this construction, consider the measurement process as an abstract function that probabilistically maps a quantum $|\psi\rangle$ to a classical outcome, visualized in Figure 2.3. The core mechanism is to transform a quantum state into a set of probabilities. We already know that a projector can have the interpretation of being probabilities, so we can start there. Like probabilities, the expectation value of a projector is bounded between $0 \leq \langle P \rangle \leq 1$.

---

[5] This is sometimes referred to as the collapse of the wave function. What this collapse actually is, or the mechanism behind it, is still a debated ontological problem.

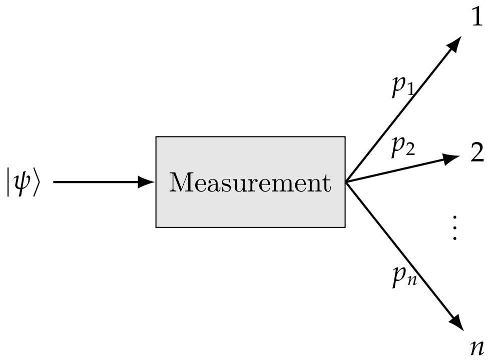


Figure 2.3: An abstract sketch of a quantum measurement. A quantum state $|\psi\rangle$ is measured, yielding classical outcomes $1, 2, \dots, n$ with respective probabilities $p_1, p_2, \dots, p_n$.

This is due to the normalization of the state vector, see Equation 2.2. The maximal expectation value occurs if $|\psi\rangle$ is the same state as the eigenstate of P with a non-zero eigenvalue, up to a global phase. The minimal value occurs if $|\psi\rangle$ is a linear combination of the other orthogonal eigenstates. Another important property of a probability distribution is that it must be normalized, expressed as $\sum_i p_i = 1$. Thus, for a set of projectors $P_i \in \mathcal{P}$, it is necessary that the resulting probabilities for any state $|\psi\rangle$ are also normalized. This leads to the following requirement for the set of projectors,

$$\sum_i p_i = \sum_i \langle\psi| P_i |\psi\rangle = \langle\psi| \sum_i P_i |\psi\rangle \overset{!}{=} 1. \tag{2.16}$$

The only operator that satisfies the last equality for all states $|\psi\rangle$ is

$$\sum_i P_i = \mathbb{1}. \tag{2.17}$$

Deferring the details to Appendix A, the set of projectors satisfying Equation 2.17 must be orthonormal, $P_i P_j = P_i \delta_{ij}$, and the number of elements in $\mathcal{P}$ is exactly $\dim(\mathcal{H})$. We observe that PVMs possess all the properties we expect from a measurement. The outstanding question is how they exemplify an ideal measurement. To address this, we need to examine the post-measurement state.

The concept of a post-measurement state in classical systems is simple. When a property of the system is determined through measurement, it is evident that the system had this property both before and after the measurement, assuming that no substantial change occurred. However, in quantum systems, the nature of the post-measurement state is more ambiguous, and as we shall see, it cannot always be determined. Quantum theory as presented so far gives no way for us to determine the post-measurement state. Sometimes, the post-measurement state is included in the postulates of quantum mechanics [40]. In experiments, it was observed that if the same measurement was repeated multiple times after an initial measurement was performed, the same outcome occurred in all subsequent measurements [4, 59, 60]. For this behavior to be consistent with the discussion previously established, the post-measurement state must have been changed to the eigenstate of the observable that was observed. That is, if the result i occurred, the post-measurement state is $|o_i\rangle$. Alternatively, the resulting eigenstates can be found using the Lünders rule [61, 62].

$$|o_i\rangle = \frac{P_i |\psi\rangle}{\sqrt{p_i}}. \tag{2.18}$$

It is notable that the measurement process itself changes the state. This should not be confused with the idea that a measurement is poorly performed, and if we were better at this, we could leave the system undisturbed. The fact that a system was measured changed it. This then leads to the unsettling answer to the quantum variant of the thought experiment "If a tree falls in a forest and no one is around to hear it, does it make a sound?", which is a resounding "no"[6].

> "*Unperformed measurements have no results.*"
>
> – Asher Peres [63]

An ideal quantum measurement is characterized by a well-defined post-measurement state, ensuring that repeated measurements in the same basis yield consistent outcomes. Consequently, all information about the original quantum state is removed and extracted through the measurement process. Currently we do not have the tools necessary to discuss what partial information extraction looks like, but PVMs will turn out to be a special case where we have complete distinguishability between outcomes, and are therefore considered ideal.

The most important example of an ideal measurement is the n-qubit *computational basis* measurement, $\mathcal{C}^n$. This type of measurement is fundamental in quantum computing and is formed using projectors onto the computational basis states described in Section 2.1,

$$P_0 = |0\rangle\langle 0| \quad \text{and} \quad P_1 = |1\rangle\langle 1|, \qquad \mathcal{C}^1 = \{P_0, P_1\}. \tag{2.19}$$

The computational basis measurement is equivalent to the ideal measurement of the Z operator in Equation 2.4. Creating PVMs for composite systems is completely analogous to creating the basis states for composite systems of qubits. For n qubits, the computational basis measurement is defined

$$\mathcal{C}^n = \{|b_i\rangle\langle b_i| \;|\; i \in \{0, 1, \ldots, 2^n - 1\}\}, \tag{2.20}$$

where b_i is the binary expansion of i. We have introduced all the tools necessary to describe the full life-cycle of an ideal quantum state, from preparation to evolution to measurements. We will now shift our focus to systems that are not ideal, open quantum systems.

2.5 OPEN QUANTUM SYSTEMS

> "*... not only does God play dice but ... he sometimes confuses us by throwing them where they cannot be seen.*"
>
> – Stephen Hawking [64]

Although unitary evolution and the state vector are fundamental to quantum computation, they represent an ideal scenario in which no information is lost to an environment. However, in reality, systems are never perfectly isolated from their surroundings, leading to noise, decoherence, and information loss. To account for these environmental interactions, it is necessary to adopt a framework that includes unknown degrees of freedom and describes open system dynamics.

---

[6] This is essentially the Schrödinger's cat thought experiment.

2.5.1 *The density matrix*

To generalize the concept of quantum states, we introduce the density matrix which stems from the need to represent systems where the actual state is uncertain. To motivate the transition from state vectors to matrices, consider a basic example in which a system outputs $|0\rangle$ and $|1\rangle$ with equal probability. To create such a system, perform an ideal computational basis measurement on the superposition state $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, but do not record which outcome was obtained. Consequently, the resulting system can be described with the statistical mixture

$$\text{System} = \begin{cases} |0\rangle & \text{with } p_0 = 0.5, \\ |1\rangle & \text{with } p_1 = 0.5. \end{cases} \tag{2.21}$$

To see that the system in Equation 2.21 is not in the superposition state $|+\rangle$, consider measuring the observable X. Since $|+\rangle$ is the $+1$ eigenstate of X, a X measurement on $|+\rangle$ will always produce the outcome $+1$. In contrast, the system in Equation 2.21 produces $|0\rangle$ and $|1\rangle$ with equal probability. The expectation value of an X measurement on this mixture is therefore

$$\langle X\rangle = p_0 \langle 0| X |0\rangle + p_1 \langle 1| X |1\rangle = 0.5 \times 0 + 0.5 \times 0 = 0. \tag{2.22}$$

Because the pure state $|+\rangle$ yields $\langle X\rangle = +1$ while the mixture yields $\langle X\rangle = 0$, the two systems are clearly distinct.

If the above exercise is repeated for different measurements, it will become apparent that no state vector is capable of describing the system in Equation 2.21. Therefore, we introduce the density matrix ρ (or density operator) that is defined for an ensemble of states $|\psi_i\rangle$ that occur with probability p_i

$$\rho = \sum_i p_i |\psi_i\rangle \langle\psi_i| , \tag{2.23}$$

where $\sum_i p_i = 1$. Defining the density matrix in this way gives it the very useful properties,

$$\begin{aligned} &1. \quad \rho = \rho^\dagger \quad \text{(Hermitian)} \\ &2. \quad \mathrm{Tr}(\rho) = 1 \quad \text{(Normalization)} \\ &3. \quad \rho \geq 1 \quad \text{(Positivity)} \end{aligned} \tag{2.24}$$

The state vector formalism is recovered if the ensemble consists of only one state $|\psi\rangle$, $\rho = |\psi\rangle\langle\psi|$. This also means that Equation 2.23 can be rewritten in terms of other density matrices

$$\rho = \sum_i p_i \rho_i. \tag{2.25}$$

To distinguish the two types of density matrices, the concept of *pure* and *mixed* states is introduced. The purity of a density matrix is defined as

$$\text{Purity} = \mathrm{Tr}\big(\rho^2\big). \tag{2.26}$$

The purity tells us how many, and how large, the non-zero eigenvalues of the density matrix are. Each non-zero eigenvalue corresponds to an amount of mixing inside the ensemble. An ensemble containing only a single state will only have one eigenstate with a non-zero eigenvalue, and this eigenvalue must be 1 due to the normalization constraint. Therefore, a density matrix with purity of 1 is equivalent to having an exact quantum state. At the opposite end of the spectrum, a purity of $1/2^n$ is the maximally mixed state, where n is the number of qubits in the system. The maximally mixed state is equivalent to having a uniformly random ensemble of orthonormal quantum states. The example in Equation 2.21 is one such state, and is written $\rho = \mathbb{1}/2$.

The density matrix introduces new degrees of freedom. A matrix satisfying the properties of Equation 2.24 has $4^n - 1$ degrees of freedom. The single-qubit Bloch sphere elegantly includes this additional degree by mapping it to the radius. The mixed state Bloch sphere is defined

$$\rho = \frac{1}{2}\left(\mathbb{1} + \vec{a}\cdot\vec{\sigma}\right), \tag{2.27}$$

where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ is the Pauli vector. The Bloch vector, $\vec{a}$, can be rephrased in terms of the spherical coordinates we used before,

$$\vec{a} = \left(r\sin(\theta)\cos(\phi), r\sin(\theta)\sin(\phi), r\cos(\theta)\right). \tag{2.28}$$

The new radial degree of freedom has a powerful interpretation of being the purity of the quantum state, where it can be shown that

$$\mathrm{Tr}\left(\rho^2\right) = \frac{1}{2}\left(1 + r^2\right). \tag{2.29}$$

Therefore, mixed states inhabit the bulk volume of the sphere, while pure states are on the surface. The maximally mixed state is in the center of the sphere.

Following this generalization of the quantum state, some of the tools introduced in Section 2.2 and Section 2.3 need to change accordingly,

$$\begin{array}{llcl} \text{1. Unitary evolution} & |\psi(t)\rangle = U(t)\,|\psi(0)\rangle & \rightarrow & \rho(t) = U(t)\rho(0)U^\dagger(t) \\ \text{2. Expectation value} & \langle O\rangle = \langle\psi|\,O\,|\psi\rangle & \rightarrow & \langle O\rangle = \mathrm{Tr}(\rho O) \end{array} \tag{2.30}$$

The PVM formalism changes accordingly with the expectation values. However, we will spend more time and introduce an even more robust measurement formalism for open systems, which is discussed Section 2.6. Before that, we will introduce a new type of evolution permitted by open quantum systems.

2.5.2 *Quantum channels*

The framework of open quantum systems offers a systematic way to describe the evolution of subsystems without the need for detailed information about the whole system. To model the effects of external interactions or incomplete information, we introduce the formalism of *quantum channels*.

As illustrated in Figure 2.4, the total system, referred to as the universe, is conceptually partitioned into a target system (S) and an environment (E). These two components may interact, exchanging both energy and information. Viewed from the system's perspective, these interactions manifest themselves as decoherence and noise. According to the system dilation

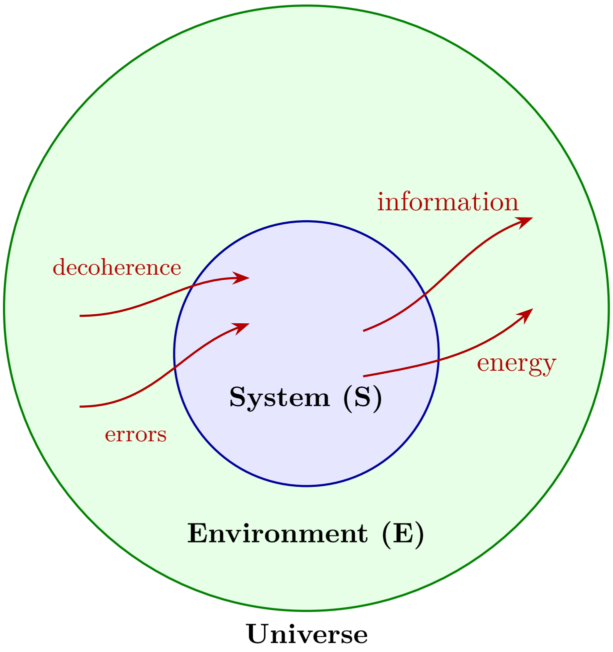


Figure 2.4: A schematic representation of an open quantum system. The total system, consisting of a system of interest (S) and an environment (E) interact by sharing energy and information between the two systems. The system experiences this interaction as decoherence and can introduce errors.

principle originally introduced by William Stinespring [65], any open system evolution can be characterized by a unitary evolution in a Hilbert space that is double the size of the original. Specifically, it states that the combined system-environment can be modeled as a pure state evolving under a unitary U, after which a partial trace can be performed over the environment DoFs. The effective evolution of the system of interest is therefore described by the channel

$$\mathcal{E}(\rho_S) = \mathrm{Tr}_E(U\rho_{SE}U^\dagger). \tag{2.31}$$

where ρ_{SE} is the joint system-environment state and $\rho_S = \mathrm{Tr}_E(\rho_{SE})$ is the reduced state of the system. The channel effectively abstracts away the environment and only captures the statistical effect of the environment on the target system. There is a rich theory surrounding the quantum channel and equivalent representations, but we will limit ourselves to the *operator-sum* representation of channels (sometimes called the Kraus representation),

$$\mathcal{E}(\rho_S) = \sum_k E_k \rho_S E_k^\dagger. \tag{2.32}$$

where the set of operators $\{E_k\}$ satisfies the following properties

1. $\sum_k E_k E_k^\dagger = \mathbb{1}$ (Trace-preserving)
2. $\mathcal{E}\left(\sum_i p_i \rho_i\right) = \sum_i p_i \mathcal{E}(\rho_i)$ (Convex-linear map)
3. For any extended Hilbert space $\mathcal{H} = \mathcal{H}_R \otimes \mathcal{H}_S$, the channel $\mathrm{id}_R \otimes \mathcal{E}_S$ is positive (Complete positivity)

(2.33)

The operator-sum representation is both mathematically convenient and has a physical interpretation for the E_k operators, where each E_k corresponds to an effective operation on the system associated with an imagined outcome k from a projective measurement on the environment.

2.5.3 *Modeling noise using channels*

One of the primary uses of quantum channels is to describe noise processes. Channels provide a flexible and powerful formalism for modeling these effects without requiring detailed knowledge of the environment. Accurate noise modeling is essential not only for understanding the system, but also for designing mitigation and estimation strategies.

One of the most widely used noise channels is the depolarizing channel, which represents a partial loss of information to the environment. It can be modeled as a probability p replacing the state with a maximally mixed state [66, 67],

$$\mathcal{E}_{\text{depol}}(\rho) = p\frac{\mathbb{1}}{2^n}\operatorname{Tr}(\rho) + (1-p)\rho, \tag{2.34}$$

where n is the number of qubits. Although this expression does not explicitly match the operator-sum form in Equation 2.33, for a single qubit it is equivalent to a channel that applies all Pauli operators X, Y, Z with equal weighting $p/3$ and the normalized identity operator $\mathbb{1}/2^n$ with weighting $1-p$ [40],

$$\mathcal{E}_{\text{depol}}(\rho) \overset{n=1}{=} \left(1 - \frac{3p}{4}\right)\rho + \frac{p}{4}\sum_{i=1}^{3} K_i \rho K_i^\dagger. \tag{2.35}$$

Further generalizations of operator-sum representation exist for multi-qubit states [67, 68].

Another important type of noise modifies the phase information in quantum systems. Phase-related errors can result from imperfect control of unitary gates, such as fluctuating pulse lengths, timing errors, or incorrectly calibrated gates. These errors can be modeled as a constant over-rotation or as a phase flip,

$$\begin{aligned} &E_\theta = e^{-\frac{i\theta}{2}X} \quad \text{(Over-rotation)},\\ &E_0 = \sqrt{p}\mathbb{1}, \quad E_1 = \sqrt{1-p}Z \quad \text{(Phase-flip)}. \end{aligned} \tag{2.36}$$

A third common channel is the amplitude damping channel, which can be used to describe energy loss due to spontaneous emission of an excited state to a ground state. The channel is modeled as a population transfer from excited states to the ground state

$$E_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix}, \qquad E_1 = \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix} \qquad \text{(Amplitude damping)}, \tag{2.37}$$

where γ is the damping rate. This model is used to describe the relaxation T_1 in superconducting qubit systems.

In practice, noise rarely falls neatly into just one of the categories mentioned above. Physical systems often experience a combination of many channels. As a result, precise modeling of noise is important not only for understanding the dynamics of the system, but also in quantum estimation tasks, where they are used to mitigate the effect of noise, and in optimal measurement strategies, where noise impacts extractable information.

2.6 GENERALIZED QUANTUM MEASUREMENTS

In Section 2.4, we introduced projective measurements as sharp and ideal quantum measurements. Projective measurements serve as a theoretically important tool, but they cannot be used to describe realistic detectors that suffer from crosstalk, noise, and distorted information extraction. To incorporate noise and imperfections caused by interactions with unknown degrees of freedom, we follow the same conceptual idea as we did for quantum channels. Noisy measurements can be modeled as ideal projective measurements on a system-environment composite, where we trace out the environment. Following this procedure, we end up with a new set of measurement operators $\mathcal{M} = \{M_i\}$ called *positive operator-valued measures* (POVMs) or generalized quantum measurements. Analogously to projective measurements, the measurement operator M_i determines the probability of obtaining outcome i for a given quantum state ρ, according to

$$p_i = \mathrm{Tr}(\rho M_i). \tag{2.38}$$

Unlike projective measurements, generalized measurement operators do not have to have rank 1, and a POVM does not have to contain $\mathsf{dim}(\mathcal{H})$ elements. In fact, projective measurements are just a special case where $M_i M_j = M_i \delta_{ij}$ and $\mathsf{rank}(M_i) = 1$. Since projective measurements are subsumed under generalized measurements, we will use the notation of POVMs exclusively moving forward.

Formally, a POVM $\mathcal{M} = \{M_i\}$ has the following properties:

$$\begin{aligned} &1. \quad \sum_i M_i = \mathbb{1} \quad \text{(Normalization)} \\ &2. \quad M_i \geq 0 \quad \text{(Positivity)} \\ &3. \quad M_i^\dagger = M_i \quad \text{(Real expectation values)} \end{aligned} \tag{2.39}$$

These properties are enough to guarantee that the expectation values of every POVM behave like a probability distribution. The generalized measurement formalism is particularly important in quantum estimation tasks, and as we will see in Chapter 3, is critical in mitigating the effect of readout errors. Before proceeding further, we will present a few examples that highlight the advantages of POVMs compared to PVMs in information processing and state estimation.

2.6.1 *Example of generalized measurements*

So far, generalized measurements have been motivated in terms of modeling noise and imperfections, but they also serve as a powerful tool for describing measurements that do not suffer from noise. *Informationally complete* (IC) POVMs are sets of measurement operators that form a complete operator basis in Hilbert space. IC POVMs are particularly useful in state estimation tasks, where a density matrix ρ can then be expressed exclusively in terms of the measurement operators

$$\rho = \sum_i c_i M_i, \tag{2.40}$$

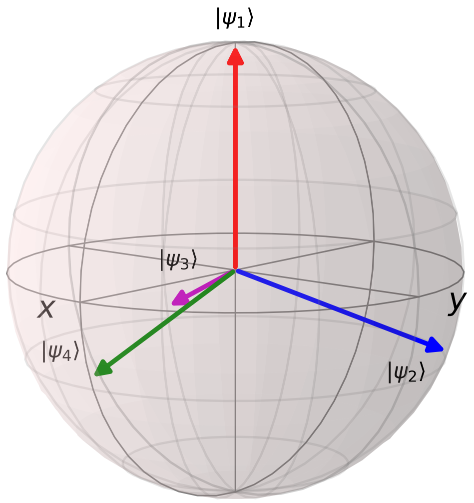


Figure 2.5: The tetrahedron states, Equation 2.41, plotted on the Bloch sphere.

where c_i are some real coefficients, discussed in more detail in Section 2.7. Projective measurements do not form an IC basis and therefore cannot be used to estimate a density matrix without more information.

Symmetric informationally complete (SIC) POVMs are perhaps the most recognized type of generalized measurement. For a single qubit, a SIC POVM can be expressed through a set of tetrahedron states on the Bloch sphere, depicted in Figure 2.5,

$$\begin{aligned}
|\psi_1\rangle &= |0\rangle\,, \\
|\psi_2\rangle &= \frac{1}{\sqrt{3}}\,|0\rangle + \sqrt{\frac{2}{3}}\,|1\rangle\,, \\
|\psi_3\rangle &= \frac{1}{\sqrt{3}}\,|0\rangle + \sqrt{\frac{2}{3}}e^{i\frac{2\pi}{3}}\,|1\rangle\,, \\
|\psi_4\rangle &= \frac{1}{\sqrt{3}}\,|0\rangle + \sqrt{\frac{2}{3}}e^{i\frac{4\pi}{3}}\,|1\rangle\,.
\end{aligned} \tag{2.41}$$

The SIC POVM, denoted $\mathcal{M}_{\text{SIC}} = \{M_i\}$, is constructed by incorporating an extra normalization factor to the projectors derived from these tetrahedron states, specifically $M_i = \frac{1}{2}\,|\psi_i\rangle\,\langle\psi_i|$. SIC POVMs are advantageous because they produce approximately the same information from each measurement due to their symmetric overlap, expressed as $\mathrm{Tr}\big(M_i M_j\big) = (2^n\delta_{ij}+1)/(2^n+1)$. In addition, they are often used because they are a minimal IC set. It is currently unknown whether SIC POVMs exist in all Hilbert-space dimensions, but it is widely conjectured that they do [69, 70].

Another important set of measurements, which are also IC but not minimal, is created with mutually unbiased bases (MUBs). These types of measurements are considered optimal in state reconstruction tasks when classical communication is not allowed and we will discuss these in more detail in Section 7.2.3. We have already seen such bases, the eigenstates of the

Pauli operators in Equation 2.5. The *Pauli-6* POVM is created from the 3 mutually unbiased Pauli-bases, which create an informationally (over)complete set of measurements,

$$\mathcal{M}_{\text{Pauli-6}} = \left\{ \frac{1}{3} |0_i\rangle \langle 0_i| , \frac{1}{3} |1_i\rangle \langle 1_i| \right\} \quad \text{for } i \in \{x, y, z\}. \tag{2.42}$$

An advantage of this generalized measurement is that it is easy to perform in an experimental setting. To implement the Pauli-6 POVM, start with the computational basis measurement $\mathcal{C}^1$ and apply one of the unitary operators $\{e^{-i\pi Y/4}, e^{i\pi X/4}, \mathbb{1}\}$ with equal probability before measuring the qubit.

In applications such as quantum state tomography, having an overcomplete POVMs can be problematic because they create an overcomplete set of linear equations. In general, it is possible to reduce any overcomplete POVMs to a minimal POVM by combining multiple outcomes. For the Pauli-6 POVM this can be done by combining all $|1_i\rangle \langle 1_i|$ outcomes into a single operator. This produces the Pauli-4 POVM,

$$\mathcal{M}_{\text{Pauli-4}} = \left\{ \frac{1}{3} |0_x\rangle \langle 0_x| , \frac{1}{3} |0_y\rangle \langle 0_y| , \frac{1}{3} |0_z\rangle \langle 0_z| , \frac{1}{3} \left(|1_x\rangle \langle 1_x| + |1_y\rangle \langle 1_y| + |1_z\rangle \langle 1_z| \right) \right\}, \tag{2.43}$$

which is both IC and minimal.

2.6.2 *Interpretation of generalized measurements*

In a similar vein to how quantum channels were introduced as a traced-out unitary evolution of a system and environment, generalized measurements can be viewed as a coupling between a system of interest and a measurement probe. This idea was first formalized by John von Neumann [58], who showed that both the measurement device and the system must be treated collectively as a quantum system. Correlations are built up over time between the system and the measurement probe, and if the probe is measured, we also get information about the system of interest. Therefore, the amount of information we can extract from a system depends on the duration of its interaction with the probe. Using such an interaction picture, one can describe partial measurements or weak measurements [71, 72]. If the interaction time is long enough, the probe-system interaction can completely entangle the system with the probe, which then recovers the projective measurement on the system of interest.

To see this in practice, consider the explicit example of measuring only the A qubit of a two-qubit Bell state,

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left(|0\rangle_A |0\rangle_B + |1\rangle_A |1\rangle_B \right). \tag{2.44}$$

If a computational basis measurement $\mathcal{C}^1$ is performed on qubit A, we can compute the post-measurement state of both possible outcomes on qubit B using the Lünders rule in Equation 2.18, and trace out the probe state A. The post-measurement states on qubit B are the same as if we had performed $\mathcal{C}^1$ on qubit B. That is, in this case, measuring the probe qubit A is equivalent to having measured the target qubit B. Therefore, if a probe is allowed to interact long enough with a target state and entangle, one recovers an ideal projective measurement on the target state by measuring the probe.

2.7 QUANTUM STATE TOMOGRAPHY

Having established the formalism of generalized measurements, we are now equipped to address one of the most important applications for measurements, *quantum state tomography* (QST). In experimental quantum systems, we rarely know the prepared state exactly. QST aims to reconstruct the state from measurement data, and thus plays a critical role in both the characterization and validation of quantum hardware [73, 74]. QST is often referred to as the holy grail of system characterization because it provides full knowledge of the quantum state, enabling both the prediction of future measurement outcomes and the computation of arbitrary state properties.

This thesis focuses primarily on the complete reconstruction of quantum states rather than partial or compressed approaches. Full state reconstruction seeks to determine every degree of freedom of the density matrix, providing a complete description of the system. Although this task scales exponentially with system size, it remains a worthwhile and insightful endeavor, as will be discussed in Part II. In the following sections, we review the primary methods for state reconstruction and examine their respective strengths and limitations.

2.7.1 *Linear inversion*

Linear inversion (LI) methods are the simplest and most straightforward approach to full state reconstruction. These methods are based on the decomposition of the density matrix in terms of an IC operator basis. This basis decomposition creates a linear set of equations that can be solved and inverted to recover the full quantum state. In the single-qubit case, one such solution can be found in the Bloch-sphere representation in Equation 2.27, where the Bloch vector can be rephrased in terms of the expectation values of the Pauli-operators, $\vec{a} = (\langle X \rangle, \langle Y \rangle, \langle Z \rangle)$. More generally, it is convenient to express the density matrix in terms of the generalized measurement operators of a POVM. If the POVM is IC and minimal, this decomposition is unique,

$$\rho = \sum_i c_i M_i, \tag{2.45}$$

where c_i are some real coefficients[7]. To find the coefficients, we insert Equation 2.45 into the expectation value of M_j,

$$\langle M_j \rangle = \mathrm{Tr}\big(\rho M_j\big) = \sum_i c_i T_{ij} = p_j, \tag{2.46}$$

where we have introduced the transfer matrix $T_{ij} = \mathrm{Tr}\big(M_i M_j\big)$. Solving for c_i and inserting it back into Equation 2.45, we have an estimator for the density matrix

$$\hat{\rho}_{\mathrm{LI}} = \sum_{ij} p_i \left(T^{-1}\right)_{ij} M_j. \tag{2.47}$$

Although LI methods are attractive, they present a significant limitation; the estimated density matrix may not be physical. This issue becomes especially problematic for pure target states, which experiments typically aim to produce. In Figure 2.6 we show the rate of invalid pure state reconstruction using LI, where less than 50 percent of the states are valid at any given

---

[7] For general expansions of ρ in terms of other matrix bases, these coefficients can be complex.

Figure 2.6: Invalidity of LI state reconstruction for 200 Haar-random single-qubit pure states. The red regions denote unphysical reconstructions, whereas the green areas represent valid state reconstructions. The same set of random states are reconstructed with increasing number of measurements used for each reconstruction. The average rate of invalid state reconstruction is about 50%. Inspired by Ref. [75].

time. An estimator producing unphysical outcomes represents a particularly undesirable trait. To remedy this, we need to move to constrained estimators.

2.7.2 *Maximum likelihood estimation*

A natural extension beyond LI state estimators is *maximum likelihood estimation* (MLE), which offers a statistically grounded approach to quantum state reconstruction. MLE seeks the density matrix $\hat{\rho}$ that is most likely to have produced the observed data set $D = \{o_1, o_2, \dots\}$, while satisfying the conditions in Equation 2.24 [76]. To formalize this, we parameterize the quantum state as $\rho(\vec{\theta})$, where $\vec{\theta} = (\theta_1, \theta_2, \dots)$ denotes a set of estimation parameters. These parameters are optimized under a set of constraints $f_i(\theta) = 0$ that ensure that the resulting state remains physical.

The likelihood function $\mathcal{L}(\rho)$ for a test state ρ is obtained by multiplying the probabilities of the observed outcomes o_i,

$$\mathcal{L}(\rho) = \prod_{o_i \in D} \mathrm{Tr}(\rho M_i), \tag{2.48}$$

where M_i denotes the measurement operator corresponding to the outcome o_i. The MLE estimate is then the state $\hat{\rho}_{\mathrm{MLE}}$ that maximizes $\mathcal{L}(\rho)$ subject to the positivity and trace-one constraints in Equation 2.24,

$$\hat{\rho}_{\mathrm{MLE}} = \operatorname*{argmax}_{\rho} \mathcal{L}(\rho) \quad \text{subject to Equation 2.24}. \tag{2.49}$$

There are two main approaches to determining $\hat{\rho}_{\mathrm{MLE}}$:

1. Physical parameterization: choosing a parameterization $\rho(\vec{\theta})$ that guaranties a physical density matrix for any $\vec{\theta}$.

2. Constrained convex optimization: maximizing $\mathcal{L}(\rho)$ directly over the space of physical states.

The first approach can be realized through the Cholesky decomposition [77],

$$\rho = \frac{AA^\dagger}{\mathrm{Tr}(AA^\dagger)}, \tag{2.50}$$

where A is a lower triangular complex matrix with real diagonal elements, accounting for the $4^n - 1$ degrees of freedom of an n-qubit density matrix. For a single qubit, the second method can be implemented using the Bloch-sphere representation shown in Equation 2.27, with the goal of maximizing $\mathcal{L}(\rho)$ while maintaining $\|\vec{a}\| \leq 1$.

A number of fast and efficient quantum state estimation algorithms based on MLE have been developed [78, 79]. In this thesis, the iterative MLE algorithm proposed by Lvovsky [80] was mainly used for state reconstruction, with implementation details provided in Appendix B.1. One of the main drawbacks of MLE-based methods is that their estimates are not always consistent with confidence regions or error bars [45]. To overcome this limitation, we turn to Bayesian approaches.

2.7.3 *Bayesian mean estimation*

Bayesian mean estimation (BME) provides the most comprehensive framework for quantum state estimation [45]. Unlike MLE, which identifies a single most likely state $\hat{\rho}$ for a data set D, BME assigns a full probability distribution over possible states. This distinction enables the definition of confidence regions and uncertainty estimates, and reveals alternative states that could plausibly explain the observed data.

BME is based on Bayesian inference, featuring an update process in which initial knowledge, represented by a prior distribution π_0, is updated to incorporate new data D, resulting in a posterior distribution π_f using the Bayesian update mechanism $\pi_0 \xrightarrow{D} \pi_f$. Detailed in Appendix B.2, the Bayesian update mechanism for quantum state estimation can be expressed as

$$\pi_f \propto \mathcal{L}(\rho)\pi_0, \tag{2.51}$$

enabling the definition of the Bayesian mean state estimate

$$\hat{\rho}_{\mathrm{BME}} = \int \rho\, \pi_f\, d\rho \propto \int \rho\, \mathcal{L}\,(\rho)\pi_0\, d\rho. \tag{2.52}$$

The proportionality arises because the likelihood function is not an actual probability distribution. To use the likelihood function for Bayesian mean estimation, normalization is necessary. In addition to being able to assign confidence regions by integration, the state is compatible with error bars. This is not always possible with either MLE and LI, as visualized in Figure 2.7, where the 3 different state estimates are compared on the single-qubit Bloch disk.

The largest drawback of BME is the numerical overhead of the representation of the normalized posterior distribution. For two qubits, the posterior distribution is a 15-dimensional distribution, which is quite challenging to represent efficiently. Resampling of such a distribution can be done somewhat efficiently using Monte Carlo methods, but is still difficult due to the sparsity of the distribution [45]. In this thesis, we used an approach based on the Metropolis-Hastings algorithm [81], which is discussed in the Appendix B.3.

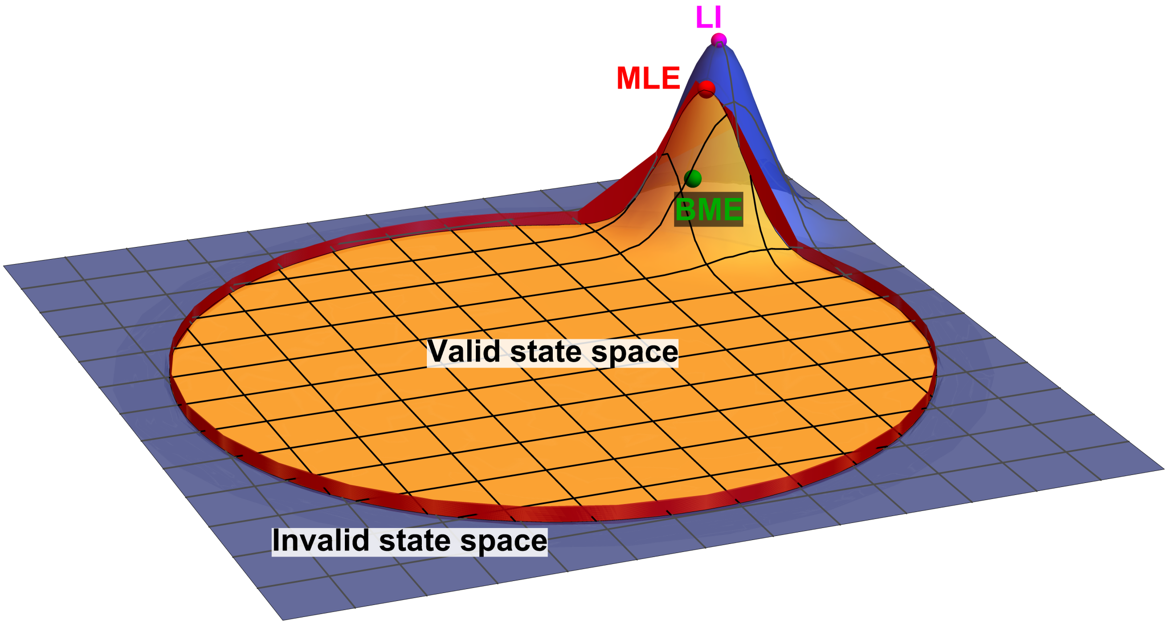


Figure 2.7: Sketch of the LI (magenta), MLE (red), and BME (green) state estimates, together with the corresponding likelihood function on a Bloch disk. The orange surface marks the region containing physically valid quantum states, while the blue surface represents invalid states with negative eigenvalues. The red circumference is the Bloch disk boundary. The LI method yields the most likely state, regardless of its physical validity, the MLE provides the most likely valid state, located here on the boundary of the Bloch disk, and the BME corresponds to the mean state averaged over all valid regions of the likelihood function. Inspired by Ref. [45].

2.7.4 *Partial state estimation*

Full state reconstruction offers complete access to every observable in the system. While valuable for characterizing and benchmarking quantum systems, it has exponential sample complexity and requires significant post-processing because of the exponential number of degrees of freedom. This scaling behavior renders most full state reconstruction methods infeasible even for a moderate number of qubits.

To overcome the exponential scaling, more restricted estimation schemes are necessary that aim only to estimate a smaller part of the Hilbert space, or only a smaller subset of properties of the full quantum state. Luckily, there are many scenarios where only a subset of properties of the quantum states are necessary, and the field of quantum metrology focuses entirely on high precision measurements of single observables. A vast array of methods for partial state estimation has been devised, including approaches such as approximate state tomography [82], neural quantum states [83–85], tensor networks [86, 87], and compressed sensing [88]. In this section, we focus on the two most relevant techniques for the thesis: classical shadow and overlapping state tomography.

Classical shadow tomography

Shadow tomography was first introduced by Scott Aaronson in Ref. [89] and is the predecessor to the more well-known classical shadows. The basic idea behind a shadow[8] of a quantum state involves performing a measurement that has two possible outcomes. The outcome of these

---

[8] The term “shadow” was originally introduced by Steven Flammia [89].

measurements serves as a shadow representation of the state projected onto a plane in Hilbert space. It was shown that it is possible to estimate an exponential number of state properties with only a polynomial number of samples.

> "*The key point is that [Shadow tomography] lets us learn the behavior of a state of exponential dimension, with respect to exponentially many different observables, using only polynomially many copies of the state.*"
>
> – Scott Aaronson [89]

The shadow protocol, although efficient in the number of samples, incurs a high circuit footprint, requiring multiple weak measurements of the same state, rendering it impractical for near-term quantum hardware.

An efficient and more tractable variant of the shadow protocol was introduced in the seminal work by Hsin-Yuan (Robert) Haung dubbed *Classical shadow* [90]. It extends the shadow formalism by introducing an efficient storage of the shadow states and a very efficient circuit to generate the shadows if one can perform random unitaries from a predefined ensemble such as the Clifford gates. Using the efficiently stored shadows, M observables O_i of the full quantum state can be predicted with an additive error of at most ϵ by performing $N \geq \log(M) \max_i ||O_i||^2_{\text{shadow}} / \epsilon^2$ measurements. One significant limitation of the protocol is its reliance on random unitaries during the measurement process, which may introduce substantial overhead in terms of the number of entangling gates necessary to be efficient. The performance guarantee of the shadow protocol is tied to the shadow norm $||O||_{\text{shadow}}$, which depends on the set of unitaries used in the protocol. Specifically for the Clifford group of random unitaries, there is an upper bound on the shadow norm given by $3\,\text{Tr}(O^2)$, which can become exponentially large in many cases, such as the non-local Pauli strings. However, this is not a deficiency of the protocol, as exceeding this bound with single-state measurements is generally unattainable [90].

Overlapping state tomography

Rather than concentrating solely on efficient approaches to estimate specific observables in a larger system, it can be beneficial to fully characterize smaller subsystems, known as *reduced density matrices* (RDMs). *Overlapping state tomography* is an approach developed to reconstruct every subsystem within a larger system [91–93]. The basic idea is rather straightforward. Reconstruction of a k-qubit subsystem within a n-qubit system requires that an IC set of measurements is performed on that k-qubit subsystem. If one only wants to reconstruct a single k-qubit RDM, then performing a k-qubit Pauli-4 measurement on that subsystem is sufficient to reconstruct the RDM. However, if one wants access to all possible k-qubit RDMs within an n-qubit system, this naive measurement strategy would require $\binom{n}{k}$ sets of k-qubit IC measurements to guarantee that all RDMs can be reconstructed. While manageable for small n, it becomes intractable for large systems.

Ideally, one should be able to strategically choose measurements so that a single measurement across all qubits contributes to many different k-qubit RDMs at the same time. Overlapping tomography elegantly solves this issue using a parallel measurement strategy based on perfect hash families [94, 95] and covering arrays [96, 97]. In this thesis, we focus exclusively on the approach put forward by Ref. [91], which uses perfect hash families. Perfect hash families can be briefly defined as a set of functions $\Phi_{N,k,v} = \{\phi_i^{k,v}\}_{i=1}^N$, where N denotes the number of functions in the family, k denotes the number of unique input values, and v denotes the number of unique output values. A k-local perfect hash family is a set where for any set of input values $\{x_1, x_2, \ldots, x_k\}$, there exists a hash function $\phi_i \in \Phi_{N,k,v}$ such that $\{\phi_i(x_1), \phi_i(x_2), \ldots, \phi_i(x_k)\}$ has k unique output values. By encoding qubit labels as inputs to perfect hash families and using the outputs as measurement instructions for the qubits, reconstruction of all k-qubit states can be performed [B, 91]. The number of measurements is proportional to the number of hash functions in the perfect hash family, which is scaled as $e^{\mathcal{O}(k)} \log N$ [91].

It is an **NP**-hard problem to find an optimal perfect hash family with the fewest hash functions, and is therefore not widely known [98]. One possible possible approach to solve this problem is to use a random measurement strategy similar to classical shadow. Instead of using a perfect hash family, one can perform a sequence of n-qubit computational basis measurement $\mathcal{C}^n$ with a random unitary gate $U = \bigotimes_{i=1}^n U_i$ gate, where each U_i is chosen at random from $\{e^{-i\pi Y/4}, e^{i\pi X/4}, \mathbb{1}\}$, similar to the implementation of the Pauli-6 POVM [36].

2.8 QUANTUM DETECTOR TOMOGRAPHY

Just as quantum state tomography reconstructs an unknown quantum state from known measurements, quantum detector tomography (QDT) reconstructs unknown measurement operators using known input states [99]. This is particularly important in experimental systems where the measurement device is imperfect. As discussed in Section 2.6, a general measurement with m outcomes on an n-qubit system is described by a collection of operators $\mathcal{M} = \{M_0, M_1, \ldots, M_{m-1}\}$, containing a total of $2^n(m-1)$ degrees of freedom [100]. Estimating these degrees of freedom can be framed as an inverse problem analogous to state reconstruction. In Section 2.7.1, the Born rule (Equation 2.38) was used to form a system of linear equations in which the unknown density matrix ρ was inferred from known measurement operators. In detector tomography the roles are reversed. The density matrix, now referred to as the calibration state, is known, while the measurement operators are the unknown quantities. To ensure complete reconstruction, an IC set of states must be prepared and measured. For example, measuring the IC and minimal set of tetrahedral states introduced in Equation 2.41 leads to the linear relation

$$p_{si} = \mathrm{Tr}(\rho_s M_i), \tag{2.53}$$

where $\rho_s = |\psi_s\rangle\langle\psi_s|$ and $s \in 1, 2, 3, 4$.[9]

As in state estimation, linear inversion methods are generally discouraged because they may yield unphysical results. Instead, this thesis adopts an iterative MLE approach following Ref. [101], which conceptually mirrors the state reconstruction procedure described in Appendix B.1.

---

[9] While pure calibration states are typically used, they are not essential. Mixed states that are linearly independent can also be employed if there is an independent method to characterize the calibration states.

2.9 QUALITY METRICS FOR INFORMATION EXTRACTION

> "*Any measurement that you make without the knowledge of its uncertainty is completely meaningless!*"
>
> – Walter Lewin [102]

Having reviewed several quantum state reconstruction techniques, we now turn to the question of how to assess their performance. The choice of quality metric is a crucial step in estimation problems and influences what is considered optimal information extraction. In QST we want to know how close our estimated state $\hat{\rho}$ is to the actual prepared state. In contrast, for quantum measurements, we seek to know how much information a POVM $\mathcal{M}$ extracts from a quantum state on average. In this section, we present the metrics that will be used for both tasks: quantum infidelity and quantum Fisher information.

2.9.1 *Quantum infidelity*

In statistics, the classical fidelity $F(p,q)$ is used as a measure of how indistinguishable two probability distributions p and q are from each other,

$$F(p,q) = \sum_i \sqrt{p_i q_i}. \tag{2.54}$$

It quantifies how likely a sample from one distribution is to be attributed to the other. *Quantum fidelity* is a generalization of classical fidelity, where the idea of indistinguishability is extended to quantum states. The quantum fidelity is the smallest possible classical fidelity between two quantum states ρ and σ, optimized over all possible generalized measurements.

$$F(\rho,\sigma) = \min_{\mathcal{M}} \sum_i \sqrt{\mathrm{Tr}(\rho M_i)\,\mathrm{Tr}(\sigma M_i)}, \tag{2.55}$$

where $M_i \in \mathcal{M}$. It is a popular quality metric for state reconstruction and was used extensively in Refs. [A–D]. The optimization in Equation 2.55 can be computed explicitly, yielding [103]

$$F(\rho,\sigma) = \left[\mathrm{Tr}\left(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right)\right]^2, \tag{2.56}$$

where $\sqrt{\rho}$ should be understood as the square root of the eigenvalues λ_i of ρ in an eigendecomposition $\sqrt{\rho} = V\sqrt{D}V^{-1}$, where $\sqrt{D} = \mathrm{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_n})$.

A related and more useful quantity is the *quantum infidelity*, which is defined as

$$I(\rho,\sigma) = 1 - F(\rho,\sigma). \tag{2.57}$$

It is simply the inverse of quantum fidelity, and quantifies how likely the states are to be identified as separate states. This means that for an ideal state estimate, the infidelity should tend towards 0, which is easier to visualize than the fidelity, which tends towards 1. The quantum fidelity simplifies in some important cases, most notably if one state is pure, where the fidelity reduces to the overlap between the two quantum states

$$F_{\text{Pure}}\left(\left|\psi_\rho\right\rangle,\sigma\right) = \left\langle\psi_\rho\right|\sigma\left|\psi_\rho\right\rangle = \operatorname{Tr}(\rho\sigma), \tag{2.58}$$

and for two single qubits

$$F(\rho,\sigma) = \operatorname{Tr}(\rho\sigma) + 2\sqrt{\det(\rho)\det(\sigma)}. \tag{2.59}$$

The quantum fidelity is not a proper distance measure, but it is related to the Bures distance $D_{\mathrm{B}}(\rho,\sigma) = 2\left(1-\sqrt{F(\rho,\sigma)}\right)$. In the limit of small infidelity $1-F(\rho,\sigma) \ll 1$, one recovers infidelity $1-F(\rho,\sigma)$ as the Bures distance, $1-\sqrt{F(\rho,\sigma)} = 1-F(\rho,\sigma) + O\left((1-F(\rho,\sigma))^2\right)$.

2.9.2 *Fisher information*

Fisher information quantifies how much information on average an observable O carries about an estimation parameter θ. This notion can be used to quantify how much information a POVM $\mathcal{M}$ carries about the parameters of a density matrix. We will explore this idea in more depth in Chapter 7, where we examine how Fisher information can be used to find optimal measurement strategies, which significantly improves the accuracies of the state reconstruction.

The Fisher information is formally defined as the variance of the score function

$$\mathcal{I}(\theta) = \operatorname{Var}(s(\theta)), \tag{2.60}$$

where the score function is defined as

$$s(\theta) = \frac{\partial \ln(\mathcal{L}(\theta))}{\partial\theta}. \tag{2.61}$$

High Fisher information means that the value of estimation parameters θ changes significantly due to small amounts of measurements, while low Fisher information means that the estimation parameters are mostly stationary with more measurements. The Fisher information can be used to establish a lower bound on the uncertainty of an estimator, as described by the Cramér-Rao bound [104–107].

$$\operatorname{Var}(\hat{\theta}) \geq \frac{1}{\mathcal{I}(\theta)}. \tag{2.62}$$

For multiple parameters $\vec{\theta} = (\theta_1,\theta_2,\dots)$ we can define the *Fisher information matrix* (FIM) with two equivalent formulations

$$[\mathcal{I}(\theta)]_{ij} = \mathbb{E}\left[\frac{\partial \ln(\mathcal{L}(\theta))}{\partial\theta_i}\frac{\partial \ln(\mathcal{L}(\theta))}{\partial\theta_j}\right] = -\mathbb{E}\left[\frac{\partial^2 \ln(\mathcal{L}(\theta))}{\partial\theta_i\partial\theta_j}\right]. \tag{2.63}$$

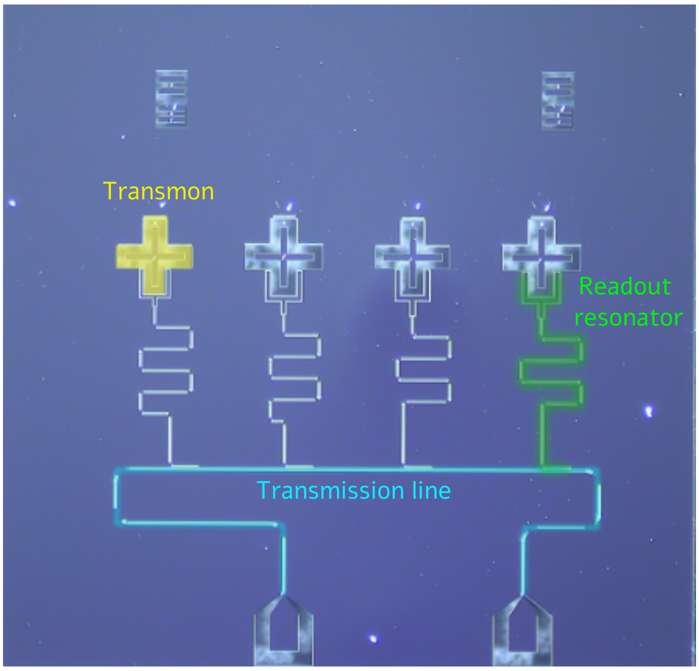


Figure 2.8: An image of four transmon qubits coupled to individual readout resonators, which are all connected to a shared transmission line. The chip was fabricated at Yale and operated at KIT. Image was modified with permission [C].

2.10 SUPERCONDUCTING QUBITS

> "*Any sufficiently advanced technology is indistinguishable from magic.*"
>
> – Arthur C. Clarke, Clarke's third law

Naturally occurring two-level quantum systems are uncommon, and those that can be manipulated are even scarcer. Consequently, the majority of techniques for producing qubits rely on constraining or manipulating the accessible Hilbert space of systems that we can precisely control. There are multiple architectures available for quantum computing and simulation that do this in different ways, such as NV-centers [108, 109], photonic systems [110, 111], Rydberg atoms [19, 112], and trapped ions [113, 114]. This thesis concentrates predominantly on superconducting qubit systems [115–118]. We will give a brief overview of superconducting qubit-based systems. For a more comprehensive review, see Ref. [119].

Superconducting qubits are based on the superconducting LC circuit, which when sufficiently cooled down to reach superconductivity, can act as a quantum harmonic oscillator by removing electrical resistance in the circuit. A quantum harmonic oscillator has a constant spacing between energy levels, giving a natural ground state $|0\rangle$ and an excited state $|1\rangle$ through the two lowest energy levels. However, due to the linear energy spacing of an harmonic oscillator, driving the transition between $|0\rangle$ and $|1\rangle$ will also drive higher excitations. To prevent this, a nonlinear component is introduced into the LC-circuit to create nonlinear transitions between the energy levels. A Josephson junction [120, 121] acts as a nonlinear inductor that generates a cosine-shaped potential, enabling varied transition energies. There are many types of superconducting qubits, differing on the explicit implementation of the nonlinearity, and physical mechanism behind the harmonic oscillator, see e.g. flux qubits, phase qubits, and fluxonium qubits. A variant called *transmon* qubits is shown in Figure 2.8. Four isolated transmon qubits are coupled via individual readout resonators to a collective readout transmission line. This chip was used for the experiments presented in Refs. [C, F].

The dominating source of errors in superconducting devices comes from parasitic two-level systems (TLS) [122], which are accidentally created qubit-like structures on the chip that can weakly couple to the primary qubits and make them decohere. Another source of errors in these chips is thermal noise [123]. Thermal noise primarily manifests during qubit operation, where the qubit state can be spontaneously excited. To minimize the circuit footprint and reduce the need for additional experimental hardware, such as amplifiers, the transmission line is shared between the four qubits. By multiplexing the readout pulses on the transmission line, and having different resonance frequencies in the resonators, it is possible to address each qubit individually and operate them in parallel. Ideally, the resonators are spaced sufficiently far apart so that a Gaussian pulse targeting one resonator does not overlap with the resonance frequency of any other resonator. However, due to challenges in hardware manufacturing, some resonator frequencies are close, causing a pulse to interact with multiple resonators [C, F]. This creates correlated readout errors that are problematic for scaling the hardware. Given the multitude of error sources, robust error handling becomes essential to both the operation and the scalability of superconducting qubit architectures, especially when aiming to reach regimes where these systems can deliver practical utility.

Part II

General readout error mitigation

3

READOUT-ERROR MITIGATION

The goal of this chapter is to introduce the necessary background to understand the methodology and main results of the series of publications [A–C, F]. Together, these works explore methods to comprehensively characterize and mitigate readout errors in superconducting qubit devices. The developed approaches are tailored to these devices, but specifically the approaches presented in Refs. [A, B] were designed to be agnostic to the underlying hardware used. The chapter begins with Section 3.1, where we give an overview of the role of error management in quantum devices and outline the two main approaches, error correction and error mitigation. We then narrow our focus down to the special case of readout errors, a class of noise that remains even with fault-tolerant hardware. In Section 3.2 we review how readout errors are modeled, the difficulties in disentangling them from state preparation errors, and the main mitigation strategy using classical stochastic maps. Throughout this section, we use a tilde, $\tilde{O}$, to indicate operators affected by noise.

3.1 ERROR MANAGEMENT IN QUANTUM DEVICES

A cornerstone of any technological system is robustness against noise and external disturbances. This is even more important in the realm of quantum technologies. Quantum systems are particularly susceptible to noise, not only because interactions and measurements tend to disrupt them, but also because of the inherently delicate nature of quantum information. Currently, the most pressing challenge facing quantum platforms is realizing fault-tolerant computation and simulation. Before introducing the main methodologies, it is important to clarify the terminology. When we refer to *noise*, we are talking about any physical processes capable of affecting a system, including thermal fluctuations, state decay, unexpected changes in magnetic fields, and out-of-qubit excitations. *Errors*, on the other hand, are consequences of noise on our system, such as physical bit flips or phase decoherence in quantum states.

3.1.1 *Error correction*

Quantum error correction (QEC) represents the most extensive and dynamic method of error management within quantum systems. Its primary aim is to redundantly encode quantum states to ensure that the original information can be recovered in the event of errors during quantum processes. In the simplest form of classical error correction, a logical bit is encoded into multiple physical bits, such as $0_L = 000$ and $1_L = 111$, where the L subscript indicates the logical bits. If an error occurs in one of the physical bits, a majority vote is conducted by measuring the bits and adjusting the one that differs. This approach is not possible for the quantum system due to the no-cloning theorem [124–126], which asserts the impossibility of duplicating a quantum state and that projective measurements collapse the state.

Quantum error correction starts out similarly to classical error correction by encoding a logical qubit into multiple physical qubits, $|0\rangle_L = |000\rangle$ and $|1\rangle_L = |111\rangle$. To enable similar fault-tolerant redundancy in quantum systems, we need a way to measure whether an error occurred without collapsing the logical quantum state we want to protect. Syndrome measurements are a special class of measurements that project the set of physical qubits onto a parity state, which tells you if the states have even or odd parity and which qubit is different, without giving you any information about the logical qubit state [40, 127]. Depending on the outcome of the syndrome measurement, a series of unitary operations are applied to the physical qubits, e.g. $|\tilde{\psi}\psi\psi\rangle$, to recover the correct logical qubit state $|\psi\rangle_L$.

In both of these simple classical and quantum error correction techniques, the protocols depend on only one error occurring at the time. If more than one classical bit flips, the majority vote gives the incorrect bit value. For quantum systems, there are more than just bit-flip errors. In addition to correcting bit-flip errors, coherent phase-flip errors must be addressed in a similar manner, resulting in a greater number of physical qubits necessary for encoding one logical qubit [127]. The number of physical qubits required to encode a logical qubit depends on the error correcting code used. Stabilizer codes can be characterized by $[[n, k, d]]$, where n is the total number of qubits that the code uses, k is the number of logical qubits it encodes, and d is the code distance [128]. The distance tells you the minimal number of cumulated physical qubit errors that can go undetected. As a general rule, the number of physical errors that can be corrected with a code is $t = \lfloor \frac{d-1}{2} \rfloor$. The physical qubit error tolerance of the code depends on the specific encoding and error model considered, where generally more complicated codes yield better thresholds. Surface codes [129, 130] require a gate error below 1% [131]. Superconducting qubits have shown error rates of about 0.1% for single-qubit gates and two-qubit gates on the order of 1.6% [132] for modest gate depth. One of the largest surface codes implemented to date with some success is the 5x5 surface code [28, 133, 134]. Although significant advancements have been made in the last decade, a fully developed error-corrected quantum computer remains somewhat distant [33, 34].

3.1.2 *Error mitigation*

The other major error management approach is *quantum error mitigation* (QEM). Designed for noisy intermediate-scale quantum devices [26], QEM techniques do not aim to directly correct physical errors during quantum operations, but seek to mitigate the impact of errors through post-processing. The objective of QEM can be phrased as removing bias in estimators caused by noise processes [35]. The bias of an observable O is defined

$$b_O = E[\hat{O}] - \mathrm{Tr}(\rho_0 O) = \mathrm{Tr}((\tilde{\rho} - \rho_0)O), \tag{3.1}$$

where ρ_0 is the noiseless state, $\hat{O}$ is an estimator for the observable and $\tilde{\rho}$ is an effective noisy state due to a noise process $\tilde{\rho} = \mathcal{E}_{\text{noise}}(\rho_0)$. The goal is to reduce the error in $\hat{O}$, which is equivalent to reducing the mean squared error (MSE) of the observable,

$$\text{MSE} = E[(\hat{O} - \text{Tr}(\rho_0 O))^2] = b_O^2 + \text{var}[\hat{O}^2] \tag{3.2}$$

The two terms in Equation 3.2 represent two different sources of errors in the estimated observable. The first term is the bias caused by the noise process in Equation 3.1, while the second is due to statistical fluctuations. The bias term can be reduced by creating an error-mitigated estimator $\hat{O}_{\text{EM}}$, while the statistical fluctuation term can generally only be reduced by performing more measurements, $\text{var}[\hat{O}^2] \propto N^a$, where N is the number of measurements and a is some scaling exponent smaller than 0. In this part, we will focus on reducing the bias term, while in Part III we will look at how the variance term can be reduced.

Error-mitigated estimators generally work by enhancing certain features of the data while suppressing others. This increased focus on a subset of the data generically means that the effect of statistical fluctuations increases, $\text{var}[\hat{O}_{\text{EM}}] > \text{var}[\hat{O}]$. This bias-variance trade-off means that error-mitigated estimators are not universally the best option for any estimation task and that some caution when choosing estimators is warranted. The literature on QEM methods is vast and a comprehensive review can be found in Ref. [35]. For the purposes of this thesis, we will limit our discussion to readout-error-mitigation methods.

3.2 READOUT ERRORS

Readout error plays a special role in error handling as it is not addressed by QEC or standard error-mitigation protocols, which focus mainly on correcting gate errors. A common and useful assumption is that readout errors can be modeled as a noise process that acts immediately prior to the measurement. Specifically, any prepared quantum state ρ is first affected by a readout noise channel $\mathcal{E}_{\text{R}}$, after which an ideal generalized measurement $\mathcal{M}$ is performed. This is schematically illustrated in Figure 3.1.

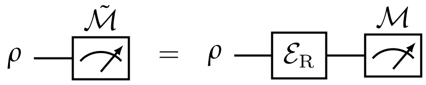


Figure 3.1: Noisy readout $\tilde{\mathcal{M}}$ can be modeled as a noise channel $\mathcal{E}_{\text{R}}$ followed by an ideal measurement $\mathcal{M}$.

Although this formulation places the noise channel on the quantum state, it is mathematically equivalent to applying a dual (adjoint) channel to the measurement operators. Using the Born rule (Equation 2.38) and the operator-sum representation (Equation 2.32), this equivalence can be demonstrated as follows:

$$p_i = \text{Tr}(\mathcal{E}_{\text{R}}(\rho) M_i) = \sum_k \text{Tr}\big(E_k \rho E_k^\dagger M_i\big) = \sum_k \text{Tr}\big(\rho E_k^\dagger M_i E_k\big) = \text{Tr}\big(\rho \mathcal{E}_{\text{R}}^\dagger(M_i)\big), \tag{3.3}$$

where we have defined the dual channel $\mathcal{E}_{\text{R}}^\dagger$ as the channel $\mathcal{E}_{\text{R}}$ with each operator complex conjugated [135]. This rearrangement is analogous to switching to the Heisenberg picture, where time evolution acts on the observables rather than on the state [135, 136].

Readout errors are often grouped with state preparation errors under the umbrella term *state preparation and measurement* (SPAM) errors [38, 137–142]. In standard experimental setups, these two error sources are inherently inseparable and one must be assumed to characterize the other. This inseparability leads to ambiguity when attributing observed imperfections to either preparation or readout processes. This issue is illustrated in Figure 3.2. In the calibration step (a), perfect state preparation is often assumed. Any imperfections in the actual preparation process, $\mathcal{E}_{\mathrm{S}}$, are therefore absorbed into the effective readout model $\tilde{\mathcal{N}}$. However, when gates or circuits are applied between preparation and readout (b), this assumption breaks down. The error channels separate, and the original calibration no longer accurately describes the readout noise process.

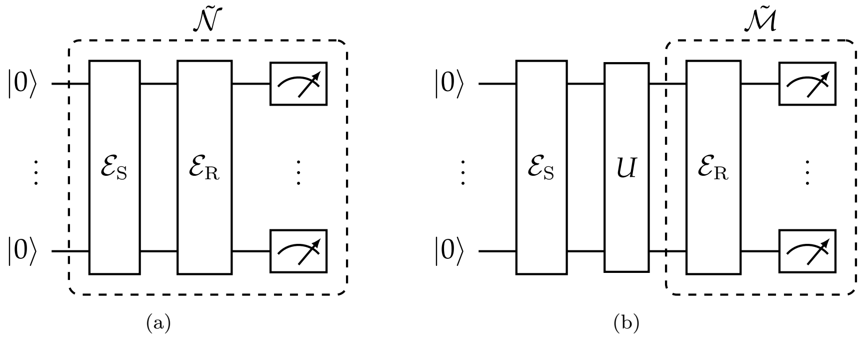


Figure 3.2: Illustration of SPAM inseparability, with a state preparation channel $\mathcal{E}_S$ and a readout-error channel $\mathcal{E}_R$. (a) Calibration assumes perfect state preparation, absorbing $\mathcal{E}_{\mathrm{S}}$ into the effective noisy measurement $\tilde{\mathcal{N}}$. (b) If operations are performed between state preparation and readout, the two error channels are separated. The calibrated measurement operator $\tilde{\mathcal{N}}$ will in general not be compatible with the measurement operator $\tilde{\mathcal{M}}$.

Despite this, in many practical cases, the contribution of state preparation errors is small in comparison to readout. For example, in superconducting qubits, readout errors are significantly larger (10-20 times) than ground state preparation and single-qubit gate errors [132]. Furthermore, there are experimental scenarios where state preparation and readout errors can be independently characterized. For example, in optical setups, the measurement process could involve a photon collecting lens which distorts the polarization. The effect of the lens can be isolated from the state preparation apparatus by removing the lens and calibrating the state preparation errors separately.

Readout errors can be broadly split into two different groups, classical and coherent errors. Classical errors can be described as a simple misidentification of different outcomes, while coherent errors involve unwanted off-diagonal contributions in the measurement operator. Classical errors are by far the most prominent in quantum systems [B, 143–145], and they are also the simplest to mitigate. In the remainder of this chapter, we will focus on techniques for mitigating classical readout errors. Coherent error-mitigation strategies will be addressed in Chapter 4.

3.2.1 *Classical readout-error mitigation*

Classical readout errors, due to bit-flip-type misassignments, are often dominant and tractable. Various classical approaches exist to model and mitigate the effect of readout errors, such as unfolding [146, 147], transfer matrix[1] inversion [143, 148], model fitting [144, 149], active error mitigation [150], compressed readout [151, 152] and detector tomography-based mitigation [36, 143]. Despite differences in methodology, these approaches fundamentally rely on inverting stochastic maps to recover the true measurement statistics.

When modeling classical errors, it is beneficial to begin by examining the effects of applying a general channel to the computational basis. Given that the end goal is to recover ideal measurement statistics, it is helpful to express the noisy measurement operators with respect to the ideal, noiseless measurement operators. Recall from Section 2.6 that projectors do not form a complete operator basis, but only span a subspace. Consequently, a general operator cannot be expressed only in terms of the computational basis. Therefore, to represent a noisy computational basis measurement $\tilde{P}_i \in \tilde{\mathcal{C}}^n$ in terms of an ideal computational basis measurement $P_i \in \mathcal{C}^n$, it is necessary to write

$$\tilde{P}_i = \sum_j T_{ij} P_j + \Delta_i, \tag{3.4}$$

where T_{ij} is a *transfer matrix*[2], and Δ_i represent errors that remove noisy measurement operators from the computational basis subspace. Classical errors may be expressed as misidentifications of measurement outcomes. Therefore, a classical error channel should map an ideal computational basis measurement $\mathcal{C}^n$ to a redistributed computational basis measurement $\tilde{\mathcal{C}}^n$. This means for a purely classical error channel, $\Delta_i = 0$ and we can rewrite Equation 3.4 to

$$\tilde{P}_i = \sum_j T_{ij} P_j. \tag{3.5}$$

In measuring a quantum state ρ, Equation 3.5 can be expressed using the noisy measurement probabilities $\tilde{p}_i$ alongside the ideal probabilities p_i,

$$\tilde{p}_i = \mathrm{Tr}\big(\rho \tilde{P}_i\big) = \sum_j T_{ij}\, \mathrm{Tr}\big(\rho P_i\big) = \sum_j T_{ij} p_j. \tag{3.6}$$

This linear set of equations can be solved by inverting the transfer matrix, yielding an expression for the ideal measurement probabilities,

$$p_i = \sum_j \left(T^{-1}\right)_{ij} \tilde{p}_j. \tag{3.7}$$

Observables in the computational basis subspace can therefore be estimated straightforwardly from the noisy measurement outcomes.

$$\langle O \rangle = f(p) = f(T^{-1}\tilde{p}). \tag{3.8}$$

The transfer matrix can be straightforwardly determined by directly estimating its elements, which are found through the preparation and measurement of computational basis states,

$$T_{ij} = \langle j| \tilde{P}_i |j\rangle\,. \tag{3.9}$$

---

[1] Sometimes called confusion, assignment or response matrix.

[2] T must be a left-stochastic map, $\sum_i T_{ij} = 1$. This property guaranties that any normalized vector v, $\|v\| = 1$, remains normalized after a stochastic transformation $\|Tv\| = 1$.

Alternatively, it is possible to perform full QDT and remove any off-diagonal elements in the reconstructed measurement operators [143].

The extreme simplicity of this noise model is not without drawbacks [143]. The above procedure is equivalent to the linear inversion estimator of quantum states discussed in Section 2.7.1, and suffers from the same problems. Since T is a left-stochastic matrix, the resulting probability vectors p_i are normalized, but their entries do not have to be positive. This can come from statistical fluctuations and coherent sources of noise. Consequently, more adjustments are necessary, resulting in the final estimate being the adjusted probability vector p_i^*, which incorporates new biases. Despite these limitations, the impressive success of this straightforward method is evident, and it has already been incorporated into the majority of existing quantum platforms [153–158].

The basic formulation above is not scalable to large numbers of qubits, since T is a $2^n \times 2^n$ matrix with all real entries for a n-qubit system. Each row is normalized, leaving $4^n - 2^n$ degrees of freedom that need to be estimated. This creates exponential overhead in both sampling and post-processing, making full classical readout-error mitigation intractable. Fortunately, readout errors in modern hardware tend to be local or restricted to certain subsets of qubits, which allows for a single-qubit tensor-product ansatz

$$T = \bigotimes_{i=1}^{n} T^i, \tag{3.10}$$

where T^i is the transfer matrix for qubit i. Recently, there has been an effort to construct models which go slightly beyond the single-qubit ansatz. Observations of correlations between qubits have been reported in superconducting qubits [C, 36, 159, 160]. By integrating local qubit correlations, clusters of qubits are formed and jointly error-mitigated [B, 36, 161].

3.2.2 *Randomized readout*

Up to this point, we have explored techniques that keep the original experiment intact and only perform an initial calibration at the beginning. However, there are also readout-error mitigation techniques that modify the measurement process itself. One such technique is *measurement randomization*, which acts as a symmetry enforcer by averaging certain noise effects. Measurement randomization can transform coherent errors into stochastic errors, which subsequently enables classical mitigation techniques. An approach to symmetrizing computational basis measurements is *Pauli-twirling* [35]. In this method, before making a measurement in the n-qubit computational basis $\mathcal{C}^n$, a random unitary operation is performed that rotates the qubits around the z-axis, expressed as $U_i = \exp(-i\theta_i Z)$, with θ_i being uniformly selected from the interval $[0, 2\pi)$. This transformation

$$\mathcal{C}^n \to \mathcal{C}^n_{\text{sym}} = \frac{1}{N}\sum_{i=1}^{N} U_i \mathcal{C}^n U_i^\dagger \tag{3.11}$$

symmetrizes the error channels so that $\mathcal{C}^n_{\text{sym}}$ predominantly contains classical errors. Another approach using randomized measurements is classical shadow tomography, already discussed in Section 2.7.4. *Robust shadow tomography* was introduced as an error-mitigated version of classical shadow tomography [162]. In this approach, random unitary operations are also applied before measurement, but they are drawn from a different distribution compared to Pauli-twirling, like random Clifford gates.

The main drawback of these approaches is the experimental difficulty in implementing random gates [163, 164]. For robust shadow tomography, the realization of global Clifford gates requires many entangling operations, which are particularly prone to noise. While experimentally demanding, randomized readout serves as a bridge between coherent and classical readout errors, offering a powerful complement to calibration-based methods. We will now move away from classical readout-error mitigation strategies and focus on one of the two major contributions of this thesis: robust coherent readout-error mitigation in quantum state tomography.

4

MITIGATING COHERENT READOUT ERRORS IN SUPERCONDUCTING QUBITS

In Chapter 3, we discuss the limitations of classical readout-error-mitigation methods and their typical reliance on simplified error models. The aim of the presented work was to come up with a comprehensive approach that addresses both coherent and classical readout errors, while avoiding the overhead associated with complete system characterization protocols, such as gate set tomography [137, 165, 166]. The essence of our readout-error-mitigated state tomography (REMST) protocol lies in the observation that to go beyond classical readout errors, one must track the noisy evolution of the full quantum state, as coherent and classical noise together impact all the degrees of freedom of the quantum state. Consequently, a coherent readout-error-mitigation strategy must encompass full quantum state characterization. Using this understanding, we bypass the use of probability vectors and integrate the noisy measurement operators directly into a likelihood-based quantum state estimator. The protocol was applied to a superconducting qubit, where experimental noise sources were deliberately boosted to test the robustness of the protocol. Characterization of the measurement operators reveals significant coherent errors, indicated by off-diagonal elements in the reconstructed computational basis measurement, corresponding to a non-zero Δ_i in Equation 3.4. The protocol allowed us to consistently obtain better state reconstruction accuracy across a variety of experiments, with up to 30 times improvement in infidelity for particularly noisy experiments.

Author participation is stated at the end of the manuscript: "*Development of the protocol and software was done by* ***A. Aasen*** *with supervision from M. Gärttner. The experimental realization and data generation was done by A. Di Giovanni with supervision from H. Rotzinger and A. Ustinov.* ***A. Aasen*** *and A. Di Giovanni prepared the draft for the manuscript. All authors contributed to the finalization of the manuscript.*".

communications physics

Article

https://doi.org/10.1038/s42005-024-01790-8

Readout error mitigated quantum state tomography tested on superconducting qubits

Adrian Skasberg Aasen [1,2,5], **Andras Di Giovanni** [3,5], **Hannes Rotzinger** [3,4], **Alexey V. Ustinov**[3,4] & **Martin Gärttner** [2]

Quantum technologies rely heavily on accurate control and reliable readout of quantum systems. Current experiments are limited by numerous sources of noise that can only be partially captured by simple analytical models and additional characterization of the noise sources is required. We test the ability of readout error mitigation to correct noise found in systems composed of quantum two-level objects (qubits). To probe the limit of such methods, we designed a beyond-classical readout error mitigation protocol based on quantum state tomography (QST), which estimates the density matrix of a quantum system, and quantum detector tomography (QDT), which characterizes the measurement procedure. By treating readout error mitigation in the context of state tomography the method becomes largely readout mode-, architecture-, noise source-, and quantum state-independent. We implement this method on a superconducting qubit and evaluate the increase in reconstruction fidelity for QST. We characterize the performance of the method by varying important noise sources, such as suboptimal readout signal amplification, insufficient resonator photon population, off-resonant qubit drive, and effectively shortened T_1 and T_2 coherence. As a result, we identified noise sources for which readout error mitigation worked well, and observed decreases in readout infidelity by a factor of up to 30.

Building quantum machines capable of harnessing superposition and entanglement promises advances in many fields ranging from cryptography[1,2], material simulation[3], and drug discovery[4,5] to finance[6–8] and route optimizations[9]. While these applications assume a large-scale fault-tolerant quantum computer, currently, we are still in the era of noisy intermediate-scale quantum (NISQ) devices[10], where noise and imperfection of qubits fundamentally limit the applicability of quantum algorithms. In this work, we refer to noise as the physical phenomenon which causes variable operation of the experiment, and error is the deviation of the measured quantities with respect to their theoretical expectation values. To harness the power of quantum computation and quantum simulation today, we need methods that are suitable for working with noisy hardware.

There are predominantly two approaches to combating errors in qubit systems. On the one hand, one can study the exact sources of noise and errors, in order to find new materials or techniques to eliminate them. For superconducting qubits, for instance, parasitic two-level systems coupling to qubits or resonators have been shown to be a limiting factor for qubit coherence[11], leaving room for further hardware improvements.

On the other hand, instead of reducing the noise inherent in the hardware, one can apply algorithmic methods to handle and reduce the errors. There are primarily two approaches to algorithmic error correction[12], of which the first and most prominent is quantum error correction. The core idea here is to encode logical qubits in multiple noisy physical qubits. Adaptive corrective operations are then performed based on syndrome measurements[13]. These methods require low enough gate errors such that the qubit number overhead does not blow up. Current gate fidelities are too low for general applicability, however, promising results have been achieved by using surface codes[14–17].

One prominent source of errors not captured in error correction are readout errors, a subset of state preparation and measurement (SPAM)-errors[18,19]. Readout errors occur in the process of reading out the state of the qubit, e.g., spin projection measurements. Experimentally, one can

---

[1]Kirchhoff-Institut für Physik, Universität Heidelberg, Im Neuenheimer Feld 227, 69120 Heidelberg, Germany. [2]Institut für Festkörpertheorie und -optik, Friedrich-Schiller-Universität Jena, Max-Wien-Platz 1, 07743 Jena, Germany. [3]Physikalisches Institut, Karlsruher Institut für Technologie, Kaiserstraße 12, 76131 Karlsruhe, Germany. [4]Institut für QuantenMaterialien und Technologien, Karlsruher Institut für Technologie, Kaiserstraße 12, 76131 Karlsruhe, Germany. [5]These authors contributed equally: Adrian Skasberg Aasen, Andras Di Giovanni. ✉e-mail: martin.gaerttner@uni-jena.de

understand the importance of such errors through the example of a superconducting resonator measurement of a transmon. When using only a few photons to read out the resonator, there is a nonzero probability of an incorrect readout. However, by increasing the number of photons, higher qubit levels will be excited, leading to a sharp drop in readout fidelity[20].

Readout errors are better handled by the second approach, quantum error mitigation. These kinds of methods do not correct the errors in execution of the quantum algorithm, but rather reduce the errors by post-processing the measured data. There exists a broad spectrum of error mitigation methods[12,21,22], a large class of which focuses on readout errors, such as unfolding[23,24], T-matrix inversion[25], noise model fitting[26,27], and detector tomography-based methods[28]. Other promising results have recently been shown on an NISQ processor[29] using zero-noise extrapolation[30,31].

Our proposed method falls into the category of detector tomography-based mitigation, similar to ref. 28, where the focus is on the correction of readout probability vectors. Detector tomography-based methods use a set of calibration states, e.g., the eigenstates of the Pauli matrices, to characterize the measurement outcomes realized by the measurement device using the positive operator-valued measure (POVM) formalism. By looking at the reconstructed measurements, one gains information about the noise present in the experiment[32], which can then be leveraged for error mitigation. We emphasize that readout error mitigation operates on the statistics of output distributions collected from multiple runs of the experiment, and is not able to correct the results of an individual projective measurement.

Our approach generalizes the ideas of previous protocols by using a complete characterization of a quantum system. This is done by quantum state tomography, which directly estimates the density matrix of the system[33,34]. The question of QST in error mitigation was also considered in ref. 28, but restricted to only classical errors, i.e., errors that can be described as a stochastic redistribution of the outcome statistics of single basis measurements. Reference 32 does perform a general reconstruction of the detectors, and suggests it can be used in QST, but does not provide any protocol that goes beyond correction of marginals with numerical inversion, mostly focusing on classical errors in their applications. Similar ideas were compared to data pattern tomography in refs. 35,36 with linear inversion. In ref. 37, some of the error mitigation methods mentioned above were compared with simulated depolarizing noise in the context of state reconstruction. The central aspect of our approach is the direct integration of QDT with QST, which leads to an error mitigation scheme that does not require matrix inversions or numerical optimizations. On top of this the protocol makes no assumptions about the error channel that represents the readout noise or about the type of quantum state (e.g., entangled, separable, pure, mixed) that it applies to. As we do not restrict the set of allowed POVMs in the detector tomography in any way, the protocol is also agnostic to the architecture (e.g., superconducting qubits, photons, ultracold-atoms, neutral atoms) and the type of readout mode (e.g., single projective measurement (photons), resonator readout (superconducting qubits)).

With this general scheme, we test the limits of readout error mitigation by subjecting the protocol to common noise sources. Previous works have been extensively tested on publicly accessible IBM quantum computers, which only allow for very limited experimental control over noise sources. In this paper, we evaluate our proposed method on a chip, where we have full control over the experiment. This allows us to systematically induce and vary important noise sources, which leads to a better understanding of the strengths and limitations of the protocol.

Results and discussion

Preliminaries

This section reviews the relevant theoretical concepts to describe general quantum measurements and state reconstruction[13,38,39]. Readers familiar with these concepts may want to skip this section.

Notation used in this paper unless stated otherwise: Objects with tilde, e.g., $\tilde{p}$, indicate that it has been subject to readout noise. Objects with a hat, e.g., $\hat{p}$, are experimentally measured quantities, which contain sample fluctuations and noise. The hat is also used to denote estimators, which will be clear from the context. Operators and objects without either accent will be considered theoretically ideal objects. The two eigenstates of the Pauli operators σ_x, σ_y, and σ_z, will be denoted by $|0_i\rangle$ and $|1_i\rangle$, where $i \in \{x, y, z\}$. If no subscript is given, the z-eigenstates are implied.

Generalized measurements. Measurements on quantum systems are described by expectation values of operators, given by Born's rule[13,38],

$$\langle O\rangle = \mathrm{Tr}(\rho O), \tag{1}$$

where ρ defines the quantum system and O represents a measurable quantity. Experiments typically perform projective measurements onto a set of basis states. The simplest example is a measurement on a computational basis, with two possible outcomes. The projective operators take on the form $P_0 = |0\rangle\langle 0|$ and $P_1 = |1\rangle\langle 1|$, where $p_i = \mathrm{Tr}(\rho P_i)$ gives the probability of obtaining the outcome i.

In this work, we are interested in noisy measurements. To be able to capture realistic readout we need to move to generalized quantum measurements, described by positive operator-valued measures (POVMs). A POVM is a particular set of operators $\{M_i\}$, sometimes called effects, that have the interpretation of yielding a probability distribution over measurement outcomes through their expectation values. In particular, any POVM has the three following properties[13]:

$$M_i^\dagger = M_i, \quad M_i \geq 0, \quad \text{and} \quad \sum_i M_i = \mathbb{1}, \tag{2}$$

where the first property guarantees that the resulting expectation values are real, the second guarantees positive expectation values, and the third guarantees that the expectation values sum to unity. Altogether, we have a natural correspondence between each operator M_i and the probability of getting outcome i of a random process,

$$p_i = \langle M_i\rangle = \mathrm{Tr}(\rho M_i). \tag{3}$$

Through this general interpretation of the effects M_i, one can assign each effect to a possible outcome from a measurement device.

We are particularly interested in POVMs that form a complete basis in their respective Hilbert space. Such a POVM is called informationally complete (IC). Using an IC POVM allows one to decompose any operator O in terms of the set $\{M_{\mathrm{IC},i}\}$,

$$O = \sum_i c_i M_{\mathrm{IC},i}. \tag{4}$$

If the POVM is minimal, meaning that its elements are linearly independent, the coefficients c_i are unique. For qubit systems to be informationally complete, we need 4^n linearly independent POVM effects, where n is the number of qubits in the system. An example of an IC POVM is the so-called Pauli-6 POVM. Experimentally, such a measurement can be performed by randomly selecting either the $x-$, $y-$, or $z-$ basis and then performing a projective measurement in that basis. This POVM is defined by the set $\left\{\frac{1}{3}|0_i\rangle\langle 0_i|, \frac{1}{3}|1_i\rangle\langle 1_i|\right\}$, where $i \in \{x, y, z\}$. Note that the Pauli-6 POVM is not minimal, and its elements are not projectors, due to the prefactor acquired by the random basis selection.

Representation of readout noise. Measurements of quantum objects suffer from imperfections caused by the errors introduced by the measurement device. This means that when intending to perform the POVM $\{M_i\}$, we actually perform the erroneous POVM $\{\tilde{M}_i\}$. In reading out the quantum state ρ we observe the distribution dictated by Born's rule

$$\tilde{p}_i = \mathrm{Tr}\left(\rho \tilde{M}_i\right), \tag{5}$$

which we will dub a passive picture of noise. This runs counter to the more common view, where the noise is applied to the quantum state rather than the readout operator, hereafter referred to as the active picture of noise,

$$\tilde{p}_i = \mathrm{Tr}\left(\tilde{\rho} M_i\right), \tag{6}$$

where $\tilde{\rho} = \mathcal{E}(\rho)$ is the ideal quantum state passed through a noise channel[13]. These two views are equivalent, since we can only access the probabilities $\tilde{p}_i$. In the remainder of this work, we will be working on the passive noise picture.

Quantum state tomography. The objective of quantum state tomography is to reconstruct an arbitrary quantum state from only the measurement results. The most basic approach to QST can be framed as a linear inversion problem. To uniquely identify the quantum state, we use the decomposition provided by an IC POVM in eq. (4),

$$\rho = \sum_i a_i M_i, \tag{7}$$

where we have dropped the subscript "IC". The task of linear inversion is to determine the real coefficients a_i through the recorded measurement outcomes. Linear inversion is often considered bad practice as it can yield unphysical estimates[40]. It is advisable to use estimators such as the likelihood-based Bayesian mean estimator (BME)[40,41], or maximal likelihood estimator (MLE)[42], which relies directly on the likelihood function

$$\mathcal{L}_M(\rho) \propto \Pi_i \mathrm{Tr}(\rho M_i)^{n_i} = \left(\Pi_i \mathrm{Tr}(\rho M_i)^{\hat{p}_i}\right)^N, \tag{8}$$

which represents the likelihood that the prepared state of the system was ρ, given that outcome i has been observed n_i times, or, equivalently with outcome frequency $\hat{p}_i = \frac{n_i}{N}$. In our notation, the subscript M indicates what POVM was used for the measurement. The estimated state is the integrated mean

$$\rho_{\mathrm{BME}} \propto \int d\rho\, \rho \mathcal{L}_M(\rho) \tag{9}$$

for BME, and the maximizer

$$\rho_{\mathrm{MLE}} = \mathrm{argmax}_\rho\, \mathcal{L}_M(\rho) \tag{10}$$

for MLE.

Noise-mitigated quantum state tomography by detector tomography

In contrast to previous works, which view error mitigation as the correction of outcome statistics (see, e.g., ref. 28), our approach integrates the error mitigation procedure directly into the quantum state tomography framework. This comes with crucial advantages, as we will see, because it allows for the mitigation of a broader class of readout errors and intrinsically yield physical reconstruction without any additional corrective steps. We stress that this and similar protocols only work on mitigation of the statistical behavior of the quantum system, and is not suited for mitigation on the level of single-shot readout.

Quantum detector tomography. QDT[43] can be summarized as reconstructing the effects associated with a given measurement outcome, based on observed outcomes from a set of calibration states. Another, perhaps more instructive, view is that QDT finds a map between the ideal and actual POVM $\{M\} \to \{\tilde{M}\}$. Such a map could be used to extract noise parameters or learn the general behavior of the measurement device. However, our protocol avoids explicitly reconstructing such a map, and does not even require the knowledge of the ideal POVM $\{M\}$.

QDT starts out by preparing a complete set of calibration states, which span the space of all quantum states, and repeatedly performs measurements on these states. An example is the (over)complete set of Pauli density operators, $\left\{|0_i\rangle\langle 0_i|, |1_i\rangle\langle 1_i|\right\}$, where $i \in \{x, y, z\}$, which spans the space of single qubit operators. With the outcomes from all of these measurements, we can set up a linear set of equations given by Born's rule,

$$\frac{n_{is}}{N} = p_{is} = \mathrm{Tr}(\rho_s M_i), \tag{11}$$

where the subscript s iterates over the set of calibration states, e.g., $\{\rho_1 = |0_x\rangle\langle 0_x|, \rho_2 = |1_x\rangle\langle 1_x|, \dots\}$. Equation (11) represents a set of $I \times S$ constraints on $\{M_i\}$ we seek to reconstruct, where S is the total number of unique calibration states prepared and I is the number of POVM elements. In addition, one has the normalization constraint on the effects, $\sum_i M_i = \mathbb{1}$. Altogether, this yields a statistical estimation problem of a very similar nature to the one outlined in state tomography. It can be solved, for example, by using a maximal likelihood estimator, such as the one outlined in ref. 44, which guarantees physical POVM reconstructions.

Readout error-mitigated tomography. The core idea of our protocol is to use QDT as a calibration step before the state reconstruction. Using the information gathered from reconstructing the measurement effects, we modify the standard state estimator using the passive picture of noise, eq. (5). In this way, the estimator is "aware" of the noise and corresponding errors present in the measurement device.

A schematic overview of the protocol is given in Fig. 1. The first step is to reconstruct the POVM of the measurement device using detector tomography. This gives us access to an estimated noisy POVM $\left\{\tilde{M}^{\text{estm}}{}_i\right\}$. This noisy POVM is fed into the quantum state estimator, giving us the likelihood function

$$\mathcal{L}_{\tilde{M}^{\text{estm}}}(\rho) \propto \Pi_i \mathrm{Tr}\left(\rho \tilde{M}_i^{\text{estm}}\right)^{n_i} = \left(\Pi_i (\tilde{p}_i)^{\hat{p}_i}\right)^N, \tag{12}$$

where we have used that $\tilde{p}_i = \mathrm{Tr}(\rho \tilde{M}_i^{\text{estm}})$. This does indeed give us an estimator that converges to the noiseless state ρ, see Supplementary Note 1 for more information. We highlight that working in the passive noise picture comes with benefits not enjoyed by other estimators. Firstly, no inversion of a noise channel is required, and secondly, the reconstructed state is guaranteed to be physical. Everything is handled internally by the estimator.

Our protocol assumes the following experimental capabilities:

- Access to an IC POVM.
- Perfect state preparation.

The first capability is required by any full state reconstruction method, which can be done by any quantum device that has the ability to perform single qubit rotations and readout. By single qubit rotations, we mean any operation on a quantum device that can be decomposed into tensor products of single qubit rotations acting on each qubit. Similarly, for readout, one only needs tensor products of single qubit POVMs. This is also important for the noise mitigation to be general, since an unambiguous characterization of the erroneous state transformation inevitably requires an IC POVM. The second assumption is more problematic, as this is not strictly fulfilled in any real experiment. To circumvent this problem, one needs to consider state preparation and readout errors in a unified framework which is an active field of research. For this work it suffices to make sure that the state preparation errors are small compared to the readout errors. We discuss this further in subsection "Protocol limitations". For the explicit implementation of the protocol used in the remainder of the paper, see "Explicit protocol realization" in the Methods section.

Experimental results

We present the results of an experimental implementation of the readout error mitigation protocol on a superconducting qubit device. Important

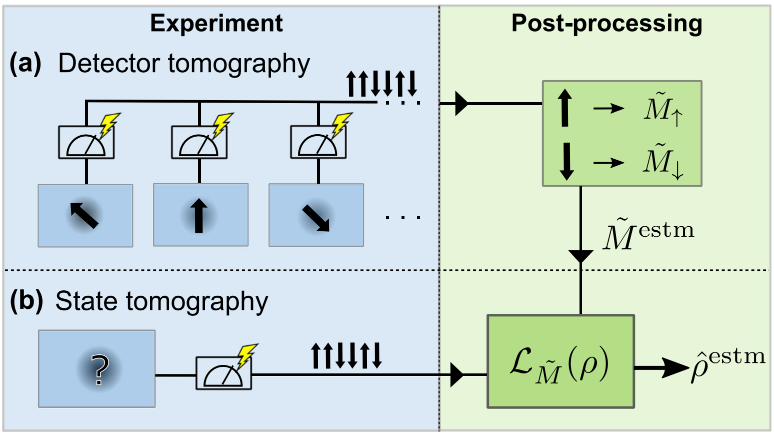


Fig. 1 | Protocol schematic overview. a Detector tomography: A complete set of basis states (e.g., the Pauli states) are prepared and measured repeatedly by the experiment. Based on the outcomes of the measurements, a positive operator-valued measure (POVM) is reconstructed. In a sense, it associates the measured outcome (here visualized as spin up/down measurements) with a measurement effect $\tilde{M}_i$. **b** State tomography: Using the reconstructed POVM, the modified likelihood function is endowed with knowledge of the operation of the measurement device. The system of interest is then prepared and measured repeatedly with the desired number of shots.

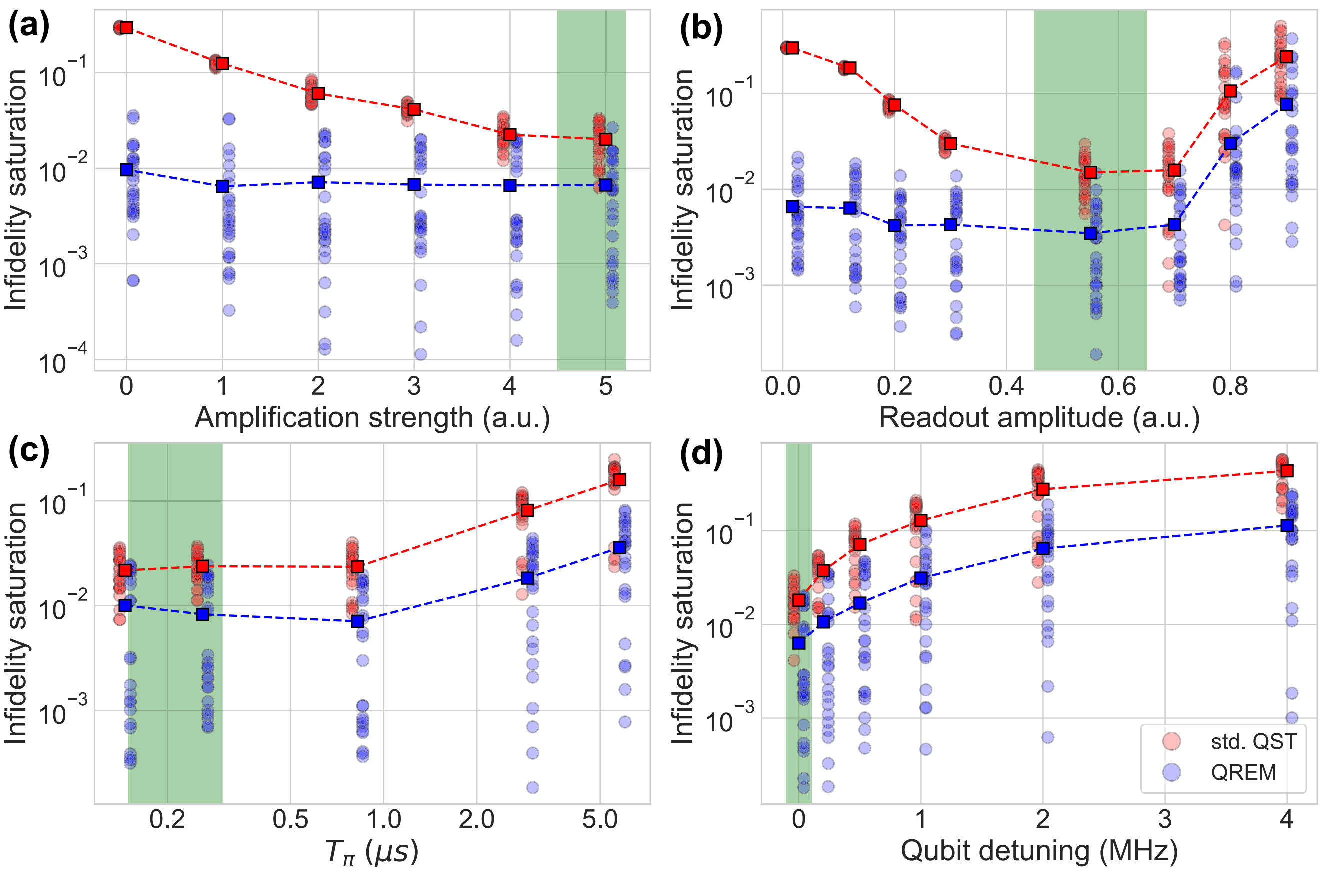


Fig. 2 | Characterizing error mitigation by detector tomography for quantum state tomography (QST) for different noise sources. Infidelity saturation refers to the last infidelity point measured over 240k single-shot measurements. The infidelity saturation for each individual run is plotted in translucent circles and shifted off center, to the left for standard QST and to the right for quantum readout error-mitigated (QREM) QST, for better visibility. The solid colored squares are the average infidelity saturation, connected by dotted lines for guidance. The green highlighted areas indicate the optimal experimental parameters. **a** Decreasing parametric amplification has a significant effect on QST through lesser distinguishability of the two states. Such errors can be mitigated very effectively by detector tomography. Zero amplification strength corresponds to having turned off the amplifier. **b** An incorrectly set readout amplitude of the resonator leads to increased infidelity in both the too weak and too strong readout regimes. Mitigation fails at higher powers, because higher levels are excited that are not taken into account. **c** Increasing manipulation timescales (T_π being the pulse time required for a bitflip) leads to more T_1 and T_2 decay events. This can be mitigated to some extent by the protocol, as seen by the smaller gradient of the mitigated infidelity curve. **d** Detuning between the qubit transition and drive frequencies can be efficiently mitigated. The mitigated infidelities rise proportionally to the unmitigated ones.

noise sources were varied to study their effect on state reconstruction accuracy. For more information on the induced noise sources and the experimental setup, see the Methods section. Quantum infidelity was used as a measure for state reconstruction accuracy. To make sure infidelity is a reliable performance metric, the infidelity is averaged over Haar-random target-states, shown as translucent circles in Fig. 2, see Methods for more information.

Insufficient readout amplification. We study the influence of insufficient signal amplification by the wideband traveling-wave parametric amplifier, in the following called amplifier. All other amplifiers operate with their optimal operational settings. We tune the amplification through the power of the microwave drive of the amplifier. This results in a reduced distinguishability of the two states. The measured saturation values of infidelity for different amplifications is given in Fig. 2a. Lowering the amplification reduces distinguishability in the IQ plane, resulting in high infidelities of around 0.3 with the amplifier turned off. The readout error mitigation protocol performs very well against this type of noise. When error mitigation is used, the reconstruction infidelity does not increase unless the amplifier is turned off completely, indicating that readout error mitigation is effective against this type of noise.

Readout resonator photon number. Ideally, superconducting qubits are read out with low resonator populations. Higher power readout will excite the qubit, potentially also to higher states outside the qubit manifold. Low readout power also enables non-demolition experiments, e.g., for active reset. If the readout power is too low, the IQ-plane state separation is not sufficient for single-shot readout, increasing the infidelity, an effect that our protocol can mitigate, as can be seen in Fig. 2b. This effectively constitutes a lower signal-to-noise ratio, similar to a low parametric amplification. At higher readout power, the higher transmon states excited by the strong readout are not captured in the two-level quantum system simulation, resulting in increased infidelities for both mitigated and unmitigated reconstruction. We can also see a higher spread of unmitigated infidelities when using readout amplitudes above and including 0.55. A possible explanation for this lies in the fact that the $0 \rightarrow 2$, $0 \rightarrow 1 \rightarrow 2$ and $1 \rightarrow 2$ processes have different frequencies, hence these transitions will be induced spontaneously by readout signals with a probability dependent on the current quantum state. This results in a state-dependence of higher state excitations, and therefore increased infidelity. We note that the infidelity ratio between standard QST and QREM becomes constant at large readout amplitudes.

In the following, we study noise sources that do not exclusively manifest as readout errors, as they also introduce significant errors in state preparation, which breaks one of the core assumptions of the protocol. It is nonetheless insightful to investigate the protocol's performance with such noise sources. Further experiments are required to determine the efficacy of the protocol under such conditions.

Shorter T_1 and T_2 times. Decoherence is modeled by exponential decay in fidelity in the z (T_1) and azimuthal (T_2) directions of the Bloch sphere. Thus, we can increase the number of decay events by increasing the length of manipulation pulses, used for state preparation and readout rotations. Experimentally, this is done by decreasing the manipulation power.

Figure 2c shows the results for varying the π-pulse lengths. At increased manipulation lengths, the qubit has more time to decay both through dephasing and energy loss, resulting in a larger error probability. Both unmitigated and mitigated state reconstruction gets progressively less accurate with increased manipulation lengths, with a noise-strength independent factor of 3–4 between them. The combined effect of decoherence on all three basis measurements is expected to be highly state-dependent, which can explain the large infidelity saturation spreads at a given noise strength.

Longer readout pulses lead to state decay already in the state preparation stages, hence the studied noise source is not only a readout error, but falls into the broader SPAM error category. One could restrict induced decay to purely affect readout, e.g. by only decreasing the manipulation power at the readout stage, but this would arguably not correspond to a realistic experimental scenario. Instead, consistently decreasing the manipulation power presents a more relevant experimental scenario.

Qubit detuning. If we consistently apply pulses detuned in frequency, we observe an apparent infidelity reduction by a factor of 4 by the protocol across all detunings, as shown in Fig. 2d. One can also see, that even a small detuning of 0.1 MHz results in lower infidelity than standard QST with perfect frequency-matching. We note that the infidelity saturation of QREM grows immediately with detuning, meaning that the protocol is not able to completely mitigate even small detunings. This can be attributed to the detector tomography step also suffering severely from effects of state preparation inaccuracies. As in the case of increased manipulation times, one could conduct an experiment where the detuning only affects the readout stage by applying resonant pulses for state preparation. As this also does not correspond to a realistic experimental scenario, the more relevant approach is to apply consistently detuned pulses.

At 4 MHz, detuning a pulse of 200 ns will result in an accumulated phase offset of 0.8, effectively scrambling σ_x and σ_y measurement outcomes. Both mitigated and conventional QST infidelity saturation depend roughly linearly on detuning, flattening out when approaching an infidelity of 0.5. The large spreads of the measured infidelity bounds stem from the fact that the effect of the noise depends on the state.

Noise that effectively manifests as a large frequency detuning can be induced, for example, by spontaneous iSWAP "operations" with neighboring two-level systems (TLS). Microwave sources may also suffer from frequency drifts. The errors coming from microwave source offsets are much smaller than the considered detunings. However, TLS-enabled jumps of more than 10 MHz have been observed on the same timescale as one run of QST takes for our experiment[45], making it a potential noise source corresponding to larger qubit detunings. It is important that these jumps do not occur after the device tomography stage, otherwise the reconstructed POVMs will not reflect the jump.

We remark that for all the shown cases where readout error mitigation does not work optimally, specifically in (c), (d), and the right part of (b) in Fig. 2, the ratio between standard QST and QREM infidelities is approximately constant. For noise sources that do not exclusively affect the readout process, a deterioration of the noise-mitigated reconstruction fidelity is expected as the noise strength increases. However, the observation of proportionality between the mitigated and unmitigated infidelities requires further investigation.

Combining noise sources. Several of the previously studied noise sources can be combined in different ways to simulate a noisy experiment. To demonstrate the efficiency of the mitigation ad extremum, we devise an experiment with comparably high levels of noise, by measuring with the off-resonant drive of $\delta\omega = 0.5$ MHz, decreased parametric amplification, and with decreased readout amplitude. While conventional QST saturates very early in the reconstruction, few-percent infidelities are possible by employing error mitigation, thereby opening the way for precise measurements with very noisy readouts, see Fig. 3. As expected (see subsection "Infidelity" in Methods), one can observe a linear dependence of infidelity on shot number on a log-log scale, corresponding to a power-law convergence.

Protocol limitations

We present an analysis of the potential limitations to the performance and reliability of the error mitigation protocol. Note that most of the discussed limitations will affect any estimator, and are not unique to our protocol. It is still important to understand their impact on protocol performance.

Sample fluctuations. A central limiting factor to the precision of any estimator are sample fluctuations, i.e., statistical fluctuations in the number of samples per effect. These fluctuations can manifest themselves in two parts of the protocol, QDT and QST, and cannot be mitigated.

QDT acts as a calibration step and is a one-time cost in terms of samples. Fluctuations in the POVM reconstruction can be viewed as a bias introduced into the state reconstruction. Therefore, it is important that the POVM reconstruction does not impose a bias larger than the expected sample fluctuation in the QST itself. In Fig. 4, we demonstrate that using less than 0.5% of experimental shots for detector tomography lowers the reconstruction infidelity to half its value. The lowest infidelity is achieved by using ca. 10% of shots for QDT, after which no meaningful further

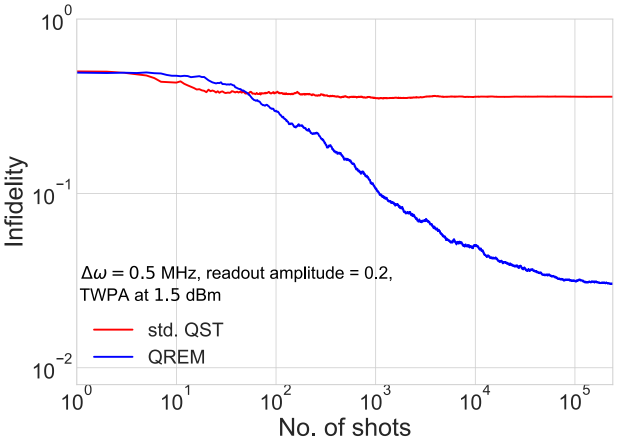


Fig. 3 | Combination of multiple strong noise sources. Conventional quantum state tomography (std. QST) saturates at around 100 shots, while using quantum readout error mitigation (QREM) enables state reconstruction with a factor of 30 lower infidelity. The infidelity is averaged over 25 Haar-random states, to see expected infidelity spread of each Haar-random state, see Supplementary Fig. 1. In this experiment, the readout amplitude was decreased compared to optimal, and the traveling-wave parametric amplifier (TWPA) was pumped with a decreased tone and the qubit was driven with a control pulse detuned in frequency by $\Delta\omega$.

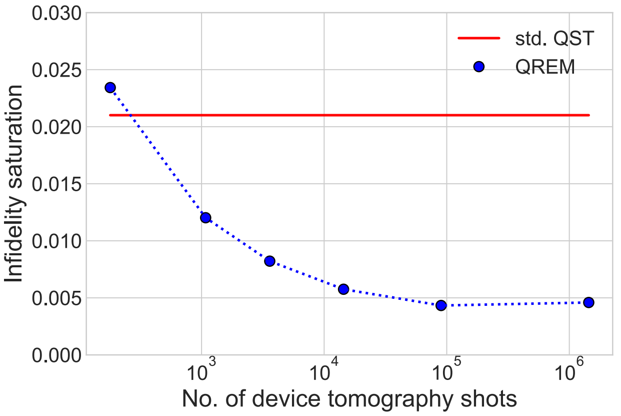


Fig. 4 | Averaged infidelity saturation in quantum state tomography (QST) as a function of calibration shots used in detector tomography. The quantum readout error-mitigated (QREM) measurements are compared to the baseline (std. QST). 240k shots were used for each QST reconstruction, averaged over 25 Haar-random states. Using 1k shots for calibration already results in a decrease in infidelity by a factor of ≈2. The curve flattens out at around 100k calibration shots.

improvement was seen in our experiment. We give the recommendation of using half the number of shots used for a single QST for QDT. When averaging over e.g., 25 quantum states for a representative benchmark, this becomes a small overhead.

Error-mitigated estimators, such as the one in eq. (12), aim to reduce the bias with respect to the prepared state caused by a noisy readout process. In reducing the bias, we are, in effect, taking into account that the state has passed through a noise channel. Since noise channels reduce the distinguishability between states, the variance of the modified likelihood function is increased with respect to the unmodified likelihood function, an effect known as the bias-variance trade-off[12,46]. In Fig. 5, we give an artistic rendition of the bias-variance trade-off under mitigation of single qubit depolarizing noise, given by the channel.

$$\mathcal{E}(\rho) = \frac{p}{2}\mathbb{1} + (1-p)\rho. \tag{13}$$

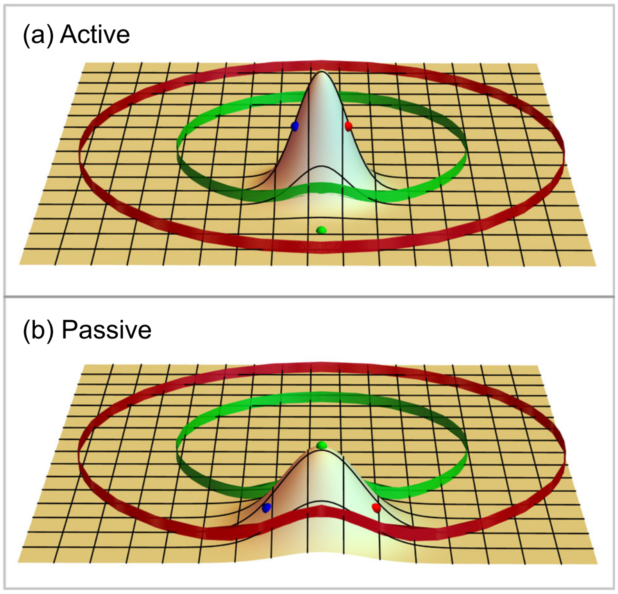


Fig. 5 | Sketch of the likelihood function on the Bloch disk. The Bloch disk represents an intersecting plane through the center of the Bloch sphere, e.g., the x-y-plane of the Bloch sphere. The red ribbon indicates the boundary of the Bloch disk, whereas the green ribbon indicates the boundary of the Bloch disk under the effect of a depolarizing channel. The green point is the true prepared state, which, due to the depolarizing readout noise, produces noisy measurement data. The red and blue points are arbitrary reference states in their active/passive noise representations. **a** An unmitigated likelihood function in the active representation of the noise (see eq. (6)). **b** An error-mitigated likelihood function in the passive representation of noise (see eq. (5)).

We compare (a) an unmitigated estimator in the active noise picture, to (b) the error-mitigated estimator in the passive noise picture. In the error-mitigated estimator the bias with respect to the true state (green point) is removed at a cost of higher variance of the likelihood function. We note that the bias due to the statistical fluctuations in the estimator remains, making it consistent with always-physical state estimators[47].

The increased variance can be combated by increasing the number of shots used for estimation. The additional number of shots required for the error-mitigated estimator to obtain the same confidence (variance of the likelihood) with respect to the true state as the unmitigated estimator with respect to the biased state is called the sampling overhead. The overhead relates to the distinguishability of quantum states and the data-processing inequality[46]. It tells us that the stronger the noise-induced distortion is, the larger the overhead. In QST, the sample overhead manifests itself as a shift in the infidelity curve. In Fig. 6 we have simulated varying depolarizing strengths and applied error mitigation. We see a clear shift as the noise strength increases. Note that such a shift would be present for any error mitigation schemes, including an inversion of the noise channel.

State preparation errors. A generic problem of QDT based methods is that they assume perfect preparation of calibration states, which is impossible in experiments. A common argument when using QDT is that the state preparation has a small error compared to the readout error itself. There are, however, methods developed to deal with the potential systematic error introduced by QDT, see e.g., refs. 41,48.

Another approach is to use the fact that preparation of the calibration states used in QDT only requires single qubit gates. If one has access to error estimates from single qubit gates, one could in principle also correct for state preparation errors in a similar manner to what is done in the protocol in this work (see e.g. randomized benchmarking[49] for gate characterization).

Despite our protocol requiring perfect state preparation, we have investigated noise sources that also affect state preparation, such as qubit

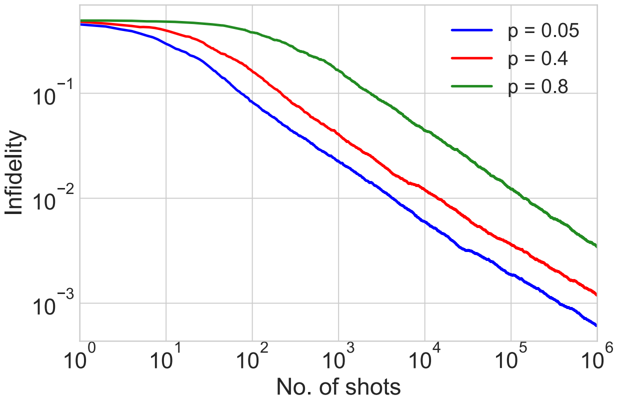


Fig. 6 | Reconstruction infidelity for mitigated simulated depolarizing noise. The various colored curves correspond to different noise strength p. The infidelity is averaged over 100 Haar-random pure states. The depolarizing channel is given in eq. (13).

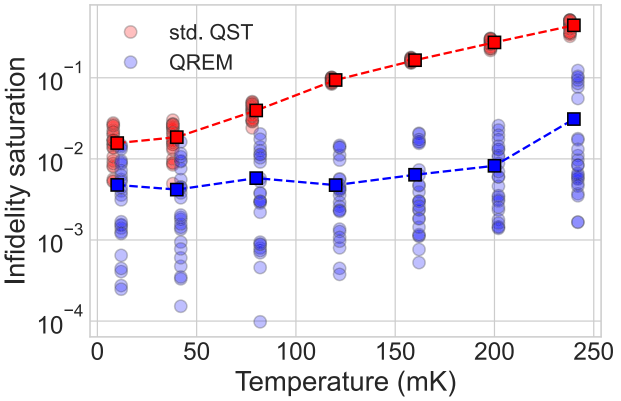


Fig. 7 | Quantum readout error mitigation (QREM) applied to increased qubit temperatures. Infidelity saturation refers to the last infidelity point measured over 240k single-shot measurements. The infidelity saturation for each individual run is plotted in translucent circles and shifted off center, to the left for standard quantum state tomography (std. QST) and to the right for QREM QST, for better visibility. The solid colored squares are the average infidelity saturation, connected by dotted lines for guidance. QREM seems to be successful, but the assumption of perfect state preparation is broken. Averaged over 25 Haar-random pure states with otherwise optimal experimental parameters.

detuning and increased T_1 and T_2 manipulation times. We will add to this investigation an experiment performed at higher qubit temperatures. We systematically increase the temperature of 10 mK up to over 200 mK, where the state in thermal equilibrium will have non-negligible contributions from the excited state. For example, a qubit temperature of 40 mK corresponds to an excited state population of 0.05 %, and a temperature of 120 mK corresponds to a 7.3 % excited state population. Hence, for our experiment without active feedback, it is not possible to reliably prepare a pure calibration state.

The results in Fig. 7 suggest that the mitigated QST is resilient to higher temperatures. The reason why we seemingly obtained a successful mitigation is that our benchmarking method does not allow us to distinguish state preparation errors from readout errors. Since our benchmarking method does not contain any additional gates between preparation and readout, the errors acquired in state preparation are interpreted as readout errors, and the protocol manages to mitigate the errors accordingly. We emphasize, however, that this is only strictly true if the combined effect of state preparation and readout errors can be viewed as a single effective error channel that is independent of the prepared state. While this is the case for finite qubit temperature, which can be modeled as depolarizing noise, it is not the case in general. In particular, for the cases of qubit detuning and increased T_1 and T_2 time, the putative effective noise channel becomes dependent on the prepared state, leading to a deterioration of the mitigation efficiency. We discuss this in more detail in Supplementary Note 2.

Experimental drift. After performing QDT, our protocol assumes that the POVM stays fixed for the remainder of the measurement sequence. However, since experimental parameters drift over time, the physically realized POVM may change with respect to the POVM used for state reconstruction. Due to this drift, one would ideally recalibrate the POVM before each reconstruction. This entails a relatively large overhead, and is not feasible.

The drift present in our experiment was small and we decided not to warrant any additional corrective measures. This is not necessarily the case in general. In these cases, we propose to add an additional step of periodically performing drift measurements and recalibration to the measurement protocol (see “Measurement protocol” in Methods for more information).

Drift measurements amount to measuring a set of well-known states, e.g. Pauli states, and reconstruct their density matrix. If the reconstruction infidelity goes beyond an acceptable threshold $I(\rho^{\text{estm}}, \rho) \geq \epsilon$, one can perform a recalibration of the measurement device, i.e. repeat the QDT step. This would ensure that the accuracy of the protocol does not degrade.

Conclusion

We have presented a comprehensive scheme for readout error mitigation in the framework of quantum state tomography. It introduces quantum detector tomography as an additional calibration step, with a small overhead cost in the number of experimental samples. After calibration, our method is able to mitigate any errors acquired at readout. Furthermore, it does not require the inversion of any error channels and guarantees that the final state estimate is physical. Comparing to most previously discussed QREM methods, our protocol is able to correct beyond-classical errors. To confirm that such errors indeed make up a significant part of the errors in our experiment, we present a selection of reconstructed POVM elements from the experiment in Supplementary Note 3. The significant off-diagonal contributions in this analysis confirm that non-classical errors are always present and often on the same order of magnitude as the diagonal classical redistribution errors.

To probe the limits of readout error mitigation, we applied our protocol to a superconducting qubit system. We experimentally subjected the qubit to several noise sources and investigated the protocol’s ability to mitigate them. We observed an improvement in the readout quality by decreasing infidelity by a factor of 5 to 30 depending on the type of readout noise. The protocol was particularly effective for lowered signal amplification and decreased resonator readout power compared to standard QST. We combined multiple noise sources in an experiment where conventional state reconstruction saturated early on, whereas our method was able to precisely reconstruct the quantum state. This opens up new possibilities for systems with noisy readouts where accurate knowledge of the quantum state is required. For noise sources which do not exclusively affect the readout stage, the protocol did not perform optimally, and we observed a constant ratio between the infidelity of mitigated and unmitigated state reconstruction.

Potential limitations of the scheme were investigated and we presented prescriptions on how to overcome them. Overall, we found that, by using readout error mitigation, one obtains accurate state estimates even under significantly degraded experimental conditions, making the readout more robust.

While being limited by exponential scaling in both memory and the required number of measurements, we expect the protocol to be feasible for up to 5–6 qubits if one replaces BME with MLE in the state reconstruction. This represents an interesting domain for error mitigation in quantum simulation as low-order correlators of a larger system offer relevant information such as correlation propagation[50] and phase estimation[51]. Recent developments in scalable approaches involving overlapping tomography[52,53] could provide a framework for a scalable version of this protocol to large qubit numbers, which we intend to

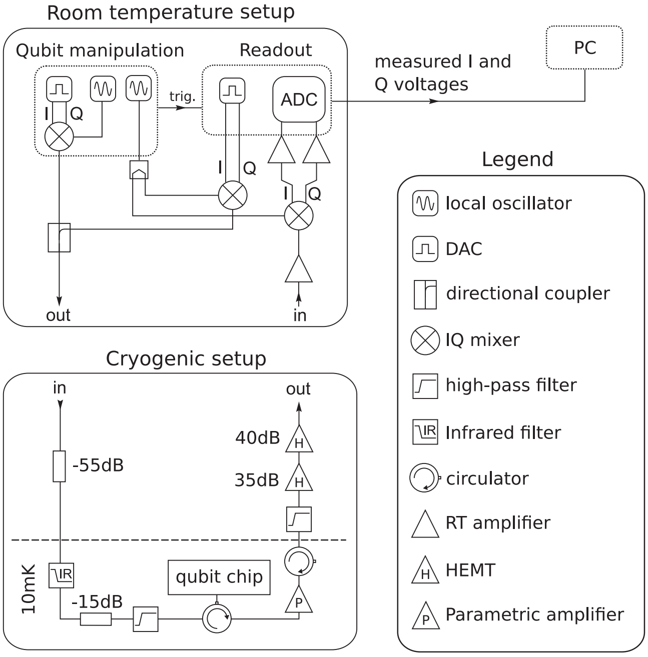


Fig. 8 | Microwave measurement setup. The local oscillator (LO) and phase-shifted pulsed IQ signals are mixed to provide the readout signal for the dispersive measurement of the qubit. A separate ω_{01} pulsed signal manipulates the qubit. The attenuators on the mixing chamber stage are carefully thermalized to mK temperatures to mitigate extra thermal noise. A wideband Josephson traveling-wave parametric amplifier (TWPA) and the two HEMT-based amplifiers provide quantum-limited readout.

explore. In future work, it could be interesting to perform a similar experiment on multiqubit systems. Implementing the protocol on a different qubit architecture with different sources of noise would be a topic of further interest. Another option of interest is to investigate adaptive noise-conscious strategies within this framework[54,55].

Methods

Experimental setup

We implemented the protocol introduced in the "Results and discussion" section on a fixed-frequency transmon qubit[56,57] with frequency $\omega_{01} = 6.3$ GHz coupled to a resonator with frequency $\omega_r = 8.5$ GHz in the dispersive readout scheme. Typical coherence times were observed to be $T_1 = 30 \pm 5\,\mu$s and $T_2 = 28 \pm 6\,\mu$s. For more information about different types of superconducting qubits, in particular their operation and noise sources affecting them, we refer to ref. 58.

A schematic diagram of the setup is shown in Fig. 8. All measurements without an explicitly stated temperature were performed at 10 mK, using a dry dilution refrigerator. At room temperature we use a time-domain setup with superheterodyne mixing of readout pulses triggered by a high-frequency arbitrary wave generator (AWG) responsible for local oscillator readout and pulsed manipulation tones. Signal lines are attenuated by 70 dB distributed over various stages, as seen in Fig. 8. The measured signals are converted into complex numbers in the so-called IQ plane, where I and Q abbreviate in-phase and quadrature components of the integrated microwave signals. These two signals can be transformed equivalently into amplitude and phase. With optimized experimental calibrations and no induced noises, we can distinguish between two non-overlapping Gaussian distributions corresponding to the two possible outcomes of the measurement, with the qubit state being projected to either ground state $|0\rangle$ or excited state $|1\rangle$. For examples of how these distributions change in the presence of noise, see Supplementary Fig. 3 in Supplementary Note 4. Unless otherwise stated, we perform measurements at the best separation settings for the classification model. This usually corresponds to approx. 98% distinguishability, meaning that 98% of shots can be correctly categorized as spin up or spin down.

Measurement protocol

or each noise source we study, we set up and execute our experiment as follows:

1. Estimation of π-pulse length (T_π) by Rabi-oscillations close to the qubit frequency.
2. Refined measurement of ω_{01} by a Ramsey experiment.
3. Measurement of T_π from a Rabi-measurement with updated ω_{01} from 2. by fitting a decaying sine function to the data.
4. Ground, $|0\rangle$, and excited states, $|1\rangle$, are measured with a single-shot readout. The location of the states in the IQ plane is learned by a supervised classification algorithm. For more information, see Supplementary Note 4.
5. QDT is performed by preparing each of the six Pauli states, and measuring them in the three bases σ_x, σ_y, and σ_z, which is described by the Pauli-6 POVM. The first two of these measurements are done as a combination of qubit rotation and subsequent σ_z readout. Quantum states are prepared using virtual Z-gates.
6. QST is performed and averaged over 25 random quantum states $U|0\rangle$, where U is a random unitary (Haar-random). Each state is measured in the three Pauli bases.

Performing this experiment for various strengths of the noise allows us to benchmark the ability of the protocol to mitigate the given noise source. For a schematic diagram of the measurement pipelines, see Fig. 9.

The outcome of two example experimental runs are shown in Fig. 10. Both standard and mitigated QST infidelities are extracted on a shot-by-shot basis. Typical features include a priori infidelity of roughly 0.5, with a rapid power-law decay, saturating at a given infidelity level for unmitigated QST. After this saturation is reached, further measurements will not improve the quantum state estimate because the measurements performed on the system are noisy. By using quantum readout error mitigation (labeled as QREM), we can significantly lower infidelity, enabling more precise state reconstruction. For all averaged experiments, the mitigated QST infidelities are consistently below the unmitigated QST infidelities.

Error sources

To probe the generality and reliability of our error-mitigation protocol, we artificially induce a set of noise sources, which introduce readout errors. In particular, we study:

1. Errors introduced by insufficient readout amplification: study through variation of the amplification of the parametric amplifier.
2. Errors introduced by low resonator photon number: study through variation of photon population in resonator by a variation of the readout amplitude.

In addition, we investigate two noise sources which manifest not only as readout errors, but also as state preparation errors. The implications of state preparation errors are discussed in "Protocol limitations" in Results and Discussion.

3. Errors introduced by energy (T_1) and phase relaxation (T_2): study through variation of the drive amplitude.
4. Errors introduced by qubit detuning: study through the variation of the applied manipulation pulse frequency.

Lastly, we combine multiple error sources to simulate a very noisy experiment, where conventional reconstruction methods fail.

Since the infidelity scaling is state-dependent, each experiment is averaged over 25 Haar-random states to get a reliable average performance (see "Explicit protocol realization" for more information). For QDT, we perform a total of $6 \times 3 \times 80{,}000$ (six states, three measurement bases) shots, for the following QST $25 \times 3 \times 80{,}000$ (25 states, three measurement bases, 80,000 shots each).

Infidelity

Quantum infidelity is well-suited as a figure of merit for successful quantum state reconstruction. We seek to minimize the infidelity, defined as

$$I(\rho,\sigma) = 1 - F(\rho,\sigma) = 1 - \left[\mathrm{Tr}\left(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right)\right]^2, \tag{14}$$

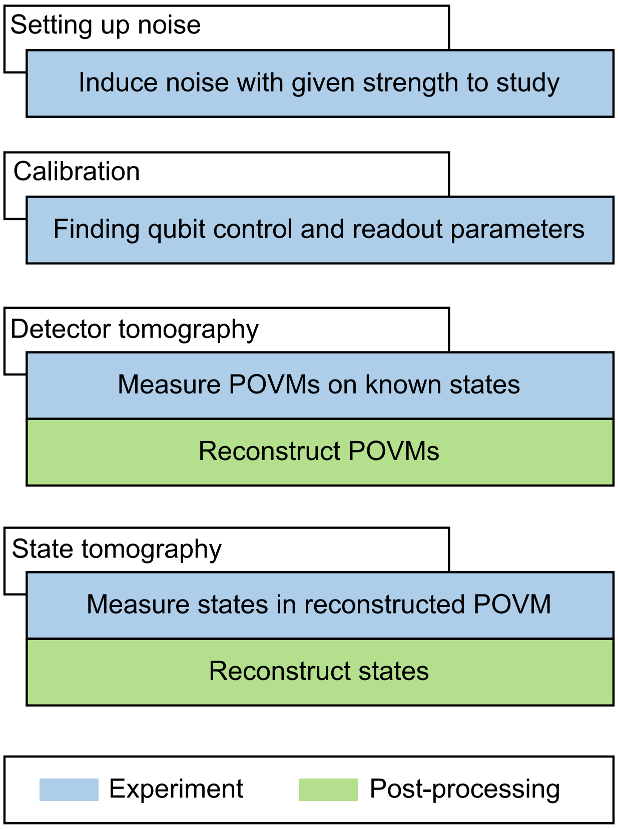


Fig. 9 | Schematic of the experimental pipeline of the protocol. In the first step, a given noise is induced with a specific strength. The experimental readout is calibrated for this noise. Detector tomography is performed, which reconstructs the noisy Pauli POVM $\{\tilde{M}_i\}$. Finally, quantum state tomography is executed, and reconstruction infidelity is evaluated and averaged over a set of randomly chosen target states.

where $F(\rho,\sigma)$ is the quantum fidelity[59]. We use the standard definition of $\sqrt{\rho}$, which is the square root of the eigenvalues of ρ in an eigendecomposition $\sqrt{\rho} = V\sqrt{D}V^{-1}$, where $\sqrt{D} = \mathrm{diag}\left(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n}\right)$. It acts as a pseudo-distance measure, and is asymptotically close to the Bures distance when $1 - F(\rho,\sigma) \ll 1$[60,61].

In this paper, we only consider the reconstruction of pure target states ρ, for which the infidelity simplifies to

$$I_{\mathrm{pure}}(\rho,\sigma) = 1 - \mathrm{Tr}(\rho\sigma). \tag{15}$$

The infidelity between a sampled target state ρ and the reconstructed state σ is expected to decrease with $I(\rho,\sigma) \propto N^{-\alpha}$ where N is the number of shots performed, and α is an asymptotic scaling coefficient that depends on features of the estimation problem[55,62,63], such as the deviation from measuring in the target states eigenbasis, and the purity of the target state.

Explicit protocol realization

We present the explicit implementation of the protocol described in the Results and discussion. A pseudo-code outline of the whole measurement and readout error mitigation is presented in Supplementary Algorithm 1. For QDT, we use the maximum likelihood estimator described in ref. 44. We follow the prescription described in the subsection "Quantum Detector Tomography" in Results and discussion, and use all of the Pauli states as calibration states. The number of times each Pauli state is measured equals the maximal number of shots used for a single spin measurement in the state reconstruction, such that the dominant source of shot noise is not QDT. For QST, we use a BME, in particular we use the implementation described in ref. 64. The bank particles are generated from the Hilbert-Schmidt measure[65]. This QST is also equipped with adaptive measurement strategies, which is a possible future extension to the current protocol.

For both QDT and QST, we use the Pauli-6 POVM, which is a static measurement strategy. It is known that the asymptotic scaling of such strategies depends on the proximity of the true state to one of the projective measurements[66,67]. To counteract this, an average over Haar-random states[68] is performed to get a robust performance estimate. In this way, we get the expected performance given that no prior information about the true state is available. We emphasize that for all random state reconstructions, the same QDT calibration is used.

Data availability

The data used for Figs. 4–7, 9, 10 and Supplementary Fig. 1 can be found on the KIT research data repository: https://radar.kit.edu/radar/en/dataset/RNbuograoVUFQNBB. For more information on the dataset, see

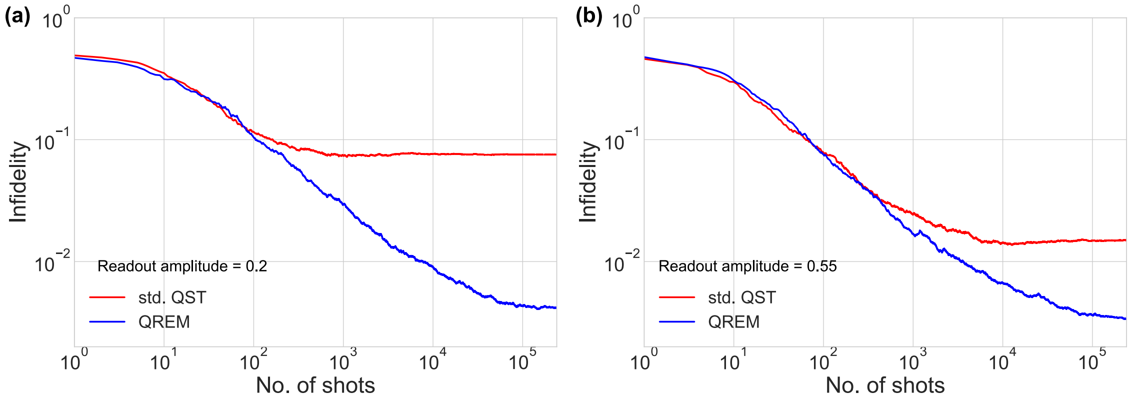


Fig. 10 | Example run of quantum state reconstruction using a Bayesian mean estimator. Mean quantum reconstruction infidelity is plotted as a function of number of shots for **a** optimal readout powers and **b** weak readout. The red lines represent unmitigated, standard quantum state tomography (std. QST), and the blue lines are obtained with quantum readout error mitigation (QREM). Each curve is an average of over 25 Haar-random states. For an indication of the spread of each Haar-random state, see Supplementary Fig. 1. QREM reaches a lower infidelity saturation value than standard reconstruction in both cases. At low readout power, error-mitigated QST can still reconstruct the state with similar accuracy, while standard QST saturates at a significantly higher value.

Supplementary Note 5. The remaining experimental data were available upon request from A. Di Giovanni.

Code availability

The code developed for this project is available on GitHub: https://github.com/AdrianAasen/EMQST. A short tutorial notebook is provided with examples of how to run the software. It can be interfaced with an experiment, or run as a simulation.

Received: 14 February 2024; Accepted: 22 August 2024;

References

1. Bennett, C. H. & Brassard, G. Quantum cryptography: public key distribution and coin tossing. *Theor. Comput. Sci.* **560**, 7–11 (2014).
2. Shor, P. W. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.* **26**, 1484–1509 (1997).
3. Ma, H., Govoni, M. & Galli, G. Quantum simulations of materials on near-term quantum computers. *npj Comput. Mater.* https://doi.org/10.1038/s41524-020-00353-z (2020).
4. Pyrkov, A. et al. Quantum computing for near-term applications in generative chemistry and drug discovery. *Drug Discov. Today* **28**, 103675 (2023).
5. Zinner, M. et al. Toward the institutionalization of quantum computing in pharmaceutical research. *Drug Discov. Today* **27**, 378–383 (2022).
6. Orús, R., Mugel, S. & Lizaso, E. Quantum computing for finance: overview and prospects. *Rev. Phys.* **4**, 100028 (2019).
7. Egger, D. J., Gutierrez, R. G., Mestre, J. C. & Woerner, S. Credit risk analysis using quantum computers. *IEEE Trans. Comput.* **70**, 2136–2145 (2021).
8. Dri, E., Giusto, E., Aita, A. & Montrucchio, B. Towards practical quantum credit risk analysis. *J. Phys. Conf. Ser.* **2416**, 012002 (2022).
9. Harwood, S. et al. Formulating and solving routing problems on quantum computers. *IEEE Trans. Quantum Eng.* **2**, 1–17 (2021).
10. Preskill, J. Quantum computing in the NISQ era and beyond. *Quantum* **2**, 79 (2018).
11. Lisenfeld, J., Bilmes, A. & Ustinov, A. V. Enhancing the coherence of superconducting quantum bits with electric fields. *npj Quantum Inform* https://doi.org/10.1038/s41534-023-00678-9 (2023).
12. Cai, Z. et al. Quantum error mitigation. *Rev. Mod. Phys.* **95**, 045005 (2023).
13. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge Univ. Press, 2012).
14. Krinner, S. et al. Realizing repeated quantum error correction in a distance-three surface code. *Nature* **605**, 669–674 (2022).
15. Kitaev, A. Fault-tolerant quantum computation by anyons. *Ann. Phys.* **303**, 2–30 (2003).
16. Dennis, E., Kitaev, A., Landahl, A. & Preskill, J. Topological quantum memory. *J. Math. Phys.* **43**, 4452–4505 (2002).
17. Raussendorf, R. & Harrington, J. Fault-tolerant quantum computation with high threshold in two dimensions. *Phys. Rev. Lett.* https://doi.org/10.1103/physrevlett.98.190504 (2007).
18. Greenbaum, D. Introduction to quantum gate set tomography. Preprint at arXiv https://arxiv.org/abs/1509.02921 (2015).
19. Geller, M. R. & Sun, M. Toward efficient correction of multiqubit measurement errors: pair correlation method. *Quantum Sci. Technol.* **6**, 025009 (2021).
20. Walter, T. et al. Rapid high-fidelity single-shot dispersive readout of superconducting qubits. *Phys. Rev. Appl.* https://doi.org/10.1103/physrevapplied.7.054020 (2017).
21. Qin, D., Xu, X. & Li, Y. An overview of quantum error mitigation formulas. *Chin. Phys. B* **31**, 090306 (2022).
22. Endo, S., Cai, Z., Benjamin, S. C. & Yuan, X. Hybrid quantum-classical algorithms and quantum error mitigation. *J. Phys. Soc. Jpn.* **90**, 032001 (2021).
23. Nachman, B., Urbanek, M., de Jong, W. A. & Bauer, C. W. Unfolding quantum computer readout noise. *npj Quantum Inform.* https://doi.org/10.1038/s41534-020-00309-7 (2020).
24. Pokharel, B., Srinivasan, S., Quiroz, G. & Boots, B. Scalable measurement error mitigation via iterative bayesian unfolding. *Phys. Rev. Research* **6**, 013187 (2024).
25. Geller, M. R. Rigorous measurement error correction. *Quantum Sci. Technol.* **5**, 03LT01 (2020).
26. Bravyi, S., Sheldon, S., Kandala, A., Mckay, D. C. & Gambetta, J. M. Mitigating measurement errors in multiqubit experiments. *Phys. Rev.* https://doi.org/10.1103/physreva.103.042605 (2021).
27. Kwon, H. & Bae, J. A hybrid quantum-classical approach to mitigating measurement errors in quantum algorithms. *IEEE Trans. Comput.* **70**, 1401–1411 (2021).
28. Maciejewski, F. B., Zimborás, Z. & Oszmaniec, M. Mitigation of readout noise in near-term quantum devices by classical post-processing based on detector tomography. *Quantum* **4**, 257 (2020).
29. Kim, Y. et al. Evidence for the utility of quantum computing before fault tolerance. *Nature* **618**, 500–505 (2023).
30. Temme, K., Bravyi, S. & Gambetta, J. M. Error mitigation for short-depth quantum circuits. *Phys. Revi. Lett.* https://doi.org/10.1103/physrevlett.119.180509 (2017).
31. Li, Y. & Benjamin, S. C. Efficient variational quantum simulator incorporating active error minimization. *Phys. Rev. X* https://doi.org/10.1103/physrevx.7.021050 (2017).
32. Chen, Y., Farahzad, M., Yoo, S. & Wei, T.-C. Detector tomography on IBM quantum computers and mitigation of an imperfect measurement. *Phys. Rev. A* https://doi.org/10.1103/physreva.100.052315 (2019).
33. Smithey, D. T., Beck, M., Raymer, M. G. & Faridani, A. Measurement of the wigner distribution and the density matrix of a light mode using optical homodyne tomography: application to squeezed states and the vacuum. *Phys. Rev. Lett.* **70**, 1244–1247 (1993).
34. xi Liu, Y., Wei, L. F. & Nori, F. Tomographic measurements on superconducting qubit states. *Phys. Rev. B* https://doi.org/10.1103/physrevb.72.014547 (2005).
35. Motka, L., Paúr, M., R^eháček, J., Hradil, Z. & Sánchez-Soto, L. L. Efficient tomography with unknown detectors. *Quantum Sci. Technol.* **2**, 035003 (2017).
36. Motka, L., Paúr, M., Řeháček, J., Hradil, Z. & Sánchez-Soto, L. L. When quantum state tomography benefits from willful ignorance. *N. J. Phys.* **23**, 073033 (2021).
37. Ramadhani, S., Rehman, J. U. & Shin, H. Quantum error mitigation for quantum state tomography. *IEEE Access* **9**, 107955–107964 (2021).
38. Sakurai, J. J. & Napolitano, J. *Modern Quantum Mechanics* (Cambridge Univ. Press, 2017).
39. Paris, M. & Reháček, J. (eds.) *Quantum State Estimation* (Springer, 2004).
40. Blume-Kohout, R. Optimal, reliable estimation of quantum states. *N. J. Phys.* **12**, 043034 (2010).
41. Gebhart, V. et al. Learning quantum systems. *Nat. Rev. Phys.* **5**, 141–156 (2023).
42. Lvovsky, A. I. Iterative maximum-likelihood reconstruction in quantum homodyne tomography. *J. Opt. B Quantum Semiclassical Opt.* **6**, S556–S559 (2004).
43. Lundeen, J. S. et al. Tomography of quantum detectors. *Nat. Phys.* **5**, 27–30 (2008).
44. Fiurášek, J. Maximum-likelihood estimation of quantum measurement. *Phys. Rev. A* https://doi.org/10.1103/physreva.64.024102 (2001).
45. Meißner, S. M., Seiler, A., Lisenfeld, J., Ustinov, A. V. & Weiss, G. Probing individual tunneling fluctuators with coherently controlled tunneling systems. *Phys. Rev. B* https://doi.org/10.1103/physrevb.97.180505 (2018).

46. Takagi, R., Endo, S., Minagawa, S. & Gu, M. Fundamental limits of quantum error mitigation. *npj Quantum Inform* https://doi.org/10.1038/s41534-022-00618-z (2022).
47. Schwemmer, C. et al. Systematic errors in current quantum state tomography tools. *Phys. Rev. Lett.* https://doi.org/10.1103/PhysRevLett.114.080403 (2015).
48. Zhang, A. et al. Experimental self-characterization of quantum measurements. *Phys. Rev. Lett.* https://doi.org/10.1103/physrevlett.124.040402 (2020).
49. Knill, E. et al. Randomized benchmarking of quantum gates. *Phys. Rev. A* https://doi.org/10.1103/physreva.77.012307 (2008).
50. Richerme, P. et al. Non-local propagation of correlations in quantum systems with long-range interactions. *Nature* **511**, 198–201 (2014).
51. Ebadi, S. et al. Quantum phases of matter on a 256-atom programmable quantum simulator. *Nature* **595**, 227–232 (2021).
52. Cotler, J. & Wilczek, F. Quantum overlapping tomography. *Phys. Rev. Lett.* https://doi.org/10.1103/PhysRevLett.124.100401 (2020).
53. Tuziemski, J. et al. Efficient reconstruction, benchmarking and validation of cross-talk models in readout noise in near-term quantum devices. Preprint at https://arxiv.org/abs/2311.10661 (2023).
54. Ivanova-Rohling, V. N., Rohling, N. & Burkard, G. Optimal quantum state tomography with noisy gates. *EPJ Quantum Technol.* https://doi.org/10.1140/epjqt/s40507-023-00181-2 (2023).
55. Huszár, F. & Houlsby, N. M. T. Adaptive Bayesian quantum tomography. *Phys. Rev. A* https://doi.org/10.1103/physreva.85.052120 (2012).
56. Koch, J. et al. Charge-insensitive qubit design derived from the cooper pair box. *Phys. Rev. A* https://doi.org/10.1103/physreva.76.042319, https://doi.org/10.1103/physreva.76.042319 (2007).
57. Kjaergaard, M. et al. Superconducting qubits: current state of play. *Ann. Rev. Condens. Matter Phys.* **11**, 369–395 (2020).
58. Krantz, P. et al. A quantum engineer's guide to superconducting qubits. *Appl. Phys. Rev.* https://doi.org/10.1063/1.5089550 (2019).
59. Jozsa, R. Fidelity for mixed quantum states. *J. Modern Opt.* **41**, 2315–2323 (1994).
60. Hübner, M. Explicit computation of the bures distance for density matrices. *Phys. Lett. A* **163**, 239–242 (1992).
61. Uhlmann, A. The "transition probability" in the state space of a *-algebra. *Rep. Math. Phys.* **9**, 273–279 (1976).
62. Massar, S. & Popescu, S. Optimal extraction of information from finite quantum ensembles. *Phys. Rev. Lett.* **74**, 1259–1263 (1995).
63. Bagan, E., Ballester, M. A., Gill, R. D., Monras, A. & Muñoz-Tapia, R. Optimal full estimation of qubit mixed states. *Phys. Rev. A* https://doi.org/10.1103/physreva.73.032301 (2006).
64. Struchalin, G. I. et al. Experimental adaptive quantum tomography of two-qubit states. *Phys. Rev. A* https://doi.org/10.1103/physreva.93.012103 (2016).
65. Zyczkowski, K. & Sommers, H.-J. Induced measures in the space of mixed quantum states. *J. Phys. A Math. Gen.* **34**, 7111–7125 (2001).
66. Bagan, E., Ballester, M. A., Gill, R. D., Muñoz-Tapia, R. & Romero-Isart, O. Separable measurement estimation of density matrices and its fidelity gap with collective protocols. *Phys. Rev. Lett.* https://doi.org/10.1103/physrevlett.97.130501 (2006).
67. Struchalin, G. I., Kovlakov, E. V., Straupe, S. S. & Kulik, S. P. Adaptive quantum tomography of high-dimensional bipartite systems. *Phys. Rev. A* https://doi.org/10.1103/physreva.98.032330 (2018).
68. Mezzadri, F. How to generate random matrices from the classical compact groups. *Notices of the American Mathematical Society, vol. 54, no. 5, pp. 592 - 604*. Preprint at https://arxiv.org/abs/math-ph/0609050 (2007).

Acknowledgements

The authors are grateful for the quantum circuit provided by D. Pappas, M. Sandberg, and M. Vissers. We thank W. Oliver and G. Calusine for providing the parametric amplifier. This work was partially financed by the Baden-Württemberg Stiftung gGmbH. The authors acknowledge support by the state of Baden-Württemberg through bwHPC and the German Research Foundation (DFG) through Grant No INST 40/575-1 FUGG (JUSTUS 2 cluster).

Author contributions

Development of the protocol and software was done by A. Aasen with supervision from M. Gärttner. The experimental realization and data generation was done by A. Di Giovanni with supervision from H. Rotzinger and A. Ustinov. A. Aasen and A. Di Giovanni prepared the draft for the manuscript. All authors contributed to the finalization of the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42005-024-01790-8.

Correspondence and requests for materials should be addressed to Martin Gärttner.

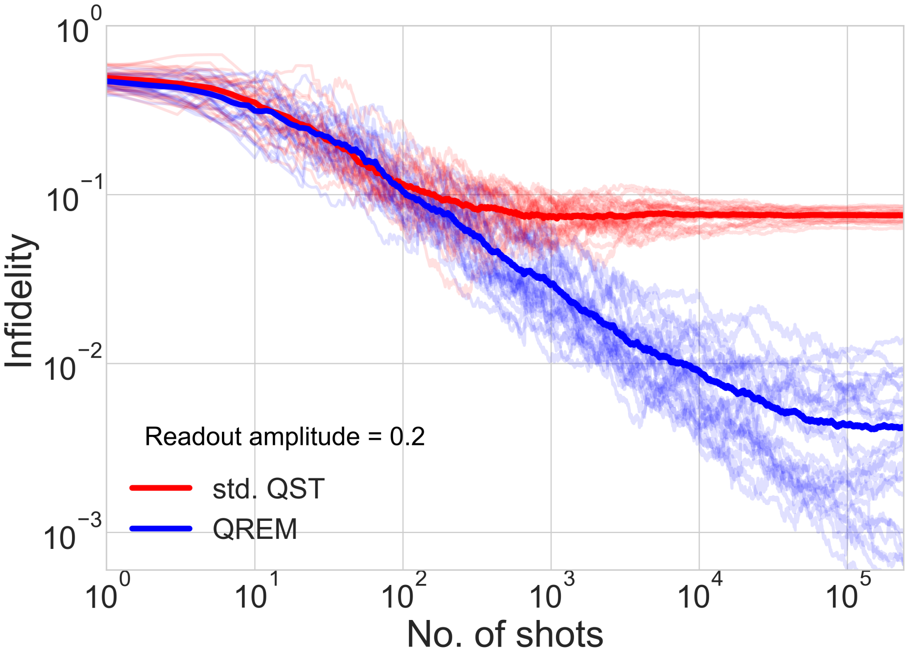
Peer review information *Communications Physics* thanks Jacob Blumoff and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Supplemental Information:
Readout error mitigated quantum state tomography tested on superconducting qubits

Adrian Skasberg Aasen,[1,2,*] Andras Di Giovanni,[3,*] Hannes Rotzinger,[3,4] Alexey V. Ustinov,[3,4] and Martin Gärttner[2,†]

[1] *Kirchhoff-Institut für Physik, Universität Heidelberg,*
Im Neuenheimer Feld 227, 69120 Heidelberg, Germany
[2] *Institut für Festkörpertheorie und -optik, Friedrich-Schiller-Universität Jena, Max-Wien-Platz 1, 07743 Jena, Germany*
[3] *Physikalisches Institut, Karlsruher Institut für Technologie, Kaiserstraße 12, 76131 Karlsruhe, Germany*
[4] *Institut für QuantenMaterialien und Technologien,*
Karlsruher Institut für Technologie, Kaiserstraße 12, 76131 Karlsruhe, Germany

(Dated: August 16, 2024)



Supplementary Figure 1. Replotted Fig. 10 **(a)** from the main text with all averaged curves plotted to show spread. The blue curves are readout error mitigated (QREM) and the red curves are standard quantum state tomography (std. QST). Each of the 25 Haar-random state reconstruction infidelities are plotted as translucent curves of the respective colors, while the solid colored curves are the average.

SUPPLEMENTARY NOTE 1: CONVERGENCE OF MODIFIED ESTIMATOR

Here we show that the protocol estimator, eq. (12) in main text, converges to the noiseless state with the modifications presented. By convergence of the estimator, we mean that the expected state converges to the true state for all quantum states, in the limit of infinite statistics. To make this explicit, we show that an estimator with the standard likelihood function, eq. (8) in the main text, converges to any given true state ρ_T. The generalization to noisy measurements follows immediately.

For simplicity, we will assume that the measurement device used implements an IC and minimal POVM. By extension, any quantum state is uniquely defined by the probabilities $p_i = \mathrm{Tr}(\rho M_i) = \langle M_i \rangle$. The natural estimator for

---

[*] These authors contributed equally to this work.
[†] martin.gaerttner@uni-jena.de

the probabilities is the frequency of each outcome,

$$\hat{p}_i = \frac{n_i}{N}. \tag{S1}$$

All we need to show is that using this estimator for the probabilities maximizes the likelihood function, eq. (8) in the main text. To be explicit, we need to show that $\hat{p}_i = \frac{n_i}{N}$ converges to the probability $p_i = \mathrm{Tr}(\rho M_i)$. This can be done by maximizing the log-likelihood function $\log(\mathcal{L}_M(\rho))$ with Lagrange multipliers. This yields that $p_i = \hat{p}_i$ maximizes the likelihood function. Therefore, if the experiment implements the POVM $\{M_i\}$, which means it samples from the probabilities $p_i = \mathrm{Tr}(\rho_T M_i)$, it is clear that we have a reconstruction $\hat{\rho}$ that converges to ρ_T.

Consider the case where the measurement device is imperfect and implements the noisy POVM $\{\tilde{M}_i\}$, and we have the reconstructed POVM from the QDT $\{\tilde{M}_i^{\text{estm}}\}$. Analogous to the noiseless case, $\hat{\tilde{p}}_i = \frac{\tilde{n}_i}{N}$ converges to $\tilde{p}_i = \mathrm{Tr}\big(\rho_T \tilde{M}_i\big)$, as long as $\langle \tilde{M}_i^{\text{estm}} \rangle$ converges to $\tilde{M}_i$. This is guaranteed by using the MLE reconstruction outlined in Ref. [1].

SUPPLEMENTARY NOTE 2: STATE PREPARATION ERRORS AS EFFECTIVE STATE-DEPENDENT READOUT ERROR CHANNELS

We present a straightforward model demonstrating how unitary gate errors can create an effective state-dependent error channel during state preparation, and how this influences the readout error mitigation protocol. By state-dependent, we mean that the effective error channel introduced during state preparation depends on specific properties of the prepared state. Our examination will be restricted to the case of a single qubit.

Consider a channel that describes the preparation of basis state ρ_S from the ground state $|0\rangle$.

$$\mathcal{E}_S(|0\rangle\langle 0|) = \rho_S. \tag{S2}$$

In the operator-sum representation a general channel can be represented as [2]

$$\mathcal{E}(\rho) = \sum_i K_i \rho K_i^\dagger, \tag{S3}$$

where $\sum K_i^\dagger K_i = \mathbb{1}$. When the state preparation is noiseless, the operator-sum representation is particularly simple

$$\mathcal{E}_{S,\text{ideal}}(\rho) = U_S \rho U_S^\dagger, \tag{S4}$$

where U_S unitary that implements the transformation $|0\rangle\langle 0| \to \rho_S$. We can decompose any single-qubit unitary [2]

$$U_S = R_Z(\theta_2) R_Y(\theta_1) R_Z(\theta_0), \tag{S5}$$

where $R_\alpha(\theta) = e^{-i\theta\sigma_\alpha/2}$ are the single qubit rotation operators about the Cartesian axis α of the Bloch sphere. We will now consider the channel $\mathcal{E}_S$ where with probability p_i implements a unitary $\tilde{U}_{S,i}$ which deviates from the ideal U_S. We want to show that $\mathcal{E}_S = \mathcal{E}_{S,\text{error}} \circ \mathcal{E}_{S,\text{ideal}}$.

Proposition 1. *The state preparation channel can be separated into $\mathcal{E}_S = \mathcal{E}_{S,error} \circ \mathcal{E}_{S,ideal}$, where $\mathcal{E}_{S,error}$ generally depends on the state S being prepared.*

Proof. Our strategy is to show that each $\tilde{U}_{S,i} = \tilde{U}_{S,i,\text{error}} U_S$, where $\tilde{U}_{S,i,\text{error}}$ is some unitary, that may or may not depend on S. From such a decomposition of each i it immediately follows that the channels also decompose. We start by expressing $\tilde{U}_{S,i}$ in the decomposition in eq. (S5) (where we drop the index i for convenience)

$$\tilde{U}_S = R_Z(\tilde{\theta}_2) R_Y(\tilde{\theta}_1) R_Z(\tilde{\theta}_0). \tag{S6}$$

This channel will always be applied to the ground state $|0\rangle$, so we can ignore $R_Z(\tilde{\theta}_0)$ as it does not affect the ground state. The remaining angles can be expressed in terms of the ideal ones as

$$\tilde{\theta}_k = \theta_k + \Delta_k. \tag{S7}$$

Since the rotation gate commutes with itself, we can split them

$$R_\alpha(\tilde{\theta}_k) = R_\alpha(\theta_k) R_\alpha(\Delta_k) = R_\alpha(\Delta_k) R_\alpha(\theta_k). \tag{S8}$$

We can then write the the noisy unitary as

$$\begin{aligned}\tilde{U}_S =&R_Z(\tilde{\theta}_2)R_Y(\tilde{\theta}_1)\\ =&R_Z(\Delta_2)R_Z(\theta_2)R_Y(\Delta_1)R_Y(\theta_1)\\ =&\underbrace{R_Z(\Delta_2)R_Z(\theta_2)R_Y(\Delta_1)R_Z^\dagger(\theta_2)}_{U_{S,\text{error}}}\underbrace{R_Z(\theta_2)R_Y(\theta_1)}_{U_S}\\ =&U_{S,\text{error}}U_S.\end{aligned} \tag{S9}$$

The error unitary generally depends on the rotation that was performed through θ_2.

Consider now the channel $\mathcal{E}_S$ where we with some probability p_i implement the unitary $\tilde{U}_{S,i}$,

$$\mathcal{E}_S(\rho) = \sum_i p_i \tilde{U}_{S,i}\rho\tilde{U}^\dagger_{S,i} \tag{S10}$$

Inserting the noisy unitary from eq. (S9) into the channel we have

$$\begin{aligned}\mathcal{E}_S(\rho) =& \sum_i p_i U_{S,\text{error}}\overbrace{U_S\rho U_S^\dagger}^{\mathcal{E}_{S,\text{ideal}}(\rho)}U^\dagger_{S,\text{error}}\\ =&\sum_i p_i U_{S,\text{error}}\mathcal{E}_{S,\text{ideal}}(\rho)U^\dagger_{S,\text{error}}\\ =&\mathcal{E}_{S,\text{error}}\circ\mathcal{E}_{S,\text{ideal}}(\rho)\end{aligned} \tag{S11}$$

and we have showed that the noise channels separate. □

This separation of the state preparation channel means that $\mathcal{E}_{S,\text{error}}$ will be interpreted as part of the readout errors in the passive picture. Importantly, we see the explicit dependence of the error channel on the prepared state and the distribution p_i. With a state-dependent error channel present, one cannot reconstruct a consistent set of POVM elements based on the $I \times S$ constraints, since each set of constrains per calibration state (S) comes from a different channel. This means, that the POVM reconstructed in the detector tomography stage, which is based on a finite set of calibration states, will in general not faithfully represent the error incurred for the calibration states themselves or any other prepared states.

With this simple and experimentally motivated model we can provide an explanation of why the protocol still works for some cases, while in others it does not. Note that the operator-sum representation is not unique, so it could still be possible to find representations where the explicit dependence vanishes, see e.g. increased temperature.

Qubit detuning

Qubit detuning has non-trivial contributions from both Δ_1 and Δ_2. We therefore expect there to be state-dependent errors present in state preparation, and subsequently our readout error mitigation protocol to under-perform.

Increased temperature

The primary source of error at elevated temperatures stems from the excitation of the qubit ground state. A secondary contribution arises from unitary gate errors, with finite temperatures interpreted as an error symmetrically distributed over fluctuating Δ_1 and Δ_2. The symmetric convex combination of these unitary rotations results in a depolarizing error channel that does not depend on the prepared state, thereby enabling us to transfer these state-preparation errors to the readout stage as well.

Shorter T_1 and T_2 times

Increased manipulation times cannot easily be viewed as effective unitary gate errors. However, it is plausible that also non-unitary errors can lead to a state-dependent effective error channel. The protocol's under-performance during extended manipulation times might suggest that state-dependent errors are indeed induced.

SUPPLEMENTARY NOTE 3: INSUFFICIENCY OF CLASSICAL ERROR MITIGATION

To demonstrate that classical errors are not sufficient to capture all error present in the measurement device, we investigate coherent errors in the POVM elements. Classical errors assume that the reconstructed POVM elements are purely diagonal in their measurement basis, i.e. only contains a statistical redistribution of the diagonal elements. Coherent errors manifest in the off-diagonal elements and are beyond classical. For a superconducting qubit the Pauli-6 POVM is performed by measuring individually the three bases of the Pauli-operators. We can then investigate how off-diagonal each of these bases are with respect to their ideal eigenbasis.

In Supplementary Figure 2 we present the reconstructed POVMs for some of the noisy runs in their ideal measurement basis (i.e. we present the POVM elements $R_{X\to Z}M_X R^\dagger_{X\to Z}$, $R_{Y\to Z}M_Y R^\dagger_{Y\to Z}$ and M_Z). For a single qubit it is sufficient to consider only the element corresponding to the outcome 0.

We found that for almost all cases, the z-basis measurement (the computational basis measurement) does not have coherence errors beyond statistical fluctuations, while for many instances the x- and y-basis measurements suffer from large coherence errors in the induced noise scenarios.

SUPPLEMENTARY NOTE 4: DIGITIZING EXPERIMENTAL DATA

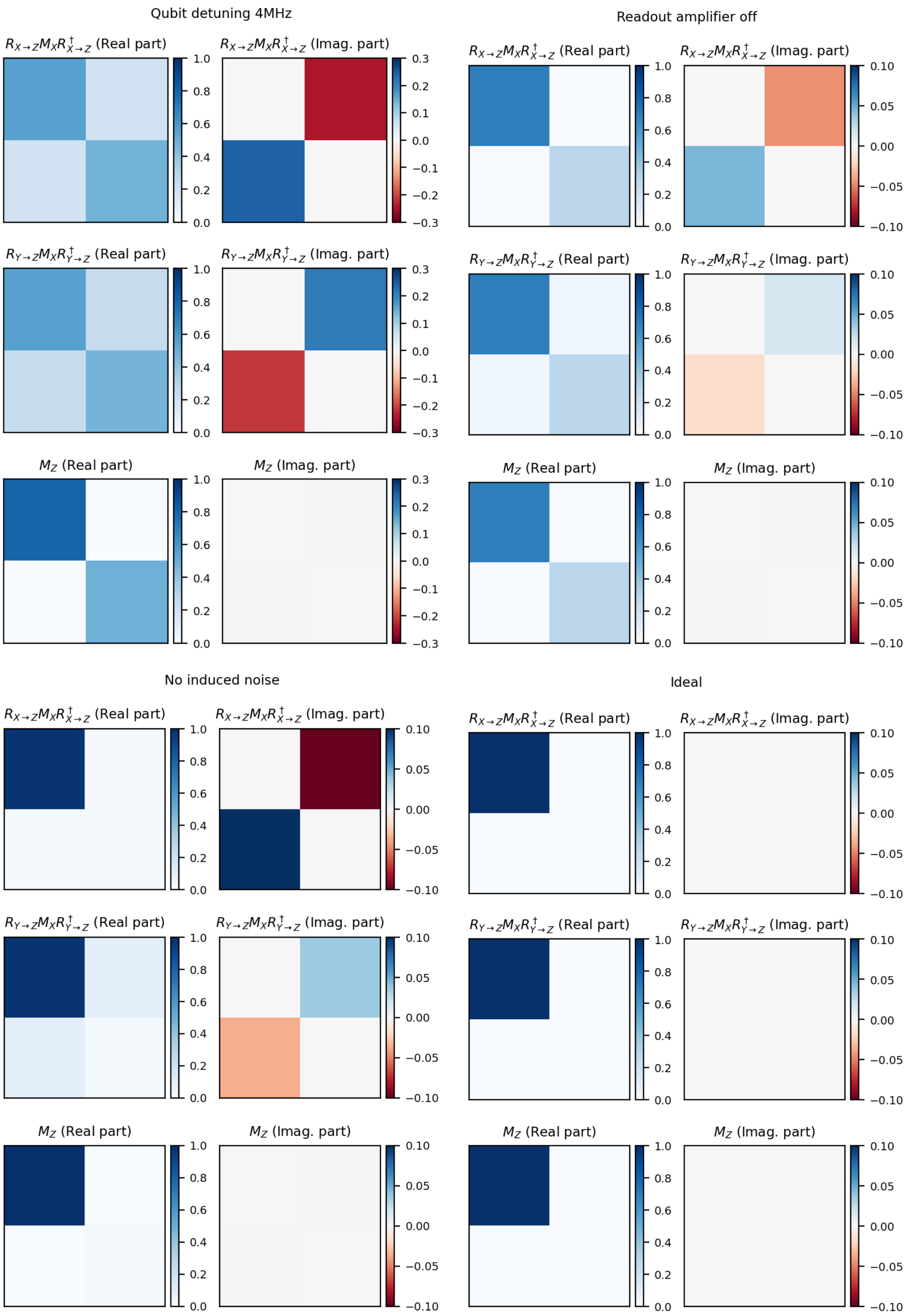
After measuring the length of a π-pulse, we use a supervised classification algorithm on the data to digitize the measured analog voltages (I and Q, see Supplementary Figure 3) into a single bit: 0 or 1. Often, two 2D Gaussian distributions can be fitted on the measured data points, but using machine-learning-based classification turned out to work more efficiently than Gaussian fitting, particularly for strong noises. We used the Python package sklearn, which runs in an intermediate level between the low-level experimental drivers and the high-level Bayesian and MLE algorithms.

SUPPLEMENTARY NOTE 5: OPEN-ACCESS DATASET

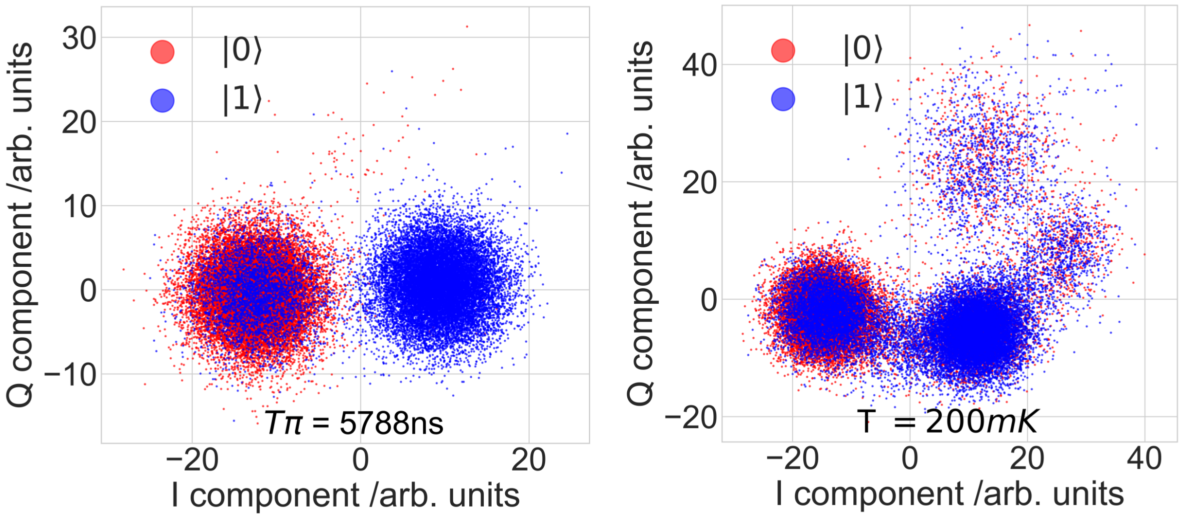
The datasets used for the experimental plots in the main text can be found on the KIT research data repository: https://radar.kit.edu/radar/en/dataset/RNbuograoVUFQNBB. The dataset contains following folders:

Figure	Folder name
Figure 2	4noises
Figure 3	combined noises
Figure 4	dt shots
Figure 6	theory depol
Figure 7	temperature
Figure 10	two different readout amplitudes
Supplementary Figure 1	0_2 ro amp spreads

Note that Figure 6 contains purely simulated data.



Supplementary Figure 2. The positive operator-valued measure (POVM) elements in their ideal eigenbasis. Non-trivial off-diagonal elements signal coherent errors that are not captured by classical error mitigation. The plots shows, starting from the top left: Qubit detuning at 4 MHz, turning off the parametric amplifier, optimal operation of the experiment, and theoretical ideal POVM. The POVMs were estimated using $N = 240$k single-shot measurements for each computational basis, which means statistical fluctuations are on the order of $\propto 3\times10^{-2}$, well below the coherence errors observed.



Supplementary Figure 3. Single-shot qubit measurements in the IQ plane, color-coded to the two qubit states. Long manipulation pulses allow the qubit to decay more, populating the ground state more often (left). Higher temperatures shift the readout to higher states, populating out-of-qubit states, too (right).

Supplementary Algorithm 1 Pseudo-code for readout error mitigation protocol

Input:
$\{\rho_B\}$: List of calibration states for QDT
$\{\rho_S\}$: (Haar-random) list of unknown states to measure.
$\{\{M_i\}\}$: A set of POVMs forming an IC POVM, e.g basis measurement in $\sigma_x, \sigma_y, \sigma_z$
Output:
$\{\rho_{\text{REM}}\}$: List of readout error mitigated reconstructed states
$\{\{\tilde{M}_i\}\}$: A set of reconstructed noisy POVMs
$\{I\}$: List of reconstruction infidelites between readout error mitigated QST and standard QST

Experiment:

```
for each measurement basis $\rho_B$ do
    for each POVM $\{M_i\}$ do
        $\text{QDT}_{\text{data}} \leftarrow \text{Measure}(\rho_B, \{M_i\})$
    end for
end for
for each Haar-random state $\rho_S$ do
    for each POVM $\{M_i\}$ do
        $\text{QST}_{\text{data}} \leftarrow \text{Measure}(\rho_s, \{M_i\})$
    end for
end for
```

Post-processing:

```
for each POVM in IC POVM do
    $\{\tilde{M}_i\} \leftarrow \text{QDT}(\text{QDT}_{\text{data}}, \rho_B)$ [1]
end for
for each unknown state $\rho_S$ do
    Construct modified likelihood function $\mathcal{L}_{\tilde{M}}(\rho)$, eq. (12) in main text
    Construct standard likelihood function $\mathcal{L}_M(\rho)$, eq. (8) in main text
    if BME then
        Integrate $\mathcal{L}_{\tilde{M}}$ and find mean state $\rho_{\text{BME}}$ [3]
        Integrate $\mathcal{L}_M$ and find mean state $\tilde{\rho}_{\text{BME}}$
    end if
    if MLE then
        Find $\rho_{\text{MLE}}$ that maximizes $\mathcal{L}_{\tilde{M}}$ [4]
        Find $\tilde{\rho}_{\text{MLE}}$ that maximizes $\mathcal{L}_M$
    end if
    $\{\rho_{\text{REM}}\} \leftarrow \rho_{\text{MLE/BME}}$
    $\{I\} \leftarrow \text{Inf}(\rho_{\text{MLE/BME}}, \tilde{\rho}_{\text{MLE/BME}})$
end for
```

SUPPLEMENTARY REFERENCES

[1] J. Fiurášek, "Maximum-likelihood estimation of quantum measurement," *Physical Review A*, vol. 64, July 2001.
[2] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information.* Cambridge University Press, June 2012.
[3] G. I. Struchalin, I. A. Pogorelov, S. S. Straupe, K. S. Kravtsov, I. V. Radchenko, and S. P. Kulik, "Experimental adaptive quantum tomography of two-qubit states," *Physical Review A*, vol. 93, Jan. 2016.
[4] A. I. Lvovsky, "Iterative maximum-likelihood reconstruction in quantum homodyne tomography," *Journal of Optics B: Quantum and Semiclassical Optics*, vol. 6, pp. S556–S559, May 2004.

5

SCALABLE READOUT-ERROR MITIGATION

Although the REMST protocol offers significant benefits, it also has notable limitations. Both the sample complexity and the post-processing complexity of the REMST protocol scale exponentially in the number of qubits in the system. Though useful as a benchmarking tool for small systems, it limits the applicability of the protocol to current state-of-the-art hardware, which already has more than 100 qubits. The work presented here is designed to alleviate this limitation while preserving most of the desirable properties of the REMST protocol. Motivation comes from the observation that many relevant applications in larger systems require only few-qubit observables, such as the extraction of phase-transition criticalities [167–169] and Lieb-Robinson bounds [170, 171]. In these cases, it is sufficient to perform readout-error mitigation on smaller subsystems.

The concrete task addressed in this work is to extract readout-error-mitigated two-qubit observables for arbitrary pairs of qubits in systems with up to 100 qubits. To achieve this, we generalize the REMST approach to enable sample-efficient reconstruction of arbitrary readout-error-mitigated reduced density matrices (RDMs). This is made possible by employing overlapping tomography, discussed in Section 2.7.4. The main challenge in scaling up the protocol lies in managing correlated readout errors between a principal pair of qubits and the rest of the system. Typically, the reconstruction of RDM assumes a Markovian environment, that is, the interaction between a principal system and the environment stays constant over time. However, when correlated noise exists between the principal system and specific qubits in the environment, this Markovian assumption is violated, as illustrated in Figure 5.1.

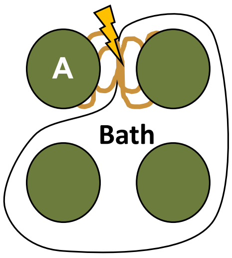


Figure 5.1: Illustration of a non-Markovian interaction between a principal system A and a bath environment. The principal system is strongly correlated to an individual qubit in the bath, represented by the orange wavy lines. This strong interaction with a single qubit disrupts the Markovian assumption, implying that the bath's influence is no longer time-independent.

To address this issue, we used correlation coefficients based on two-qubit POVMs to identify strongly correlated qubits [36, 161]. Clustering algorithms were then used to determine which subsets of qubits required joint mitigation. After identifying limitations in standard clustering approaches, we proposed a new and more reliable clustering algorithm based on hierarchical clustering. The resulting correlated REMST protocol was thoroughly tested using numerical simulations that included both highly correlated and experimentally extracted noise models. To validate the scalability of our approach, we simulated experimentally obtained noise from a superconducting qubit system on up to 100 qubits, observing an order-of-magnitude improvement over readout-error-mitigation protocols that neglect correlated readout errors.

Author participation is stated in the acknowledgment of the manuscript: "***A.A.*** [the author] *developed the protocol and software with the supervision of M.G. A.D., H.R. and A.U. provided the experimental characterization of the measurement device used for sample generation.* ***A.A.*** *wrote the manuscript with contributions from A.D and M.G. All authors contributed to the finalization of the manuscript.*".

PHYSICAL REVIEW A **112**, 022601 (2025)

Mitigation of correlated readout errors without randomized measurements

Adrian Skasberg Aasen,[1,2,*] Andras Di Giovanni,[3] Hannes Rotzinger,[3,4] Alexey V. Ustinov,[3,4] and Martin Gärttner[2,†]
[1]*Kirchhoff-Institut für Physik, Universität Heidelberg, Im Neuenheimer Feld 227, 69120 Heidelberg, Germany*
[2]*Institut für Festkörpertheorie und -optik, Friedrich-Schiller-Universität Jena, Max-Wien-Platz 1, 07743 Jena, Germany*
[3]*Physikalisches Institut, Karlsruher Institut für Technologie, 76131 Karlsruhe, Germany*
[4]*Institut für QuantenMaterialien und Technologien, Karlsruher Institut für Technologie, 76344 Eggenstein-Leopoldshafen, Germany*

(Received 1 April 2025; accepted 9 July 2025; published 4 August 2025)

Quantum simulation, the study of strongly correlated quantum matter using synthetic quantum systems, has been the most successful application of quantum computers to date. It often requires determining observables with high precision, for example when studying critical phenomena near quantum phase transitions. Thus, readout errors must be carefully characterized and mitigated in data postprocessing, using scalable and noise-model agnostic protocols. We present a readout error-mitigation protocol that uses only single-qubit Pauli measurements and avoids experimentally challenging randomized measurements. The proposed approach captures a very broad class of correlated noise models and is scalable to large qubit systems. It is based on a complete and efficient characterization of few-qubit correlated positive operator-valued measures, using overlapping detector tomography. To assess the effectiveness of the protocol, observables are extracted from simulations involving up to 100 qubits employing readout errors obtained from experiments with superconducting qubits.

DOI: 10.1103/6p6s-t8b7

I. INTRODUCTION

A significant challenge in scaling up quantum devices to regimes where they can demonstrate a quantum advantage is their sensitivity to noise [1]. Various error mitigation protocols for near-term quantum hardware have been developed to extend the reach of non-fault-tolerant devices [2–5]. Although the value of these short-term solutions has recently been subject to debate [6], there are arguments suggesting that they will remain crucial for reaching practical quantum advantage [7]. What is more, the mitigation of *readout* errors takes a distinct and important role in many applications, as these errors persist even when fault tolerance is implemented. As quantum hardware scales up, unintended readout correlations between qubits become a mounting issue that has only recently been discussed [8–11]. In superconducting qubits, readout correlations have been observed in multiplexed dispersive readout between pairs of qubits [12]. Early efforts to reduce such readout correlations focused on fine-tuning quantum control [13] or mitigating them postmeasurement with specialized hardware [14].

We have identified four properties which are desirable in practical near-term readout error-mitigation protocols:

(A) scalability to large qubit numbers,
(B) robustness against correlated and coherent errors,
(C) access to relevant observables,
(D) practicality of experimental implementation.

Among the most studied readout errors are classical errors, which amount to a statistical redistribution of measurement outcomes. These errors are commonly modeled using a confusion matrix [15,16] or addressed by unfolding [17]. The methods used to mitigate these types of errors have the advantage of being extremely simple to implement and typically account for the largest part of readout errors observed in experiments. These error-mitigation methods can also be made scalable [8], which makes them satisfy properties (A), (C), and (D), but fail to address coherent errors. Randomized measurement strategies, such as classical shadows [18], are of particular interest because many nonclassical errors are transformed to classical ones under randomization. Classical shadows have been made resistant to noise by the introduction of robust shadow tomography [19] and related methods [20–22]. These techniques satisfy properties (A)–(C), but are arguably difficult to implement due to the need for randomized global Clifford measurements. Recent efforts have been made to alleviate some of the difficulties of randomized measurements [23,24], but they often require entangling gates, such as CNOT gates, to be effective. Other methods that avoid randomized measurement have been proposed [25], but are not scalable, satisfying only properties (B)–(D). Correlated readout errors have been avoided entirely with compression readout [26,27], which involves only measurements of a single ancilla qubit, but requires relatively complicated compression circuits, satisfying properties (A)–(C). To the authors' knowledge, there is no readout error-mitigation method that satisfies properties (A)–(D) simultaneously.

In this work, we extend the readout error mitigated state tomography (REMST) protocol introduced in Ref. [25] to

---

[*] Contact author: adrian.aasen@uni-jena.de
[†] Contact author: martin.gaerttner@uni-jena.de

be scalable while maintaining desirable properties with only small caveats to (B) and (C). The REMST protocol is based on the calibration of the measurement device using detector tomography [28] and the integration of the reconstructed positive operator-valued measure (POVM) directly into a likelihood-based state estimator. Here, we develop a scheme that uses this protocol for the readout error-mitigated extraction of few-body observables that is scalable if readout-induced correlations are limited to small subsystems of qubits. Under this assumption, we can create a tiling of the whole qubit system with at most n_{corr} qubits in each tile. The tiling for a qubit device can be constructed efficiently by performing detector overlapping tomography [8,29] and extracting readout correlation coefficients [8,30] for all pairs of qubits. Using the correlation coefficients, one can create correlated noise clusters using a heuristic-free hierarchical clustering algorithm.

To demonstrate the utility and necessity of the protocol, we consider different realizations of readout error mitigation without randomized measurements and compare how well they perform in relevant tasks, such as reconstructing few-body reduced density matrices and extracting arbitrary two-point observables. Comparisons are made by numerically simulating measurements from noisy POVMs for system sizes between 16 and 100 qubits. To demonstrate the robustness of our scheme, we consider both artificial POVMs with correlated noise and noisy POVMs extracted from superconducting qubits by performing detector tomography. We find that our protocol performs well across all scenarios tested. When no knowledge of the prepared state is available, our protocol outperforms other alternatives by up to an order of magnitude in mean-squared errors of arbitrary two-point observables.

Our protocol will yield crucial advantages in quantum simulation experiments [31] where precise extraction of few-body observables is necessary but may be limited by readout errors. For example, preparing ground states for studying quantum phases of matter using variational quantum circuits requires measuring the expectation value of the system Hamiltonian which, for quantum spin models or Hubbard models, typically consists of sums of low-weight operators. In addition, for quantum simulations of the critical behavior around phase transitions [32–34], two-point correlators need to be precisely measured to extract critical exponents. When studying nonequilibrium quantum dynamics, the accurate determination of two-point correlators is key for determining the rate of information spreading and probing Lieb-Robinson bounds [35,36]. Furthermore, the tools and approach presented can be used separately for characterization of the readout correlation structure, which can be used to optimize experimental readout parameters [37]. Thus, we expect our scalable readout error-mitigation scheme to be of wide use to quantum simulation experiments, opening up new opportunities for studying quantum many-body phenomena.

II. PRELIMINARIES

A. Generalized quantum measurements

A generalized quantum measurement, also called a positive operator-valued measure (POVM), is a set of Hermitian matrices $\mathbf{M} = \{M_i\}$ that satisfies the following properties:

$$M_i \geqslant 0, \quad M_i^\dagger = M_i, \quad \text{and} \quad \sum_i M_i = \mathbb{1}. \tag{1}$$

Each POVM element M_i is associated with an outcome i from a measurement process. The Born rule provides the probabilities for the different possible outcomes when measuring quantum state ρ:

$$\mathrm{Tr}(\rho M_i) = \langle M_i \rangle = p_i. \tag{2}$$

This provides an operational interpretation of the expectation value of POVM elements as a probability distribution. Thus, the three properties in Eq. (1) are equivalent to positivity, realness, and normalization of the probability distribution, respectively.

The benefit of working with POVMs rather than just projective measurements is that it allows us to capture the process of noise at readout, both classical noise, which manifests itself as a redistribution of the eigenvalues of the POVM elements, and coherent errors, which occur in the off-diagonal elements.

B. Readout error mitigated state tomography

In Ref. [25], a readout error-mitigation protocol was introduced that allows correction of beyond-classical readout errors in quantum state reconstruction. This work is a direct continuation and will provide a scalable protocol. The basic outline of the approach is as follows: Quantum detector tomography is used as a calibration step, where the effective POVM $\mathbf{M}$ implemented by the measurement device is reconstructed [38]. The reconstructed POVM is then integrated into the likelihood function

$$\mathcal{L}(\rho) \propto \Pi_i \mathrm{Tr}(\rho M_i)^{n_i}, \tag{3}$$

where n_i is the number of occurrences of outcome i. Using a Bayesian mean estimator [39] or a maximum likelihood estimator [40], a most probable state can be found that is guaranteed to be physical.

C. Correlation coefficients and traced-out POVMs

To account for correlated readout errors, it is essential to have a precise method to assess how the readout of one qubit affects the readout of another. To do this, one can check how separable a POVM is between two subsystems. In contrast to computing reduced quantum states, extracting a traced-out POVM from a larger POVM lacks a unique definition and necessitates knowing the quantum state being measured. To create an unambiguous notion of a reduced POVM we follow Ref. [8]. Consider a bipartite quantum system, A and B, with POVM $\mathbf{M}^{\mathrm{AB}} = \{M^{\mathrm{AB}}_{i_\mathrm{A} i_\mathrm{B}}\}$, where i_A and i_B label the outcome in subsystems A and B, respectively. To derive the POVM that acts on the subsystem A, we define the traced-out POVM $M^{\mathrm{A},\rho_\mathrm{B}}$ conditioned on subsystem B being prepared in state ρ_B as

$$M^{\mathrm{A},\rho_\mathrm{B}}_{i_\mathrm{A}} = \sum_{i_\mathrm{B}} \mathrm{Tr}_\mathrm{B}\big[M^{\mathrm{AB}}_{i_\mathrm{A} i_\mathrm{B}}(\mathbb{1}_\mathrm{A} \otimes \rho_\mathrm{B})\big]. \tag{4}$$

By specifying the state of subsystem B, the relative occurrence of the different possible reduced POVM elements on subsystem A is fixed, and therefore unambiguous.

As initially outlined in Refs. [8,30], correlation coefficients can be formulated using traced-out two-qubit POVMs. These coefficients are defined as the maximum distance between two reduced POVMs, obtained by preparing the traced-out subsystem in different pure states. Following to the two-qubit "worst-case" definition from Ref. [8], the correlation coefficient obtained from the computational basis POVM is explicitly given by

$$c_{\mathrm{B}\to\mathrm{A}} = \sup_{\rho_{\mathrm{B}},\sigma_{\mathrm{B}}} \left\| M_0^{\mathrm{A},\rho_{\mathrm{B}}} - M_0^{\mathrm{A},\sigma_{\mathrm{B}}} \right\|_\infty , \tag{5}$$

where it suffices to consider the 0 element of the reduced POVM of subsystem A, as it fully specifies the one-qubit POVM due to the normalization condition $M_0 + M_1 = \mathbb{1}$. This coefficient quantifies how strongly the readout of subsystem B impacts the readout of subsystem A. Since this definition of the correlation coefficient is not symmetric, we define a symmetric version

$$c_{\mathrm{A}\leftrightarrow\mathrm{B}} = \frac{c_{\mathrm{A}\to\mathrm{B}} + c_{\mathrm{B}\to\mathrm{A}}}{2} \tag{6}$$

which we will refer to as the correlation coefficient from now on. Certain limitations of the correlation coefficients are addressed in Appendix A.

D. Overlapping detector tomography and perfect hash families

Our protocol requires the extraction of all pairwise correlation coefficients between qubits. To achieve this efficiently for large-scale qubit systems while minimizing the experimental cost, we employ overlapping detector tomography [8]. Various forms of overlapping tomography have been introduced [29,41,42], each grounded in the principles of covering arrays or perfect hash families (PHF). For simplicity, this discussion will focus on the variant using PHFs proposed by Ref. [29], as it offers the greatest flexibility within our framework. For our purposes, a hash family can be denoted as $\Phi_{N,k,v} = \{\phi_i^{k,v}\}_{i=1}^N$, which is a set of N hash functions ϕ that assign to k possible input values one of v possible output values [43,44]. For a hash family to be perfect, it requires that for any arbitrary subset of $t \leqslant v$ input values $\{x_1, x_2, \ldots, x_t\}$, there exists a hash function $\phi_i \in \Phi_{N,k,v}$ such that $\{\phi_i(x_1), \phi_i(x_2), \ldots, \phi_i(x_t)\}$ has t unique output values. We refer to these as t-local PHFs. In other terms, a t-local PHF contains hash functions that ensure that for any collection of t input values, there exists at least one hash function that assigns distinct output values to each of the t inputs. For the rest of this work, we will consider PHFs that have $v = t$. In this work, it is only necessary to use a PHF with $t = 2$, which can be generated analytically [29],

$$\phi_i^{k,2}(x) = i\text{th digit of the binary expansion of } (x-1), \tag{7}$$

where x denotes the input value.

In overlapping detector tomography, the PHFs are used to guarantee that all possible two-qubit POVMs can be reconstructed, which means that, for all possible two-qubit pairs, an informationally complete set of calibration states needs to be measured between them. Specifically, a minimal informationally complete set of calibration states for a single qubit comprises four calibration states, $\{\psi_1, \psi_2, \psi_3, \psi_4\}$, and for two qubits, all 16 combinations of two single-qubit calibration states are necessary. As an example of how PHFs are used to generate a complete set of calibration states for all qubits, consider the 2-local PHF described in Eq. (7). In this context, the input values x are the indices of each qubit, $x \in \{1, 2, 3, \ldots, n_{\text{qubits}}\}$. For each hash function $\phi_i^{k,2}$, each qubit is assigned one of two possible output values 0 or 1. An informationally complete set of calibration states is prepared between the qubits that were assigned 0 and 1. Specifically, all possible unique combinations of the single-qubit calibration states are prepared between the two groups, with no repeated calibration state, that is, for qubits with label (0,1) the calibration states $\{(\psi_1, \psi_2), (\psi_1, \psi_3), (\psi_1, \psi_4), (\psi_2, \psi_1), \ldots, (\psi_4, \psi_3)\}$ are prepared and measured, where all cases with equal calibration states prepared on each qubit group are removed. Once this is done for all hash functions, four final calibration states are measured with all qubits prepared in ψ_1, ψ_2, ψ_3, and ψ_4, respectively. This procedure is straightforwardly generalized to a t-local PHF, which ensures that any t-qubit POVM can be reconstructed. If the calibration states are replaced by single-qubit basis measurements $\{X, Y, Z\}$, one recovers the overlapping tomography prescription in Ref. [29].

The total number of measurements required to reconstruct any t-qubit POVM from a k-qubit system is proportional to the number of hash functions in the PHF. The expected number of hash functions scales as $e^{\mathcal{O}(t)} \log k$ [29], that is, exponential in the subsystem size but logarithmic in the total size of the system, making overlapping detector tomography scalable.

There exist situations where larger t-local PHF are useful and we do use a PHF with $t = 3$ for fairer comparisons later. Creating an optimal PHF is a hard problem, and few analytical methods are available to generate them [43,44]. We provide some PHFs generated by a density algorithm for different k and t in Ref. [45]. An example of the $\Phi_{16,15,3}$ PHF is presented in Appendix B. Although the proposed measurement strategy may not be resource optimal, it offers significant adaptability. Additional optimizations could be achieved through the use of covering arrays [41,46]

III. METHOD

This section introduces our main result, a protocol specifically devised to address correlated readout errors in any arbitrary two-qubit observable, with the potential to extend to higher-order observables in certain subsystems. As discussed in Sec. I, high-precision measurements of low-order observables in the presence of correlated readout errors are crucial for many applications in quantum simulation. The correlated quantum readout error-mitigation (QREM) protocol is depicted in Fig. 1. The protocol can be broadly split into two separate parts, readout correlation characterization and noise-cluster-based REMST. The first part is used to sample efficiently extract the correlated structure of the qubit readout which enables construction of correlated noise clusters. The obtained correlated noise structure is used to extract error mitigated observables in the subsequent part.

A. Noise-cluster characterization

In general, readout correlations may exist among all qubits in any quantum system, rendering any complete characterization nonscalable. Therefore, in order to efficiently characterize

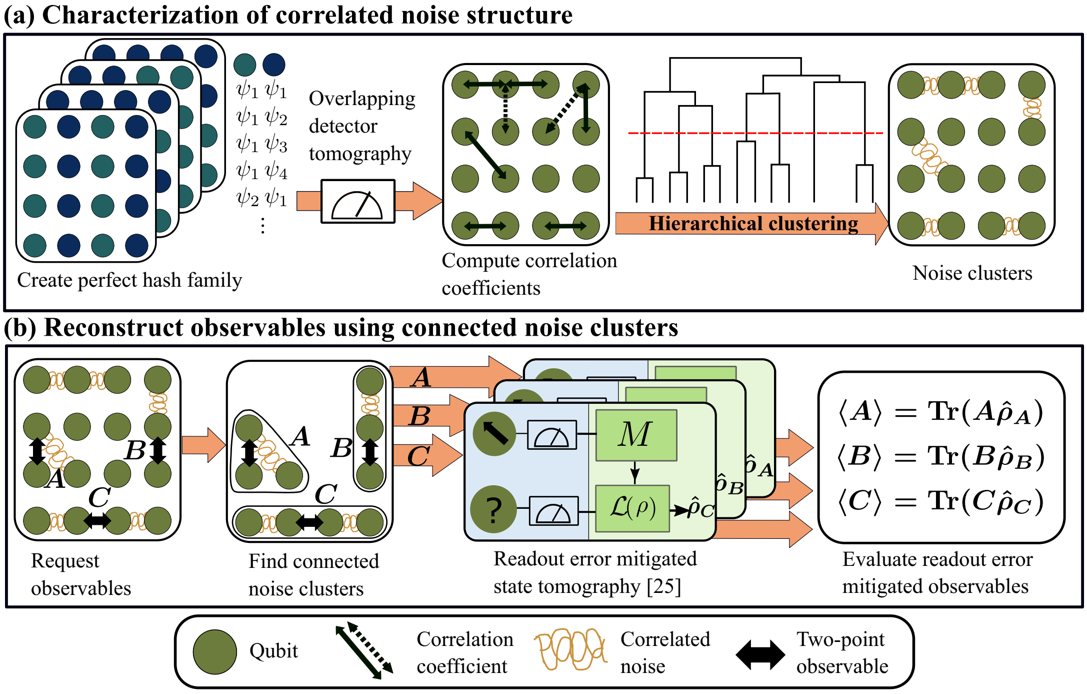


FIG. 1. Schematic overview of the correlated readout error-mitigation protocol. (a) The correlated noise is efficiently characterized for the relevant qubit system. A perfect hash family is generated in such a way that all possible two-qubit reduced POVMs can be reconstructed by utilizing parallel measurements. Every hash function assigns one of two distinct labels to each qubit, indicating which calibration state is assigned to each group. Each group label receives an informationally complete set of calibration states between them which is repeatedly measured. Overlapping detector tomography can then be performed where every possible two-qubit POVM is reconstructed. For each two-qubit POVM, the symmetric correlation coefficients are calculated [see Eq. (5)]. These coefficients are represented with bidirectional arrows. The dotted bidirectional arrow indicates weaker correlation coefficients. By employing a hierarchical clustering algorithm, the dominant readout noise clusters within the system can be identified. (b) Reconstructing observables with mitigated readout errors. Each requested observable is matched with readout noise clusters to form connected noise clusters, which must be fully characterized. On these connected noise clusters, readout error-mitigated state tomography is carried out to reconstruct the error-mitigated states for each connected cluster. The observables are then computed from the readout error-mitigated states.

the noise in the entire system, assumptions about the correlation structure have to be made. Fortunately, the extent of correlation in contemporary systems is typically limited, and readout correlations occur only within smaller groups, such as for superconducting qubits that share the same readout microwave line. Therefore, a maximum number of qubits that can be correlated n_{corr} can be set.

To find which qubits should be grouped together into noise clusters, we rely on two-qubit correlation coefficients, which are computed from two-qubit POVMs [see Eq. (5)]. Reconstructing all possible two-qubit POVMs would require an exhaustive amount of measurements if done naively by measuring all two-qubit POVMs one after another. It can be done significantly more efficiently by employing overlapping detector tomography and perfect hash families to measure all qubits in parallel (see Sec. II D).

A 2-local PHF is generated to ensure that each set of two qubits is measured with an informationally complete set of calibration states. Higher k's can be used to create a "measure first, ask questions later" approach, where the initial calibration measurement can be reused in noise cluster characterization later. In the representation in Fig. 1(a), there are 16 qubits, for which four hash functions are sufficient. In each sheet all the qubits of the same colors received the same measurement instructions such that between the two groups there exists an informationally complete set of measurements. Once all four sheets are measured, one can reconstruct all possible two-qubit POVMs. With the two-qubit POVMs the correlation coefficients can be computed as described in Eq. (5).

The correlation coefficients form a fully connected graph, facilitating the use of clustering algorithms to identify groups of qubits with the strongest readout correlations. Hierarchical clustering techniques consistently identified qubit clusters across different noise models and experimentally extracted POVMs. These methods provide the benefit of eliminating the need for heuristic parameters, which often differ significantly between different experimental setups or noise models. Although further manual cluster tuning is possible with extra insight into the noise sources, it is not essential for the protocol to function effectively. For more details on the clustering algorithm, see Appendix C.

B. Reconstructing readout error-mitigated observables

The identified noise clusters are used to perform readout error mitigation [cf. Fig. 1(b)]. For a given two-qubit observable to be measured, all noise clusters containing those qubits are grouped to form a *connected* noise cluster. The readout error-mitigated state is reconstructed using the REMST protocol [25] on the entire connected noise cluster. In Appendix D we show that, in the case of a general prepared state, state reconstruction on connected noise clusters is required if both correlated and coherent readout errors are to be mitigated. The reconstruction of the noisy POVMs can be carried out on each noise cluster individually.

The observables initially requested are calculated using the readout error-mitigated states received from the REMST protocol. The connected noise clusters can be reused to compute further higher-order observables, even if not originally requested. For details on the implementation, see Appendix E.

C. Extraction of experimental POVMs

To extract realistic POVMs to test the protocol, detector tomography was performed on a chip with four individual transmon (Xmon) qubits. Each qubit is dispersively coupled to a separate readout resonator. The qubit control and the resonator readout is done with frequency division multiplexing [47,48]. The chip is measured in a dilution cryostat below 15 mK. For more information on the device used, see Ref. [49]. Some of the qubits have larger readout correlations as a result of the smaller frequency spacing of the resonators, leading to potential crosstalk.

IV. RESULTS

To demonstrate that all the desired properties stated in the Introduction are satisfied by the protocol, measurements were simulated on up to 100 qubits. These simulations were based on readout noise extracted from a superconducting qubit system. The properties examined were as follows:

(A) scalability to large qubit numbers,
(B) robustness against correlated and coherent errors,
(C) access to relevant observables.

The property (D) is satisfied by construction as only single-qubit Pauli measurements were used. To make the sampling and noise simulation numerically feasible, we used tensor products of four-qubit states as test states for the QREM protocols and applied noise channels that potentially overlap multiple four-qubit states. This should not be viewed as a limitation of the protocol because the results and protocol complexity would not change if general 16- and 100-qubit states were prepared (see Appendix D for more details).

The protocol is tested by reconstructing two-qubit observables. The target states are tensor products of four-qubit Haar-random states, $\rho = \bigotimes_{i=1}^{n_{\text{chunks}}} \rho_i$, where $n_{\text{chunks}} = n_{\text{qubits}}/4$. The Haar-random pure states ρ_i are generated by drawing a random unitary matrix U_i from the unitary group $U(2^4)$ and applying it to the all-zero state $\rho_0 = |0000\rangle\langle 0000|$, resulting in $\rho_i = U_i \rho_0 U_i^\dagger$. The considered observables are products of Pauli operators $O = A \otimes B$, where $A, B \in \{X, Y, Z\}$. In Appendix F 2 we include a simulation with mixed target states.

Multiple layers of averaging were performed to achieve representative performance estimates. Specifically, each simulation involved numerical measurement samples from $N_s = 10$ randomly chosen target states. For each state, N_p random pairs of qubits were selected, and for each selected pair, all two-qubit Pauli observables were reconstructed. As figures of merit, the mean-squared error (MSE) and state reconstruction infidelity were used. The mean-square error is computed between the reconstructed observables $\hat{O} = \text{Tr}(O\hat{\rho})$ and their ideal values $\langle O\rangle$ from the Haar-random target states,

$$\text{MSE} = \frac{1}{9}\sum_{i=1}^{9}(\langle O_i\rangle - \hat{O}_i)^2, \tag{8}$$

where the sum is performed over nine possible two-qubit Pauli observables. Unless stated otherwise, the MSE is also averaged over the N_s random target states and N_p pairs of qubits, which is denoted as the averaged MSE. State reconstruction infidelity is computed between the reconstructed error-mitigated state $\hat{\rho}$ and the target state ρ, defined as

$$\text{Infidelity} = 1 - F(\rho, \hat{\rho}) = 1 - \left[\text{Tr}\left(\sqrt{\sqrt{\rho}\hat{\rho}\sqrt{\rho}}\right)\right]^2. \tag{9}$$

Since only pure target states ρ are used, infidelity simplifies to

$$\text{Infidelity} = 1 - \text{Tr}(\rho\hat{\rho}). \tag{10}$$

As for the MSE, infidelity is averaged over N_s random states and N_p random pairs unless otherwise stated.

To evaluate the statistical robustness of the protocol, each simulation was repeated 10 times and averaged. The standard deviations of these repetitions are shown as error bars in each plot. For all simulations, the maximal size of the noise cluster was set to $n_{\text{corr}} = 3$.

A. 16-qubit example with experimental noise

To simplify our analysis, we initially focus on a 16-qubit system to test the performance of the protocol. Five different implementations of the protocol are compared in scenarios where some are expected to perform optimally and where no advantage is anticipated. The protocols considered are as follows:

(i) *No QREM*: No readout error mitigation is applied.

(ii) *Factorized QREM*: The REMST protocol in which each qubit is treated individually. The reconstructed noisy POVMs are tensor products of single-qubit POVMs.

(iii) *Two-qubit QREM*: The REMST protocol is applied to the two qubits that are part of the observable. The correlated readout errors are only mitigated for the two qubits participating in the observable.

(iv) *Classical correlated QREM*: The correlated QREM protocol is applied where only classical errors are corrected. The reconstructed noisy POVMs are produced using the full correlated QREM protocol, but discarding all off-diagonal terms in the POVMs.

(v) *Correlated QREM*: The complete correlated QREM protocol is applied.

Additional details on the different protocols can be found in Appendix E 4.

All protocols were evaluated for $N_s = 10$ random realizations of the 16 qubits extracting correlators between $N_p = 20$ random pairs of qubits. The readout noise was sampled at random from a large set of experimentally extracted POVMs (see Sec. III C). In Fig. 2 the MSE is plotted as an average

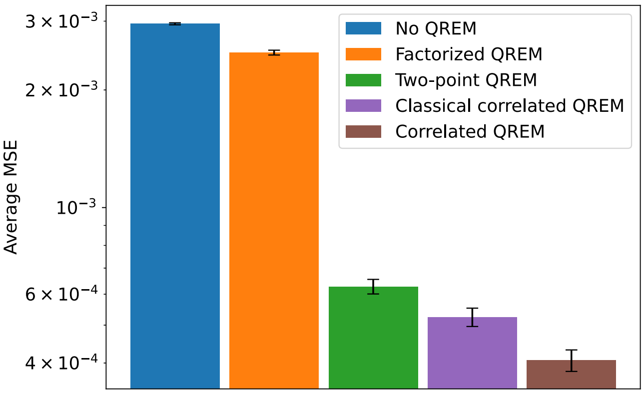


FIG. 2. Comparing different scalable readout error-mitigation protocols. Averaged mean-squared error of two-qubit Pauli observables is plotted for five different protocols. The averages are performed over $N_p = 20$ randomly chosen pairs for $N_s = 10$ random realizations of the 16-qubit states. The error bars are one standard deviation from 10 repeated simulations of the entire protocol with the same readout noise, random states, and pairs chosen.

over all states, pairs, and observables. For a fair comparison, each protocol received the same measurements in both noise calibration, when applicable, and readout error-mitigated state reconstruction. A 3-local PHF was utilized for calibration, resulting in a total of 9×10^6 calibration measurements performed. For each requested observable, $\approx 10^5$ measurements were performed. The correlated QREM protocol performed the best overall, indicating that both beyond-classical readout errors and correlated readout errors were present in the experimentally extracted POVMs. To better understand how the protocols work, specific pairs of qubits are examined in Fig. 3, where both state reconstruction infidelity and MSE are averaged only over randomly chosen target states. The [0, 1] pair featured correlated noise only between the qubits constituting the observable. Therefore, the factorized QREM protocol performed worse than the two-point and correlated QREM protocols, which worked equally well. The classical correlated QREM protocol performs slightly worse, which is expected, as small amounts of nonclassical errors are present in the experimental POVMs. For the pair [4, 8], there was no correlated readout noise between the two qubits appearing in the observable, but significant correlated readout noise between the adjacent qubits, specifically for (4, 5) and (8, 9, 10). It is therefore expected that the two-point QREM protocol does not perform too well while both correlated QREM protocols perform well. For the two remaining cases, [6, 11] and [7, 13], we observe similar performance between all QREM protocols, indicating that there were no significant correlated readout errors.

B. Strong artificial noise

In order to illustrate the applicability of the protocol to correlated and beyond-classical noise sources [property (B)], two artificial noise models were examined in a 16-qubit simulation: One characterized by significant correlated errors and the other by pronounced coherent errors. Both MSE and state reconstruction infidelity were averaged over $N_s = 10$ randomly chosen target states and $N_p = 20$ random pairs. The number of calibration measurements and state reconstruction measurements is the same as in Sec. IV A.

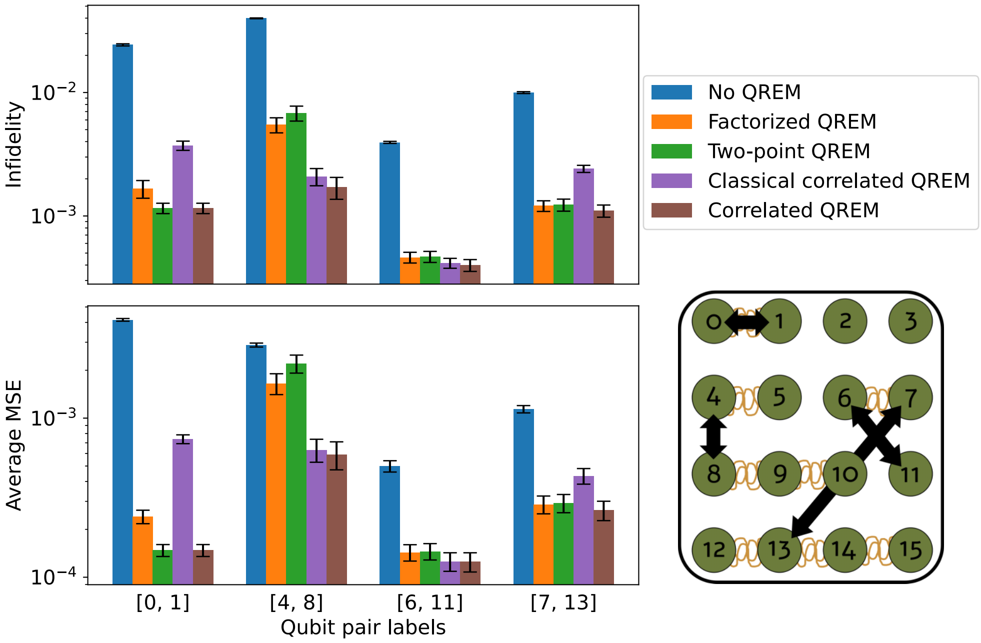


FIG. 3. Average readout error-mitigated observables for selected two-qubit pairs. In a 16-qubit simulation, state infidelity and average mean-squared error are evaluated across five distinct implementations of the scalable QREM protocol. The measurement setup is the same as for Fig. 2. In the bottom right the experimental n-qubit POVMs applied are visualized by orange squiggly lines, indicating potential correlated readout errors. The specific qubit pairs plotted are indicated by black bidirectional arrows.

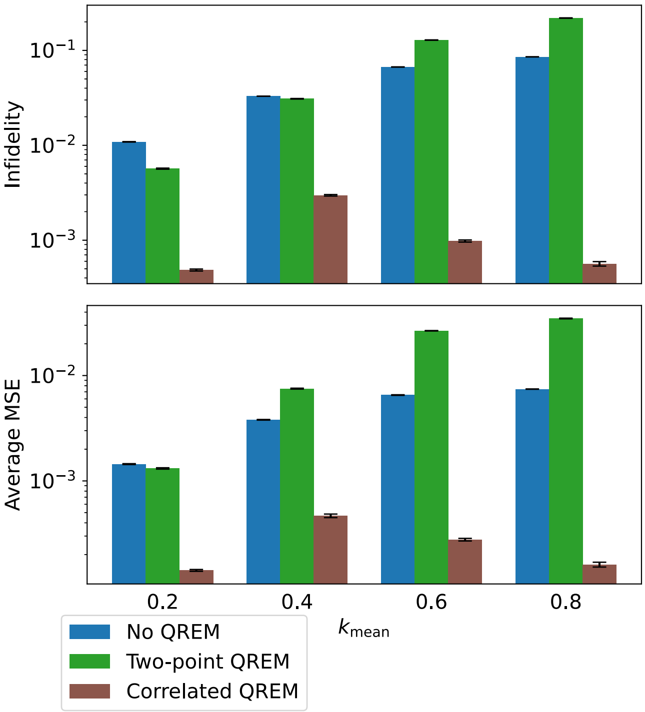


FIG. 4. Comparing protocol performance for increasing correlated noise strength. Readout error-mitigated infidelity and averaged mean-squared error for all two-qubit Pauli observables are plotted for the two-point QREM and the correlated QREM protocol. The error bars are the standard deviation from 10 repetitions of the simulation with identical initial conditions.

1. Strong correlated noise

Correlated noise can be generated by applying a noisy entangling gate between two qubits. The iSWAP gate is the native entangling gate for superconducting qubits and generates symmetric correlation between the two qubits, unlike the CNOT gate, which is directional. The simulated correlated noise is modeled as a probabilistic iSAWP channel

$$\mathcal{E}(M_i) = (1-k)M_i + kU_{\text{iSWAP}}M_iU_{\text{iSWAP}}^{\dagger}, \tag{11}$$

which can be understood as preserving the quantum state with probability k, while with probability $1-k$ an iSWAP gate is applied to the state, correlating the qubits. In the simulation, only neighboring pairs of qubits were correlated, i.e., pair $\{(0,1),(2,3),\dots,(14,15)\}$. For each pair of qubits, the noise strength k was randomly drawn from a uniform distribution around a mean value $k \in [k_{\text{mean}} - 0.1, k_{\text{mean}} + 0.1]$.

In Fig. 4 correlated QREM is compared to no QREM and two-point QREM for different mean noise strengths k_{mean}. The two-point QREM protocol has weaker assumptions than the factorized QREM protocol, and is a stronger candidate to compare to the correlated QREM protocol. The two-point QREM protocol accounts for correlations between the qubits in the two-point observable, but not other qubits.

With increasing noise strength, both the two-point protocol and the no-mitigation protocol performed significantly worse than the correlated QREM protocol. In particular, the two-point QREM protocol yielded marginally poorer results compared to when the no-mitigation protocol. This can be attributed to the fact that most random pairs of qubits involved qubits that belonged to distinct noise clusters. By overlooking correlated noise, the effective environment for these qubits exhibits non-Markovian behavior, causing a foundational assumption of the two-point protocol to break down, rendering it ineffective. When the mean noise strength is at $k_{\text{mean}} = 0.8$, the averaged MSE and the infidelity decrease for the correlated QREM protocol. This is likely because the reconstructed POVM is very close to being a swapping of the two qubits, which is a simpler error to mitigate than a more equal mixture of swapped and not swapped measurements. We note that the correlation coefficients generated by this model with $k_{\text{mean}} = 0.4$ are comparable to the highest correlation coefficients observed in state-of-the-art superconducting qubit systems [8] (see Fig. 7 in Appendix C for more details).

2. Strong coherent noise

Coherent errors can occur when unitary gates are applied incorrectly. As an example for superconducting qubits, single-qubit unitary rotations are applied by sending a resonant driving pulse for a specified time t. Inaccuracies in how long this pulse is applied cause either underrotation or overrotation of the qubit. A combination of correlated qubit readout and coherent errors can be modeled by single-qubit overrotations and neighboring two-qubit XX interactions [19]. Such a noise model can be formulated as a n-qubit channel with noise strength θ:

$$\mathcal{E}_{\theta}^{n}(M_i) = R_x^n(\theta)M_iR_x^{n\dagger}(\theta), \tag{12}$$

where

$$R_x^n(\theta) = \begin{cases} \exp\left(-\frac{i\theta}{2}\sum_{\langle j,k\rangle} X^j \otimes X^k\right), & \text{when } n > 1 \\ \exp\left(-\frac{i\theta}{2}X\right), & \text{when } n = 1. \end{cases} \tag{13}$$

The 16-qubit correlation structure was reused from the example in Fig. 2 (inset to the bottom right). For each noise cluster, the channel in Eq. (12) was applied with the same noise strength θ for all clusters. Although this noise model creates a four-qubit noise cluster, we continue to limit the maximum size of the cluster to $n_{\text{corr}} = 3$. This means that the correlated QREM protocol will not be able to capture all correlations in the system.

In Fig. 5 both correlation-conscious versions of the protocol are compared to no QREM. Classical correlated QREM only considered a statistical redistribution of the outcomes within the same cluster and ignored any off-diagonal entries in the POVM reconstruction, which is equivalent to using a confusion matrix to mitigate readout errors [15]. Both classical correlated and no QREM performed similar across the board, while correlated QREM achieved about an order of magnitude better results. There is a clear trend for the correlated QREM protocol upwards, which is due to the four-qubit correlated noise that is not captured with the maximum cluster size $n_{\text{corr}} = 3$. Given that this noise model is strong and unrealistic and is not anticipated in any plausible experiments, the performance of the protocol is considered satisfactory, as it still surpasses all other alternatives. The readout errors associated with a noise strength $\theta = 0.314$ are similar to the coherent errors observed between the qubit pair [0, 1] shown in Fig. 3,

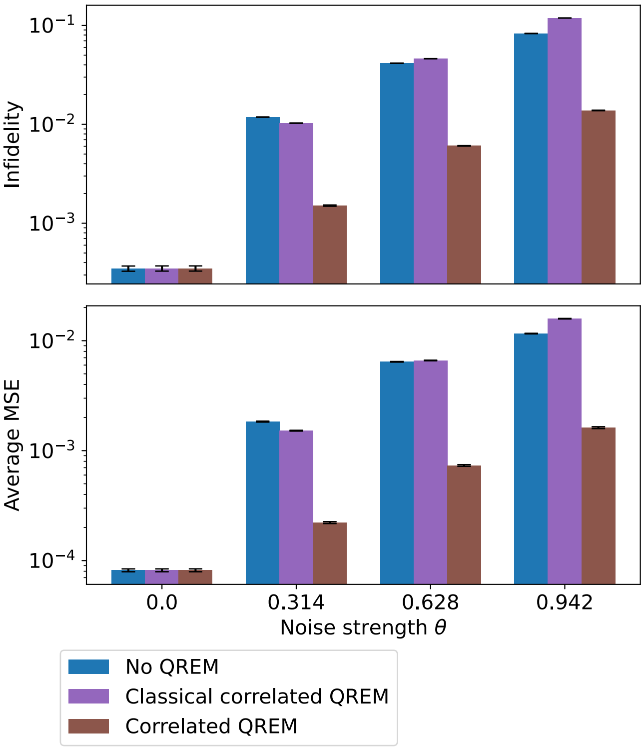


FIG. 5. Comparing protocol performance for increasing coherent noise. Readout error-mitigated infidelity and averaged mean-squared error for all two-qubit Pauli observables are plotted for the classical correlated QREM and the correlated QREM protocol. The error bars are the standard deviation from 10 repetitions of the simulation with identical initial conditions.

when comparing classical correlated QREM to correlated QREM.

For this simulation a different hierarchical clustering method was used than for all the other simulations due to an artifact in how the correlation coefficients are computed. This is discussed in detail in Appendix C 1. In Appendix F 1 a simulation is shown with the same clustering method as the other simulations.

C. 100-qubit readout error mitigation with experimental noise

To demonstrate scalability [property (A)], $N_s = 10$ random 100-qubit simulations were performed with readout noise sampled from experimentally extracted POVMs. The calibration step used a total of 11.2×10^6 measurements, and $\approx 5 \times 10^4$ measurements were performed for each requested observable. The correlated QREM protocol is compared to the two-point QREM protocol and no QREM. Figure 6 shows $N_p = 100$ randomly selected pairs of qubits sorted by increasing two-point QREM averaged MSE. The correlated QREM model consistently shows performance that is on par with or superior to the two-point model across the data set. A more distinct separation between the three methods appears in the more correlated pairs on the right side, highlighting the respective utility of the three protocols as no mitigation, partial mitigation, and full mitigation. For a system of this size, perfect clustering is unattainable, as detailed in Appendix C, leading to potential inaccuracies in identifying the appropriate noise cluster for some observables. Nonetheless, the correlated protocol consistently matches or exceeds the performance of other protocols. In Appendix F we include two additional 100-qubit simulations, one with mixed target states and the other with single-qubit observables included in the average MSE.

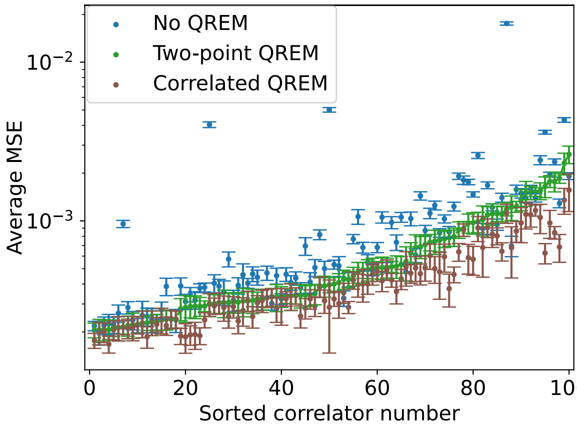


FIG. 6. 100-qubit example with POVMs extracted from superconducting qubits. The correlated QREM protocol is compared to the two-point QREM protocol. Random pairs are selected and the mean-squared error is computed over all two-qubit Pauli observables for each pair. The pairs are sorted by increasing MSE for the two-point QREM protocol. The error bars are the standard deviation from 10 repetitions of the simulation with identical initial conditions.

V. CONCLUSION AND OUTLOOK

We have presented a scalable readout error-mitigation protocol that takes into account the correlated readout structure. Correlated readout errors are efficiently characterized by combining overlapping detector tomography and hierarchical clustering to group the most correlated qubits into noise clusters. For a desired few-body observable, the involved noise clusters are combined to a connected noise cluster, on which the REMST protocol [25] is executed to reconstruct observables with mitigated readout errors. We demonstrate the effectiveness and generality of our protocol by comparing it with previously proposed QREM protocols. Simulations of up to 100 qubits were performed with both experimentally obtained readout noise and simulated correlated readout noise. In every scenario tested, the correlated QREM protocol shows a performance that is, at worst, on par with all competing protocols, but in most cases it substantially exceeds them in terms of reduction of the mean-squared errors of the averaged observables.

Although the protocol's sample complexity is logarithmic in the number of qubits, it is not guaranteed to be optimal due to challenges in identifying minimal measurement strategies. Finding a better parallel measurement scheme would allow for more efficient characterization of correlated readout, and also a more practical "measure first, ask questions later" approach, potentially by using covering arrays [46]. The correlation coefficients perform well as an indicator for correlated readout, but have certain drawbacks. When the strength of correlations within the system is large, it can assign correlation coefficients to uncorrelated qubits that are of the same order of magnitude as other weakly correlated qubits. These incorrect correlation coefficients can cause weakly correlated qubits to be incorrectly assigned to other groups. Although these incorrect assignments lower the average performance of the correlated QREM protocol, it still shows a significant improvement over other candidates. Finding better approaches to quantify the correlated readout would improve the correlated QREM protocol and is a potential future direction.

Looking forward, we will further explore the application of this protocol to large-scale experimental setups. Notably, in tasks such as evaluating correlation propagation in large quantum spin systems, we expect our protocol to yield a precision gain leading to new insights on nonequilibrium quantum dynamics.

ACKNOWLEDGMENTS

A.S.A. thanks K. Hansenne for fruitful discussions on overlapping tomography and covering arrays, N. Euler for suggesting a cluster verification approach, R. Dougherty for help with generating the perfect hash families, and M. D'Achille for technical assistance. The authors acknowledge support by the state of Baden-Württemberg through bwHPC. This research is supported by funding from the German Research Foundation (DFG) under the project identifier Grant No. 398816777-SFB 1375 (NOA).

A.S.A. developed the protocol and software with the supervision of M.G., A.D., H.R., and A.V.U. provided the experimental characterization of the measurement device used for sample generation. A.S.A. wrote the manuscript with contributions from A.D. and M.G. All authors contributed to the finalization of the manuscript.

DATA AVAILABILITY

The data that support the findings of this article are openly available [45].

APPENDIX A: LIMITATION OF THE CORRELATION COEFFICIENTS

Although the correlation coefficients provide a useful proxy for readout-induced correlations, they are not a strict correlation measure. They can and do assign nontrivial correlation coefficients to uncorrelated qubits. The origin of these unphysical correlations lies in the definition of the correlation coefficients. As emphasized in Sec. II C, defining a reduced POVM from a larger subsystem in the presence of correlated readout is a complex task. The starting point for the correlation coefficients are two-qubit POVMs which are directly reconstructed from measurements. One is therefore implicitly tracing out all the other qubits in the system. Recalling the definition of traced-out POVMs from Eq. (4), constructing reduced POVMs directly assumes the reference state as the maximally mixed state $\rho_{\mathrm{B}} = \mathbb{1}/2^{N-2}$. This is because the calibration states on the traced subsystem are an equal mixture of all possible symmetric and informationally complete calibration states [see Eq. (E3)], leading to reduced POVM,

$$M^{\mathrm{A}}_{i_{\mathrm{A}}} = \frac{1}{2^{N-2}} \sum_{i_{\mathrm{B}}} \mathrm{Tr}_{\mathrm{B}}\big(M^{\mathrm{AB}}_{i_{\mathrm{A}} i_{\mathrm{B}}}\big), \tag{A1}$$

where N is total number of qubits in the system, and the reference state label is omitted on the reduced POVM.

TABLE I. Perfect hash family $\Phi_{15,16,3}$. This perfect hash family was used to generate all samples used for the QDT step in Secs. IV A and IV B. It was generated by using the density algorithm presented in Ref. [44]

	k															
N	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0
2	0	1	2	2	1	0	1	0	2	2	1	0	1	2	0	1
3	0	1	2	2	0	1	2	1	0	1	2	0	0	2	1	1
4	0	1	2	2	0	1	2	0	1	2	0	2	1	0	2	0
5	0	1	2	1	0	2	0	2	1	2	1	0	1	0	1	2
6	0	1	2	1	0	2	0	2	1	0	2	1	2	1	0	1
7	0	1	2	2	1	0	2	1	0	1	2	0	0	0	0	0
8	0	1	2	2	1	0	2	2	1	0	0	1	2	1	2	0
9	0	1	2	1	0	2	2	0	0	1	2	2	0	1	1	0
10	0	1	2	2	1	0	1	2	0	1	2	1	1	2	2	2
11	0	1	2	0	2	0	0	0	1	1	0	0	2	1	1	0
12	0	0	1	2	1	1	1	2	0	0	1	1	2	2	1	1
13	0	0	1	2	1	0	2	2	2	1	1	1	2	2	2	0
14	0	0	0	0	1	1	1	2	2	2	2	2	2	2	0	0
15	0	0	0	0	0	0	0	0	1	1	1	1	1	2	2	2

This choice ignores the potential correlations that occur between the A and B subsystems. Therefore, if there were substantial correlations between these subsystems, the residual discrepancy would be attributed to the correlation between two qubits in subsystem A. This becomes apparent if there are groups of strongly correlated qubits and other groups of weakly correlated qubits which are mutually uncorrelated. The correlation coefficients between a qubit from the weakly correlated group and a qubit from the strongly correlated group can be larger than the correlation coefficients between two weakly correlated qubits. In effect, this causes an unphysical correlation to be assigned to pairs of qubits which may not share any correlation.

APPENDIX B: PERFECT HASH FAMILIES

The PHF $\Phi_{15,16,3}$ used for the 16-qubit simulations is presented in Table I. In Ref. [45] we provide a set of other PHF, specifically $\Phi_{36,16,4}$, $\Phi_{37,100,3}$, and $\Phi_{87,50,4}$.

APPENDIX C: HIERARCHICAL CLUSTERING ALGORITHM AND LIMITATIONS

In this work, we used hierarchical agglomerative clustering algorithms to group correlated qubits. We will only give a brief and qualitative description of the clustering algorithm (for more details see Ref. [50]). In simple terms, a hierarchical agglomerative clustering algorithm uses the mutual distance between all data points to procedurally combine the data points with the smallest pairwise distance into clusters.

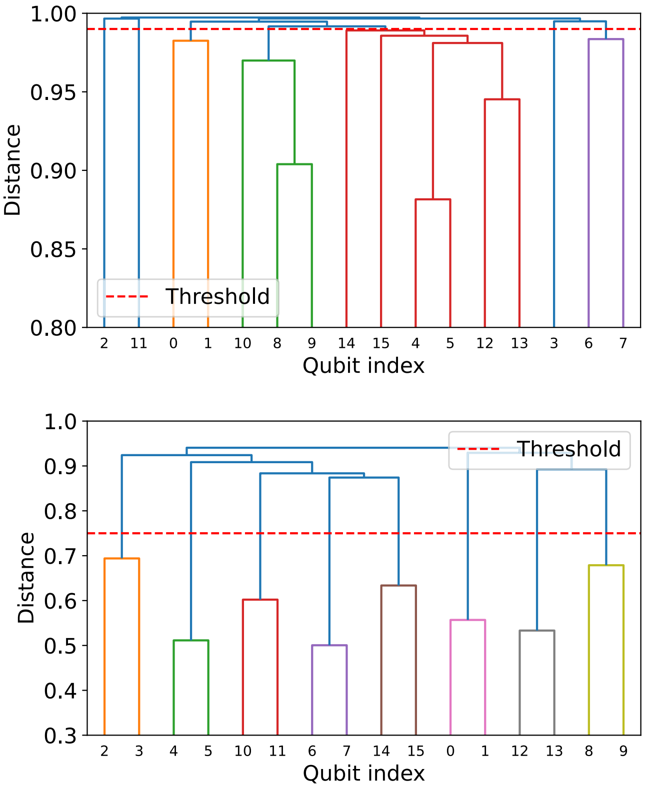


FIG. 7. Examples of automatic and manual dendrogram tuning. Distance refers to d^2. Top**:** Dendrogram from Sec. IV A. No tuning was performed and the automatic threshold $\gamma = 1 - 1/\sqrt{n_{\text{calib}}}$ was used. The largest cluster was iteratively split into $\{(14), (15), (4, 5), (12, 13)\}$. Bottom**:** Dendrogram from Sec. IV B 1 with $k = 0.4$. Manual threshold was set due to clear separation of correlation structure. The characteristic value of the correlation coefficients for this noise model are comparable with the worst qubit correlations observed in Ref. [8] (IBM Cusco $d^2 \approx 0.7$ and Rigetti Aspen-M-3 $d^2 \approx 0.6$.).

This process continues until a singular cluster is achieved that includes all data points. For our purposes the data points are the qubits and the mutual distances are derived from the symmetric correlation coefficients (see Sec. II C), and are defined as

$$d_{ij}^2 = 1 - c_{i\leftrightarrow j}. \tag{C1}$$

This gives a distance where $d = 1$ means completely uncorrelated and $d = 0$ means maximally correlated. To limit how large the clusters can grow, a threshold value γ can be chosen such that no clusters with intracluster distance larger than γ will be grouped together. The threshold value is upper bounded by 1 and should be set at least below the expected statistical fluctuations $\gamma \leqslant 1 - 1/\sqrt{n_{\text{calib}}}$, where n_{calib} is the number of measurements used for each calibration state. If the readout noise is large, one may in addition manually tune this threshold if there are obvious separations of correlator scale. Examples of tuned and untuned dendrograms are shown in Fig. 7. A dendrogram is a visual representation of the clustering process. Read from the bottom upward, each qubit starts in their own node, indicated by each qubit index having its own branch. Moving upward, the nodes are joined by horizontal lines at their characteristic intercluster distance. The threshold value indicated at what distance between clusters the clustering stops, and the noise clusters can be read off by following each branch downward. In Fig. 7 each noise cluster is given a distinct color. If noise clusters exceed the maximum size n_{corr} after applying the threshold, as shown in the top plot, they are iteratively split by following dendrogram branches downward to the first intersection point, dividing the cluster into two smaller ones. This does not affect the threshold for other clusters.

The numerical implementation used the scipy hierarchy package with the "complete" linkage method [51]. The "complete" linkage method groups together clusters by minimizing the largest distance between any of the qubits in the new joined cluster. In Sec. IV B 2, the "average" linkage was used instead due to a limitation of the correlation coefficients. The average linkage method instead groups together qubits based on the lowest average intercluster distances. We also found very good error-mitigation performance using the "ward" method which generally produced dendrograms where it was easier to visually see the separation of correlation scale when correlations were large. However, we advise against its use, as the Ward method is only guaranteed to work with Euclidean distances [50], while the distance described in Eq. (C1) is non-Euclidean.

Limitation on the coherent error model

The coherent error model used in Sec. IV B 2 used the "average" linkage method, as opposed to the "complete" linkage method used for all other simulations. This choice stems from the way the noise model creates correlations among qubits within three- and four-qubit noise clusters. To demonstrate this problem, we will specifically analyze the case of three qubits.

The error unitary for a three-qubit system reads as

$$\begin{aligned} R_x^3(\theta) &= e^{-i(X\otimes X\otimes \mathbb{1}_2 + \mathbb{1}_2\otimes X\otimes X)\theta/2} \\ &= \cos^2\left(\frac{\theta}{2}\right)\mathbb{1}_8 + \sin^2\left(\frac{\theta}{2}\right) X\otimes \mathbb{1}_2\otimes X \\ &\quad - i\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right)(X\otimes X\otimes \mathbb{1}_2 + \mathbb{1}_2\otimes X\otimes X), \end{aligned} \tag{C2}$$

where a matrix exponential expansion was performed. The second term, which correlates the first and third qubits, is suppressed by a factor $\sin^2\left(\frac{\theta}{2}\right)$, which, at small θ, makes the correlation coefficients for these two qubits small, while the correlation between the first and second and the second and third remain large. This causes issues using the "complete" linkage method, which optimizes the smallest largest distance between all qubits in the clusters that should be joined. Therefore, if two qubits have very weak correlations, they are very unlikely to be combined into a noise cluster, even though they are very correlated to a mutual qubit. This, combined with the drawbacks of the correlation coefficients discussed in Appendix A, makes it very unlikely that all three qubits are clustered together. A solution to this was to change the linkage method to "average" which instead optimizes for the smallest average distance between qubits.

APPENDIX D: NECESSITY OF CONNECTED NOISE CLUSTERS

This appendix demonstrates that the correlated QREM protocol, outlined in Sec. III, is the minimal procedure required to mitigate both correlated and coherent readout errors. To begin with, it is convenient to introduce a practical representation for the readout error-mitigated quantum states. When it is necessary to denote the subsystem to which operators belong, superscripts with capital letters are used.

1. Representing the readout error-mitigated state

A density matrix can be represented as a linear combination of any complete basis of Hermitian operators. An informationally complete POVM forms such a basis. Using the index sum notation, the density matrix can be decomposed as

$$\rho = c_i M_i, \tag{D1}$$

where c_i is some complex coefficient. An expression for c_i can be found by using the probability of receiving outcome i from a measurement,

$$p_i = \mathrm{Tr}(M_i\rho) = \mathrm{Tr}(M_i M_j)c_j = T_{ij}c_j, \tag{D2}$$

where we have defined the overlap matrix $T_{ij} = \mathrm{Tr}(M_i M_j)$. Solving Eq. (D2) for c and inserting it into Eq. (D1) one finds a linear inversion representation of the density matrix

$$\rho = p_i (T^{-1})_{ij} M_j. \tag{D3}$$

This representation is convenient as it directly relates the quantum state to the measurement operators and the measurement outcomes $p_i = n_i/N$, where n_i is the recorded number outcome i, and N is the total number of measurements.

If N measurements were performed by a noisy measurement device $\tilde{\mathbf{M}}$, a linear inversion estimator for the prepared state would be

$$\hat{\rho} = \frac{n_i}{N}(\tilde{T}^{-1})_{ij}\tilde{M}_j. \tag{D4}$$

Therefore, if the noisy POVM $\tilde{\mathbf{M}}$ is known, Eq. (D4) is an estimator of the readout error-mitigated state. In the limit of infinite measurements, linear inversion aligns with both the maximum likelihood estimator and the Bayesian mean estimator, which are used in our protocols. Therefore, in the remainder of this section, we assume $n_i/N \rightarrow p_i = \mathrm{Tr}(\rho\tilde{M}_i)$.

2. Minimality of correlated QREM protocol

The necessity of the correlated QREM protocol can be phrased as an observation.

Observation D1. If no assumptions about the prepared state can be made, then readout error mitigation must be performed on the connected noise cluster.

This observation states that an error-mitigated state can generally not be reduced beyond the connected noise clusters of an observable without losing accuracy. We will show this by considering a simplified three-qubit case illustrated in Fig. 8. The noisy correlated readout is described by two error channels $\mathcal{E}_1$ and $\mathcal{E}_2$ which operate on a one-qubit and two-qubit Hilbert space, respectively, and the prepared state is a general three-qubit state. The process can be written as follows:

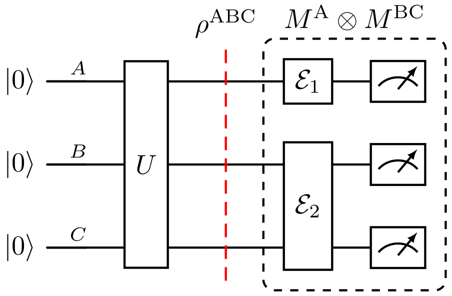


FIG. 8. Example quantum circuit with a prepared general three-qubit sate between correlated qubit readout.

$$\begin{aligned}|000\rangle \rightarrow & \rho^{\mathrm{ABC}} = U|000\rangle\langle 000|U^\dagger \\ \rightarrow & \mathcal{E}_1 \otimes \mathcal{E}_2(\rho^{\mathrm{ABC}}).\end{aligned} \tag{D5}$$

We want to show that if a state is sampled by a noisy POVM $\mathbf{M} = \{M^{\mathrm{A}} \otimes M^{\mathrm{BC}}\}$ (described by $\mathcal{E}_1$ and $\mathcal{E}_2$), we will not obtain the correct readout error-mitigated state ρ^{AB} if we do not first reconstruct the full state ρ^{ABC}, which spans the full connected noise cluster. We do this by comparing the reconstruction of $\mathrm{Tr}_C(\rho^{\mathrm{ABC}})$ to a reconstruction of ρ^{AB} that is aware of the measurement outcomes on the C subsystem.

It is useful to define a reduced measurement operator which retains the information about the outcome on the traced-out subsystem as

$$\mathrm{Tr}_{\mathrm{B}}\big(M^{\mathrm{AB}}_{ij}\big) = M^{\mathrm{A|B}}_{ij}, \tag{D6}$$

which is the measurement operator for outcome i on subsystem A conditioned on outcome j occurring on subsystem B.

These conditioned measurement operators occur naturally when the C subsystem is traced out from the fully reconstructed state on ABC from Fig. 8. To see this consider the full ρ^{ABC},

$$\rho^{\mathrm{ABC}} = p^{\mathrm{ABC}}_{ijk}\big(T^{\mathrm{ABC}^{-1}}\big)_{ijkmno} M^{\mathrm{A}}_m \otimes M^{\mathrm{BC}}_{no}, \tag{D7}$$

and then trace out the subsystem C,

$$\begin{aligned}\mathrm{Tr}_{\mathrm{C}}(\rho^{\mathrm{ABC}}) &= p^{\mathrm{ABC}}_{ijk}\big(T^{\mathrm{ABC}^{-1}}\big)_{ijkmno} M^{\mathrm{A}}_m \otimes \mathrm{Tr}_{\mathrm{C}}\big(M^{\mathrm{BC}}_{no}\big) \\ &= p^{\mathrm{ABC}}_{ijk}\big(T^{\mathrm{ABC}^{-1}}\big)_{ijkmno} M^{\mathrm{A}}_m \otimes M^{\mathrm{B|C}}_{no}.\end{aligned} \tag{D8}$$

If one instead starts directly with the conditioned measurement operators $\mathbf{M} = \{M^{\mathrm{A}} \otimes M^{\mathrm{B|C}}\}$, which primarily operates in the AB subsystem while keeping track of the outcomes that occurred in the environment, one has

$$\rho^{\mathrm{AB|C}} = p^{\mathrm{ABC}}_{ijk}\big(T^{\mathrm{AB|C}^{-1}}\big)_{ijkmno} M^{\mathrm{A}}_m \otimes M^{\mathrm{B|C}}_{no}. \tag{D9}$$

The only difference between Eqs. (D8) and (D9) is the inverse overlap matrix, which in general

$$\begin{aligned}T^{\mathrm{AB|C}}_{ijkmno} &= \mathrm{Tr}_{\mathrm{AB}}\big(\big(M^{\mathrm{A}}_i \otimes M^{\mathrm{B|C}}_{jk}\big)\big(M^{\mathrm{A}}_m \otimes M^{\mathrm{B|C}}_{no}\big)\big) \\ &\neq T^{\mathrm{ABC}}_{ijkmno} = \mathrm{Tr}_{\mathrm{ABC}}\big(\big(M^{\mathrm{A}}_i \otimes M^{\mathrm{BC}}_{jk}\big)\big(M^{\mathrm{A}}_m \otimes M^{\mathrm{BC}}_{no}\big)\big)\end{aligned} \tag{D10}$$

because, for general matrices Q and R

$$\mathrm{Tr}_{\mathrm{AB}}(QR) \neq \mathrm{Tr}_{\mathrm{A}}[\mathrm{Tr}_{\mathrm{B}}(Q)\mathrm{Tr}_{\mathrm{B}}(R)]. \tag{D11}$$

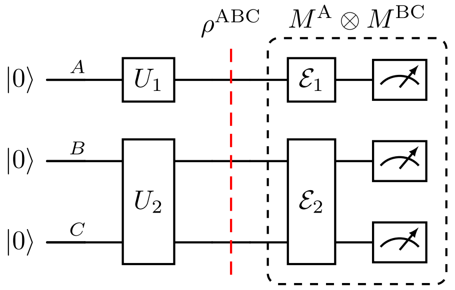


FIG. 9. Example quantum circuit with a product state between correlated qubit readout.

Since the matrices are not equal, their inverse cannot be equal. Therefore, one cannot recover the reduced error-mitigated state if one does not consider the full state between the two error clusters.

Observation D 1 can be understood as a worst-case state reconstruction requirement to get any error-mitigated observable. That is, if the largest cluster size allowed is n_{corr} qubits, then the largest state that might need to be reconstructed for a two-qubit observables is a $2n_{\mathrm{corr}}$-qubit state.

3. Product state assumption

Significant simplifications are possible if some structure is known about the prepared state. In particular, if the prepared state is a product between noise clusters, readout error mitigation can be performed on the noise clusters separately, which means that no connected noise clusters are necessary.

To show this, we need the following property of the overlap matrix.

Proposition D 1. If a POVM is a product of two smaller POVMs, $\mathbf{M} = \{M^{\mathrm{A}} \otimes M^{\mathrm{B}}\}$, then

$$\left(T^{\mathrm{AB}^{-1}}\right)_{ijmn} = \left(T^{\mathrm{A}^{-1}}\right)_{im}\left(T^{\mathrm{B}^{-1}}\right)_{jn}. \tag{D12}$$

Proof. The overlap matrix for a product POVM can be written out explicitly,

$$\begin{aligned} T^{\mathrm{AB}}_{ijmn} &= \mathrm{Tr}\left(\left(M_i^{\mathrm{A}} \otimes M_j^{\mathrm{B}}\right)\left(M_m^{\mathrm{A}} \otimes M_n^{\mathrm{B}}\right)\right) \\ &= \mathrm{Tr}\left(M_i^{\mathrm{A}} M_m^{\mathrm{A}}\right)\mathrm{Tr}\left(M_j^{\mathrm{B}} M_n^{\mathrm{B}}\right) \\ &= T^{\mathrm{A}}_{im} T^{\mathrm{B}}_{jn}. \end{aligned} \tag{D13}$$

From this one can then directly conclude that

$$T^{\mathrm{AB}}_{ijmn}\left(T^{\mathrm{A}^{-1}}\right)_{mk}\left(T^{\mathrm{B}^{-1}}\right)_{nl} = \mathbb{1}_{ijkl}. \tag{D14}$$

■

Consider the simplified three-qubit example in Fig. 9. It is known that the prepared state is a product state between the noise clusters. We want to show that reconstructing the states for each noise cluster individually, $\rho^{\mathrm{A}} \otimes \rho^{\mathrm{BC}}$, is equivalent to reconstructing the whole ρ^{ABC}.

The two readout error-mitigated states can be written as follows:

$$\begin{aligned} \rho^{\mathrm{A}} &= p_i^{\mathrm{A}}\left(T^{\mathrm{A}^{-1}}\right)_{im} M_m^{\mathrm{A}}, \\ \rho^{\mathrm{BC}} &= p_{jk}^{\mathrm{BC}}\left(T^{\mathrm{BC}^{-1}}\right)_{jkno} M_{no}^{\mathrm{A}}. \end{aligned} \tag{D15}$$

Using Eq. (D12), we can combine the two

$$\rho^{\mathrm{A}} \otimes \rho^{\mathrm{BC}} = p_i^{\mathrm{A}} p_{jk}^{\mathrm{BC}}\left(T^{\mathrm{ABC}^{-1}}\right)_{ijkmno} M_m^{\mathrm{A}} \otimes M_{no}^{\mathrm{BC}}. \tag{D16}$$

Since the prepared state is a product state of subsystems A and BC, $p_i^{\mathrm{A}} p_{jk}^{\mathrm{BC}} = p_{ijk}^{\mathrm{ABC}}$. Thus, $\rho^{\mathrm{A}} \otimes \rho^{\mathrm{BC}} = \rho^{\mathrm{ABC}}$.

APPENDIX E: DETAILS ON THE PROTOCOL

In this appendix, we provide implementation details on various parts of the protocol, including measurement simulations and POVM reconstruction. In addition, a pseudocode representation of the correlated QREM protocol in Fig. 1 is provided.

1. Simulating measurements

The process of simulating the measurements draws inspiration from the way measurements are performed on superconducting qubit chips. The native measurement is the computational basis measurement. In order to achieve an informationally complete (IC) measurement for a single qubit, the qubit must be rotated into the x basis and y basis, as well as left undisturbed. This process effectively corresponds to the computational measurements performed along the x, y, and z axes, respectively. The natural way to simulate readout noise is to have it act only in the computational basis readout, and not include rotation gate errors. This closely resembles superconducting qubits, as individual qubit gates can be executed with greater precision than reading out the computational basis. Simulating measurements in this way simplifies the mitigation step, as performing detector tomography on the computational basis is sufficient. The mitigated IC POVM is created by applying ideal rotations to the mitigated computational basis POVM. Note that this does not mean that only classical errors are simulated, each basis is characterized and simulated with coherent errors.

Due to noncommutativity of unitary operators, the application of noise channels such as the one in Sec. IV B 2 to the computational basis does not yield the same IC POVM as applying the noise channel directly to an IC POVM such as the Pauli-6 POVM.

2. State reconstruction

Within the REMST protocol, quantum state tomography (QST) is essential to mitigate coherent errors. In the correlated QREM protocol, QST is performed for every connected noise cluster. The states are reconstructed using an iterative maximum-likelihood algorithm [40]. The likelihood function used for the connected noise clusters is defined as

$$\mathcal{L}(\rho) \propto \Pi_i \mathrm{Tr}\left(\rho M_i^{C}\right)^{n_i}, \tag{E1}$$

where n_i is the observed number of outcomes i and $\mathbf{M}^{C}$ is the reconstructed connected noise-cluster POVM.

The primary limitation affecting the protocol's numerical runtime is quantum state reconstruction. The maximum state size that can be potentially reconstructed is a $2n_{\mathrm{corr}}$-qubit state, whereas the largest POVM involves only n_{corr} qubits. Consequently, reconstructing the POVM is much quicker than reconstructing the state. Thus, to determine an adequate

TABLE II. Recorded runtimes for connected cluster reconstructions for various numbers of qubits in the cluster. These measurements were conducted on a laptop with 10 cores, using iterative MLE [40].

Number of qubits	Approximate runtime
3	1 s
4	30 s
5	2.5 min
6	15 min
7	30 min
8	14 h

maximum cluster size, the runtimes of QST were examined. Table II displays the numerical runtimes for state reconstruction in three to eight qubits. These runtimes were recorded on a 10-core laptop with a 2.5-GHz clock frequency. It can be concluded that setting $n_{\mathrm{corr}} = 3$ results in satisfactory runtimes suitable for the benchmarking performed in this study. In the existing implementation, choosing $n_{\mathrm{corr}} = 4$ is possible, albeit challenging. Substantial literature on more refined state reconstruction approaches (see, e.g., Refs. [52,53]) suggests these could considerably accelerate reconstruction times compared to the current setup; however, they probably would not permit much larger cluster sizes.

3. Detector tomography

In this work we follow the prescription given by Ref. [38]. Detector tomography for a n-qubit system is performed by preparing and measuring an informationally complete set of calibration states, ρ_s, M times and solving the set of equations

$$\frac{m_{si}}{M} = \mathrm{Tr}(\rho_s M_i), \tag{E2}$$

where m_{si} is the number of times outcome i was observed when measuring calibration state ρ_s. The calibration states are all possible combinations of tensor products of the single-qubit calibration states $\{|\psi_i\rangle\langle\psi_i|\}_{i=1}^{4}$. Explicitly, we use pure states pointing to the corners of a tetrahedron in the Bloch sphere,

$$\begin{aligned}
|\psi_1\rangle &= |0\rangle,\\
|\psi_2\rangle &= \frac{1}{\sqrt{3}}|0\rangle + \sqrt{\frac{2}{3}}|1\rangle,\\
|\psi_3\rangle &= \frac{1}{\sqrt{3}}|0\rangle + \sqrt{\frac{2}{3}}e^{i\frac{2\pi}{3}}|1\rangle,\\
|\psi_4\rangle &= \frac{1}{\sqrt{3}}|0\rangle + \sqrt{\frac{2}{3}}e^{i\frac{4\pi}{3}}|1\rangle.
\end{aligned} \tag{E3}$$

The four calibration states form a set of symmetric and informationally complete rank-1 projectors for a single qubit.

In order to have a fair comparison between all protocols, we have opted to use 3-local PHF to generate the calibration states. Using a 3-local PHF guarantees that any three-qubit POVM can be reconstructed, which is the largest used for the results presented in this work. As an example of how the calibration states are generated, Table III shows the calibration states written explicitly for the first four qubits using the first hash function of the 3-local PHF in Table I. Each hash function assigns one of three possible values to each qubit. An informationally complete set of calibration states is prepared between the three groups of qubits, where each qubit with label $(0, 1, 2)$ is assigned $\{(\psi_1, \psi_1, \psi_2), (\psi_1, \psi_1, \psi_3), (\psi_1, \psi_2, \psi_1), \ldots, (\psi_3, \psi_3, \psi_2)\}$. The cases where all qubits receive the same calibration states are removed from each hash function and measured at the end.

TABLE III. Shortened list of the 60 calibration states generated from the first hash function in $\Phi_{15,16,3}$ PHF, focusing on the initial four qubits. The notation $|i\rangle$ is an abbreviation for the calibration state $|\psi_i\rangle$.

Qubit index	1	2	3	4
Hash value	0	1	2	0
1	$\|1\rangle$	$\|0\rangle$	$\|0\rangle$	$\|1\rangle$
2	$\|2\rangle$	$\|0\rangle$	$\|0\rangle$	$\|2\rangle$
3	$\|3\rangle$	$\|0\rangle$	$\|0\rangle$	$\|3\rangle$
4	$\|0\rangle$	$\|1\rangle$	$\|0\rangle$	$\|0\rangle$
5	$\|1\rangle$	$\|1\rangle$	$\|0\rangle$	$\|1\rangle$
6	$\|2\rangle$	$\|1\rangle$	$\|0\rangle$	$\|2\rangle$
7	$\|3\rangle$	$\|1\rangle$	$\|0\rangle$	$\|3\rangle$
8	$\|0\rangle$	$\|2\rangle$	$\|0\rangle$	$\|0\rangle$
9	$\|1\rangle$	$\|2\rangle$	$\|0\rangle$	$\|1\rangle$
10	$\|2\rangle$	$\|2\rangle$	$\|0\rangle$	$\|2\rangle$
⋮	⋮	⋮	⋮	⋮
57	$\|3\rangle$	$\|2\rangle$	$\|3\rangle$	$\|3\rangle$
58	$\|0\rangle$	$\|3\rangle$	$\|3\rangle$	$\|0\rangle$
59	$\|1\rangle$	$\|3\rangle$	$\|3\rangle$	$\|1\rangle$
60	$\|2\rangle$	$\|3\rangle$	$\|3\rangle$	$\|2\rangle$

4. Details on the other protocols

In Sec. IV, alongside the correlated QREM protocol, three alternative protocols were examined. We provide further details on how each of these protocols diverges from the full correlated QREM protocol.

a. Factorized QREM

The facotrized QREM protocol only mitigated individual qubit readout errors. From the detector tomography measurements shared between all protocols, the noisy POVM is reconstructed for each qubit. For a requested two-qubit observable, the POVM for the two qubits entering the observable are tensored together and used as the reconstructed noisy POVM in the REMST protocol [25].

b. Two-point QREM

The two-point QREM protocol is a naive implementation of the REMST protocol [25]. For a given two-qubit observable, the two qubits entering the observable are isolated, and the noisy POVMs are reconstructed for the two-qubit system. In this way correlated readout between the two qubits is captured, but no other correlations the two qubits may have are mitigated. A drawback with using this method is that effects from other correlated qubits are incorrectly assumed to be due to correlations between the two qubits, which can cause the REMST protocol to perform worse than performing no QREM (see, for example, the right side of Fig. 6).

c. Classical correlated QREM

The classical correlated QREM protocol operates exactly like the correlated QREM protocol; the only difference is that once the noise-cluster POVMs are reconstructed, the off-diagonal elements in the computational basis are set to zero. In this way, only statistical distributions of the different outcomes are captured, which is equivalent to using a confusion matrix [15].

5. Pseudocode for the correlated QREM protocol

A pseudocode representation of the two primary components of the correlated QREM protocol is presented. Noise characterization is detailed in Algorithm 1, while observable reconstruction is described in Algorithm 2.

ALGORITHM 1. Correlated noise structure characterization.

```
Input:
$\Phi$: 2-local perfect hash family
$\{\psi_1, \psi_2, \psi_3, \psi_4\}$: Single-qubit calibration states
Output:
$s$: List of noise clusters
M: List of reconstructed noise-cluster POVMs

for each hash function $\phi$ in $\Phi$ do
    for each qubit $k$ do
        Assign label $\phi(k)$ to qubit $k$
    end for
    Assign calibration state $\{(\psi_1, \psi_2), (\psi_1, \psi_2), \dots, (\psi_4, \psi_3)\}$
      to qubits with label $(1, 2)$
    Measure calibration states
end for
for each calibration state $\psi_n$ do
    Measure state $\psi_n$ on all qubits
end for
Create list $\alpha$ of all possible qubit pairs
for each pair $i$ in $\alpha$ do
    Reconstruct POVM $\mathbf{M}_i$
    Compute correlation coefficient $c_i$ using $\mathbf{M}_i$
end for
$s \leftarrow$ hierarchical clustering using $c$
for each noise cluster $j$ in $s$ do
    Measure informationally complete set of calibration states
      on noise cluster $j$
    $\mathbf{M}_j \leftarrow$ reconstruct POVM on noise cluster $j$
end for
```

ALGORITHM 2. Reconstructing readout error-mitigated observables.

```
Input:
$O$: List of requested observables
$s$: List of noise clusters
M: List of reconstructed noise-cluster POVMs
Output:
$O^{\mathrm{QREM}}$: List of readout error-mitigated observables
$\rho^{\mathrm{QREM}}$: List of readout error-mitigated states

for each observable $O_i$ in $O$ do
    Find connected noise cluster $C_i$ from $s$
    Create connected noise-cluster POVM $M_i^C$ using M and $C_i$
    $\rho_i^{\mathrm{QREM}} \leftarrow$ Perform REMST [25] on $C_i$ using $M_i^C$
    $O_i^{\mathrm{QREM}} \leftarrow \mathrm{Tr}(\rho_i^{\mathrm{QREM}} O_i)$
end for
```

APPENDIX F: ADDITIONAL RESULTS

1. Coherent error model

Figure 10 illustrates another simulation from Sec. IV B 2, using the "complete" method for the hierarchical clustering linkage map. This method optimizes the clustering obtained based on the minimal "largest distance" between the nodes that should be joined. As expected, the correlated QREM protocol performs significantly worse than when the "average" linkage map is used, which optimized joined nodes with the smallest average distance. This comes from the fact that only two-qubit correlated errors are correctly captured and three- and four-qubit correlations are missed. This issue is discussed in more detail in Appendix C 1.

2. 100 qubits with mixed target states

Previous simulations have used random pure states as target states. Although most experiments focus on generating pure states, circuit errors can result in output states being mixed even before noisy measurements are performed. To confirm that our protocol also works well for mixed states, we reran the simulations from Sec. IV C using mixed target states, with the findings depicted in Fig. 11. The target states

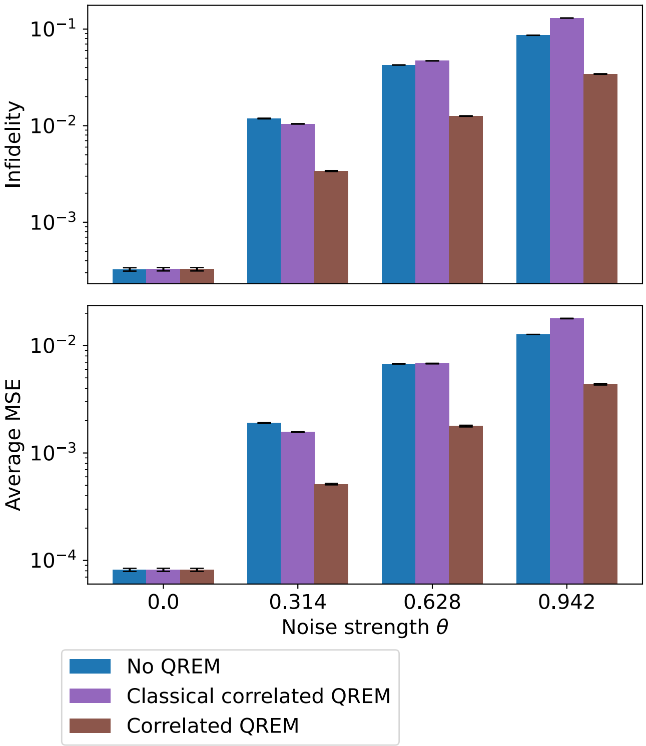


FIG. 10. The coherent noise example from Sec. II C with the "complete" linkage method. Everything else is identical to Fig. 5.

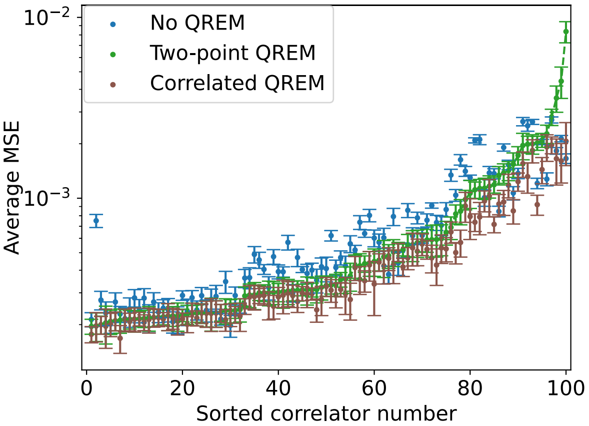


FIG. 11. 100-qubit example with mixed target states. Everything else is identical to Fig. 6.

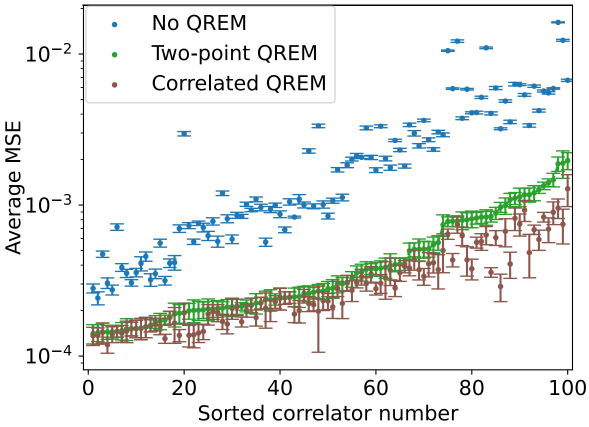


FIG. 12. 100-qubit example with the identity operator included in the list of Pauli operators. Everything else is identical to Fig. 6.

are generated from Haar-random pure states by applying a depolarizing channel. For each four-qubit pure state, the channel is applied with a strength p uniformly selected from the range $[0.03, 0.07]$, leading to states with an average purity of approximately 0.9. The four-qubit depolarizing channel is given by

$$\mathcal{E}(\rho) = (1-p)\rho + \frac{p}{2^4}\mathbb{1}. \tag{F1}$$

The overall performance for all readout error-mitigation methods considered is slightly better than with the pure target states in Fig. 6, but qualitatively the same. This is expected since pure states are harder to estimate than mixed states [54], resulting in a higher overall state reconstruction accuracy for mixed states.

3. 100-qubit simulation including single-qubit observables

Throughout the simulations presented in Sec. IV, the MSE was computed for all two-qubit Pauli observables and averaged to give representative protocol performance. In Fig. 12 the 100-quibt example is recomputed with the identity operator included in the list of Pauli observables. The effect of including the identity operator is that the single-qubit observables are also averaged in the MSE. A slight decrease in the performance of the no QREM protocol can be seen, while both the two-point and correlated QREM protocols improve significantly on average, indicating that both protocols are also well suited to capture all lower-order observables.

[1] E. Altman, K. R. Brown, G. Carleo, L. D. Carr, E. Demler, C. Chin, B. DeMarco, S. E. Economou, M. A. Eriksson, K.-M. C. Fu, M. Greiner, K. R. Hazzard, R. G. Hulet, A. J. Kollár, B. L. Lev, M. D. Lukin, R. Ma, X. Mi, S. Misra, C. Monroe *et al.* Quantum simulators: Architectures and opportunities, PRX Quantum **2**, 017003 (2021).

[2] K. Temme, S. Bravyi, and J. M. Gambetta, Error mitigation for short-depth quantum circuits, Phys. Rev. Lett. **119**, 180509 (2017).

[3] X. Bonet-Monroig, R. Sagastizabal, M. Singh, and T. E. O'Brien, Low-cost error mitigation by symmetry verification, Phys. Rev. A **98**, 062339 (2018).

[4] E. van den Berg, Z. K. Minev, A. Kandala, and K. Temme, Probabilistic error cancellation with sparse Pauli–Lindblad models on noisy quantum processors, Nat. Phys. **19**, 1116 (2023).

[5] Z. Cai, R. Babbush, S. C. Benjamin, S. Endo, W. J. Huggins, Y. Li, J. R. McClean, and T. E. O'Brien, Quantum error mitigation, Rev. Mod. Phys. **95**, 045005 (2023).

[6] Y. Quek, D. Stilck França, S. Khatri, J. J. Meyer, and J. Eisert, Exponentially tighter bounds on limitations of quantum error mitigation, Nat. Phys. **20**, 1648 (2024).

[7] Z. Zimborás, B. Koczor, Z. Holmes, E.-M. Borrelli, A. Gilyén, H.-Y. Huang, Z. Cai, A. Acín, L. Aolita, L. Banchi, F. G. S. L. Brandão, D. Cavalcanti, T. Cubitt, S. N. Filippov, G. García-Pérez, J. Goold, O. Kálmán, E. Kyoseva, M. A. C. Rossi, B. Sokolov, I. Tavernelli, and S. Maniscalco, Myths around quantum computation before full fault tolerance: What no-go theorems rule out and what they don't, arXiv:2501.05694.

[8] J. Tuziemski, F. B. Maciejewski, J. Majsak, O. Słowik, M. Kotowski, K. Kowalczyk-Murynka, P. Podziemski, and M. Oszmaniec, Efficient reconstruction, benchmarking and validation of cross-talk models in readout noise in near-term quantum devices, arXiv:2311.10661.

[9] Z. Zhou, R. Sitler, Y. Oda, K. Schultz, and G. Quiroz, Quantum crosstalk robust quantum control, Phys. Rev. Lett. **131**, 210802 (2023).

[10] M. R. Geller and M. Sun, Toward efficient correction of multiqubit measurement errors: pair correlation method, Quantum Sci. Technol. **6**, 025009 (2021).

[11] M. Sarovar, T. Proctor, K. Rudinger, K. Young, E. Nielsen, and R. Blume-Kohout, Detecting crosstalk errors in quantum information processors, Quantum **4**, 321 (2020).

[12] J. Heinsoo, C. K. Andersen, A. Remm, S. Krinner, T. Walter, Y. Salathé, S. Gasparinetti, J.-C. Besse, A. Potočnik, A. Wallraff, and C. Eichler, Rapid high-fidelity multiplexed readout of superconducting qubits, Phys. Rev. Appl. **10**, 034040 (2018).

[13] P. C. de Groot, A. F. van Loo, J. Lisenfeld, R. N. Schouten, A. Lupaşcu, C. J. P. M. Harmans, and J. E. Mooij, Low-crosstalk bifurcation detectors for coupled flux qubits, Appl. Phys. Lett. **96**, 123508 (2010).

[14] B. Lienhard, A. Vepsäläinen, L. C. G. Govia, C. R. Hoffer, J. Y. Qiu, D. Ristè, M. Ware, D. Kim, R. Winik, A. Melville, B. Niedzielski, J. Yoder, G. J. Ribeill, T. A. Ohki, H. K. Krovi, T. P. Orlando, S. Gustavsson, and W. D. Oliver, Deep-neural-network discrimination of multiplexed superconducting-qubit states, Phys. Rev. Appl. **17**, 014024 (2022).

[15] F. B. Maciejewski, Z. Zimborás, and M. Oszmaniec, Mitigation of readout noise in near-term quantum devices by classical post-processing based on detector tomography, Quantum **4**, 257 (2020).

[16] P. D. Nation, H. Kang, N. Sundaresan, and J. M. Gambetta, Scalable mitigation of measurement errors on quantum computers, PRX Quantum **2**, 040326 (2021).

[17] B. Nachman, M. Urbanek, W. A. de Jong, and C. W. Bauer, Unfolding quantum computer readout noise, npj Quantum Inf. **6**, 84 (2020).

[18] H.-Y. Huang, R. Kueng, and J. Preskill, Predicting many properties of a quantum system from very few measurements, Nat. Phys. **16**, 1050 (2020).

[19] S. Chen, W. Yu, P. Zeng, and S. T. Flammia, Robust shadow estimation, PRX Quantum **2**, 030348 (2021).

[20] H. Jnane, J. Steinberg, Z. Cai, H. C. Nguyen, and B. Koczor, Quantum error mitigated classical shadows, PRX Quantum **5**, 010324 (2024).

[21] D. E. Koh and S. Grewal, Classical shadows with noise, Quantum **6**, 776 (2022).

[22] A. Arrasmith, A. Patterson, A. Boughton, and M. Paini, Development and demonstration of an efficient readout error mitigation technique for use in NISQ algorithms, arXiv:2303.17741.

[23] H.-Y. Hu, A. Gu, S. Majumder, H. Ren, Y. Zhang, D. S. Wang, Y.-Z. You, Z. Minev, S. F. Yelin, and A. Seif, Demonstration of robust and efficient quantum property learning with shallow shadows, Nat. Commun. **16**, 2943 (2025).

[24] E. Onorati, J. Kitzinger, J. Helsen, M. Ioannou, A. H. Werner, I. Roth, and J. Eisert, Noise-mitigated randomized measurements and self-calibrating shadow estimation, arXiv:2403.04751.

[25] A. S. Aasen, A. Di Giovanni, H. Rotzinger, A. V. Ustinov, and M. Gärttner, Readout error mitigated quantum state tomography tested on superconducting qubits, Commun. Phys. **7**, 301 (2024).

[26] C. Ding, X.-Y. Xu, Y.-F. Niu, S. Zhang, W.-S. Bao, and H.-L. Huang, Noise-resistant quantum state compression readout, Sci. China Phys. Mech. Astron. **66**, 230311 (2023).

[27] H.-L. Huang and C. Ding, Quantum state compression shadow, Eur. Phys. J.: Spec. Top (2025), doi:10.1140/epjs/s11734-025-01715-8.

[28] J. S. Lundeen, A. Feito, H. Coldenstrodt-Ronge, K. L. Pregnell, C. Silberhorn, T. C. Ralph, J. Eisert, M. B. Plenio, and I. A. Walmsley, Tomography of quantum detectors, Nat. Phys. **5**, 27 (2009).

[29] J. Cotler and F. Wilczek, Quantum overlapping tomography, Phys. Rev. Lett. **124**, 100401 (2020).

[30] F. B. Maciejewski, F. Baccari, Z. Zimborás, and M. Oszmaniec, Modeling and mitigation of cross-talk effects in readout noise with applications to the quantum approximate optimization algorithm, Quantum **5**, 464 (2021).

[31] I. Georgescu, S. Ashhab, and F. Nori, Quantum simulation, Rev. Mod. Phys. **86**, 153 (2014).

[32] A. Keesling, A. Omran, H. Levine, H. Bernien, H. Pichler, S. Choi, R. Samajdar, S. Schwartz, P. Silvi, S. Sachdev, P. Zoller, M. Endres, M. Greiner, V. Vuletić, and M. D. Lukin, Quantum Kibble–Zurek mechanism and critical dynamics on a programmable Rydberg simulator, Nature (London) **568**, 207 (2019).

[33] J. Zhang, S. H. Cantú, F. Liu, A. Bylinskii, B. Braverman, F. Huber, J. Amato-Grill, A. Lukin, N. Gemelke, A. Keesling, S.-T. Wang, Y. Meurice, and S.-W. Tsai, Probing quantum floating phases in Rydberg atom arrays, Nat. Commun. **16**, 712 (2025).

[34] A. Schuckert, O. Katz, L. Feng, E. Crane, A. De, M. Hafezi, A. V. Gorshkov, and C. Monroe, Observation of a finite-energy phase transition in a one-dimensional quantum simulator, Nat. Phys. **21**, 374 (2025).

[35] P. Richerme, Z.-X. Gong, A. Lee, C. Senko, J. Smith, M. Foss-Feig, S. Michalakis, A. V. Gorshkov, and C. Monroe, Non-local propagation of correlations in quantum systems with long-range interactions, Nature (London) **511**, 198 (2014).

[36] P. Jurcevic, B. P. Lanyon, P. Hauke, C. Hempel, P. Zoller, R. Blatt, and C. F. Roos, Quasiparticle engineering and entanglement propagation in a quantum many-body system, Nature (London) **511**, 202 (2014).

[37] A. Di Giovanni, A. S. Aasen, J. Lisenfeld, M. Gärttner, H. Rotzinger, and A. V. Ustinov, Correlation-conscious optimization of multiplexed qubit readout (unpublished).

[38] J. Fiurášek, Maximum-likelihood estimation of quantum measurement, Phys. Rev. A **64**, 024102 (2001).

[39] R. Blume-Kohout, Optimal, reliable estimation of quantum states, New J. Phys. **12**, 043034 (2010).

[40] A. I. Lvovsky, Iterative maximum-likelihood reconstruction in quantum homodyne tomography, J. Opt. B: Quantum Semiclassical Opt. **6**, S556 (2004).

[41] G. García-Pérez, M. A. C. Rossi, B. Sokolov, E.-M. Borrelli, and S. Maniscalco, Pairwise tomography networks for many-body quantum systems, Phys. Rev. Res. **2**, 023393 (2020).

[42] X. Bonet-Monroig, R. Babbush, and T. E. O'Brien, Nearly optimal measurement scheduling for partial tomography of quantum states, Phys. Rev. X **10**, 031064 (2020).

[43] R. A. Walker, II and C. J. Colbourn, Perfect hash families: Constructions and existence, J. Math. Cryptol. **1**, 125 (2007).

[44] R. Dougherty, Hash families and applications to t-restrictions, Ph.D. thesis, Arizona State University, 2019.

[45] A. Aasen, CREMST, 2025, https://github.com/AdrianAasen/CREMST.

[46] K. Hansenne, R. Qu, L. T. Weinbrenner, C. de Gois, H. Wang, Y. Ming, Z. Yang, P. Horodecki, W. Gao, and O. Gühne, Optimal overlapping tomography, arXiv:2408.05730.

[47] M. Jerger, S. Poletto, P. Macha, U. Hübner, E. Il'ichev, and A. V. Ustinov, Frequency division multiplexing readout and simultaneous manipulation of an array of flux qubits, Appl. Phys. Lett. **101**, 042604 (2012).

[48] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, A quantum engineer's guide to superconducting qubits, Appl. Phys. Rev. **6**, 021318 (2019).

[49] A. D. Giovanni, A. S. Aasen, J. Lisenfeld, M. Gärttner, H. Rotzinger, and A. V. Ustinov, Multiplexed qubit readout

quality metric beyond assignment fidelity, arXiv:2502.08589 [Phys. Rev. Appl. (to be published)].

[50] D. Müllner, Modern hierarchical, agglomerative clustering algorithms, arXiv:1109.2378.

[51] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey *et al.*, Scipy 1.0: fundamental algorithms for scientific computing in python, Nat. Methods **17**, 261 (2020).

[52] J. Shang, Z. Zhang, and H. K. Ng, Superfast maximum-likelihood reconstruction for quantum tomography, Phys. Rev. A **95**, 062336 (2017).

[53] Z. Hou, H.-S. Zhong, Y. Tian, D. Dong, B. Qi, L. Li, Y. Wang, F. Nori, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, Full reconstruction of a 14-qubit state within four hours, New J. Phys. **18**, 083036 (2016).

[54] D. H. Mahler, L. A. Rozema, A. Darabi, C. Ferrie, R. Blume-Kohout, and A. M. Steinberg, Adaptive quantum state tomography improves accuracy quadratically, Phys. Rev. Lett. **111**, 183601 (2013).

6

GENERALIZED READOUT QUALITY METRIC FOR SUPERCONDUCTING QUBITS

Building on the numerical tools developed for the correlated REMST protocol in Chapter 5, we searched for a direct application to improve the characterization of readout quality in superconducting qubit systems. *Assignment fidelity* is the standard quality metric for superconducting qubits and quantifies the probability of correctly identifying a prepared computational basis states,

$$\mathcal{F} = \frac{p(\text{outcome } 0 \mid |0\rangle) + p(\text{outcome } 1 \mid |1\rangle)}{2}. \tag{6.1}$$

In terms of generalized measurement operators, this becomes

$$\mathcal{F} = \frac{\mathrm{Tr}\left(\tilde{P}_0 \left|0\right\rangle \left\langle 0\right|\right) + \mathrm{Tr}\left(\tilde{P}_1 \left|1\right\rangle \left\langle 1\right|\right)}{2}, \tag{6.2}$$

which is 1 when $\tilde{P}_i$ matches the ideal projectors $P_i = \left|i\right\rangle \left\langle i\right|$. While assignment fidelity is practical, it is limited in that it only partially detects classical readout errors and is typically computed independently for each qubit, leaving out both correlated and coherent errors. Optimizing experiments based exclusively on single-qubit assignment fidelity is therefore suboptimal [F].

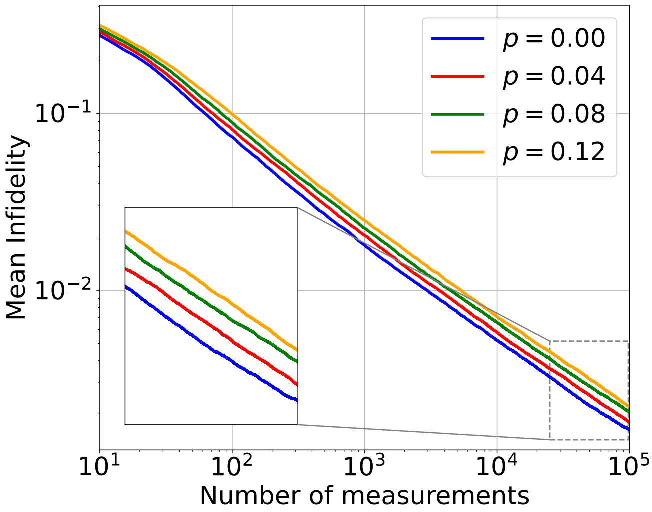


Figure 6.1: Simulated readout-error-mitigated state reconstruction with different levels of depolarizing noise, Equation 2.34, applied at readout.

In this work, we present a more comprehensive readout quality metric based on error-mitigated state reconstruction accuracy. This metric is motivated by the observation that the number of measurements required for an error-mitigated estimator to achieve a specified accuracy increases proportionally with the system's noise level. This phenomenon is visualized in Figure 6.1, where the stronger the depolarizing readout noise is, the more measurements are required to reach the same state reconstruction infidelity threshold. It is therefore possible to use the number of measurements required for an estimator to cross an accuracy threshold as a proxy for how much noise there is in the system.

We compared the new readout quality metric to the standard assignment fidelity using our in-house superconducting qubits and found that they performed similarly at finding optimal readout settings. Additionally, we investigated how a fixed shot budget should be distributed between calibrating the nosy measurements and performing state reconstruction. We found that less shots are required for calibration than a naive estimate based on degrees of freedom of the measurement operators, suggesting that there is structure in the measurement operators that simplifies the estimation task.

Author participation is stated in the acknowledgment of the manuscript: "*A.D.G. performed the experiments with supervision from H.R., J.L., and A.U. The theory and numerics were implemented by* ***A.S.A.*** [the author] *with supervision from M.G. A.D.G. and* ***A.S.A.*** *performed the data analysis. The first manuscript was prepared by A.D.G. and* ***A.S.A.*** *All authors contributed to the finalization of the manuscript* ".

PHYSICAL REVIEW APPLIED **24,** 044043 (2025)

Benchmarking the quality of multiplexed qubit readout beyond assignment fidelity

Andras Di Giovanni,[1,*] Adrian Skasberg Aasen,[2,3] Jürgen Lisenfeld,[1] Martin Gärttner,[3]
Hannes Rotzinger,[1,4] and Alexey V. Ustinov[1,4]

[1] *Physikalisches Institut, Karlsruher Institut für Technologie, Kaiserstraße 12, 76131 Karlsruhe, Germany*
[2] *Kirchhoff-Institut für Physik, Universität Heidelberg, Im Neuenheimer Feld 227, 69120 Heidelberg, Germany*
[3] *Institut für Festkörpertheorie und -optik, Friedrich-Schiller-Universität Jena, Max-Wien-Platz 1, 07743 Jena, Germany*
[4] *Institut für QuantenMaterialien und Technologien, Karlsruher Institut für Technologie, Kaiserstraße 12, 76131 Karlsruhe, Germany*

(Received 21 February 2025; revised 30 May 2025; accepted 15 September 2025; published 15 October 2025)

The accurate measurement of quantum two-level objects (qubits) is crucial for developing quantum computers. Over the last decade, the measure of choice for benchmarking readout routines for superconducting qubits has been assignment fidelity; however, this method only focuses on the preparation of computational basis states and therefore does not provide a complete characterization of the readout. Here, we expand the focus to the use of detector tomography to fully characterize multiqubit readout of superconducting transmon qubits. The impact of different readout parameters on the rate of information extraction is studied using quantum-state reconstruction infidelity as a proxy. The results are then compared with assignment fidelities, showing good agreement for separable two-qubit states. We therefore propose the rate of infidelity convergence as a validation tool for assignment fidelity and a more comprehensive benchmark for single- and multiqubit readout optimization. To make the best use of experimental resources, we investigate the most efficient distribution of a limited shot budget between detector tomography and state reconstruction within the context of single- and two-qubit experiments. To address the growing interest in three-qubit gates and test the scalability of the validation tool, we perform three-qubit quantum-state tomography that goes beyond conventional readout-error-mitigation methods and find a factor-of-30 reduction in quantum infidelity. Our results demonstrate that qubit readout correlations are not induced by a significantly reduced state distinguishability. Consequently, correlation coefficients can serve as a valuable tool in qubit readout optimization.

DOI: 10.1103/dpft-lxtx

I. INTRODUCTION

State-of-the-art quantum computing experiments have started pushing the boundary to fault tolerance [1,2]. Although gate-error-mitigation approaches such as Pauli twirling or zero-noise extrapolation [3,4] remain important for most experiments, these methods have been shown to be of limited use to extract noiseless expectation values [5]. Significant research effort has gone into optimizing single- and multiqubit gate fidelities [6–9]; however, it is becoming increasingly clear that readout errors also pose an important obstacle to general-purpose quantum computing [10]. On the analog quantum simulation side, precise readout is just as important to prevent biasing of reconstructed observables. Therefore, recent research has focused on understanding [11,12] and improving the multiplexed dispersive readout for superconducting qubits [13,14]. In this work, we aim to contribute to understanding and mitigating readout errors with the help of quantum detector tomography [15–18]. This approach assumes that state preparation in the quantum system is accurate, and it seeks to mitigate readout errors by reconstructing the experimental measurement operator instead of assuming that the readout is accurately modeled by projective computational basis measurements.

In particular, correlated readout errors have recently been observed in superconducting qubit systems. References [19,20] consider correlated bitflip errors and energy relaxation, respectively. Both discuss physical mechanisms

---

[*] Contact author: andras.digiovanni@gmail.com

that may occur during readout, effectively manifesting as correlated readout, which would be captured by readout correlation coefficients. A further example could be fluctuations in the mutual electric ground. Numerical simulations of correlated readout errors from experiments and noise models have been reported in Ref. [21]. Furthermore, Ref. [22] discusses methods to decrease electromagnetic crosstalk enabled readout correlations in superconducting flux qubits. Reference [14] provides experimental evidence for qubit dephasing due to correlated readout errors.

We apply a comprehensive readout-error-mitigation protocol, introduced in Ref. [23], to a multiqubit quantum-state tomography experiment. The protocol mitigates both coherent and correlated readout errors, going beyond traditional techniques that address only classical readout errors [17,24,25] and enables very precise state reconstruction. Our experiments are implemented on a superconducting qubit chip with four transmons, which we manipulate and read out using frequency-division multiplexing [26]. This architectural choice simplifies wiring, but it can be prone to significant readout crosstalk, and it is hence an ideal testbed for multiqubit readout-error mitigation. By employing a traveling-wave parametric amplifier (TWPA) [27], we can perform single-shot readout of the qubits. We introduce a readout quality metric that uses the number of shots necessary to reach a specified state reconstruction infidelity, quantifying the efficiency of extracting information from the system. The proposed metric takes into account coherent errors, also referred to as quantum errors, which we compare with assignment fidelities, defined as the probabilities of correctly assigning labels 0 or 1 to the prepared computational basis states.

Given the substantial number of repetitive measurements required for detector tomography, it is important to avoid spending time on overly precise calibration of the reconstructed positive operator-valued measures (POVMs). Therefore, we also investigate different distributions of a fixed shot budget between quantum detector tomography (QDT) and quantum-state tomography (QST) to determine which results in the lowest achievable infidelity. Furthermore, we demonstrate two- and three-qubit density matrix reconstruction that takes into account coherent and correlated readout noise [23], which opens the way to reducing readout errors in complete benchmarking of two- and three-qubit gates. This is motivated by renewed interest in multiqubit entangling gates in the community [28–30].

An advantage of detector tomography is that it explicitly reconstructs the POVM corresponding to the multiqubit readout process. This enables extraction of multiqubit readout correlation coefficients of the system, regardless of their origin, which could include off-resonant drive of the resonator, mechanical vibrations [19], or ionizing radiation [31]. In multiplexed readout, a perfect measurement of a qubit should not depend on the state of its neighbors, which makes readout correlation a suitable indicator of readout quality [22].

The readout correlation coefficients should be independent of the state distinguishability in the readout plane. By lowering the readout pulse amplitude, effectively lowering state distinguishability, we experimentally demonstrate that readout correlation coefficients satisfy this property.

II. PRELIMINARIES

A. Generalized measurements

To characterize a noisy detector, it is essential to adopt a framework that goes beyond projective measurement. This approach can be formulated by a set of operators possessing the following properties:

$$M_i^\dagger = M_i, \quad M_i \geq 0, \quad \text{and} \quad \sum_i M_i = \mathbb{1}, \tag{1}$$

where $\boldsymbol{M} = \{M_i\}$ is the POVM with elements M_i. When taking the expectation values of the measurement operators, one can retrieve the probabilities of obtaining different outcomes i in a measurement process,

$$p_i = \mathrm{Tr}(\rho M_i). \tag{2}$$

B. Detector tomography

Characterization of the noisy measurement device is achieved by performing detector tomography [15]. The goal of detector tomography is to reconstruct the generalized measurement operators $\boldsymbol{M} = \{M_i\}$ that satisfy the constraints in Eq. (1) through experimentally measured outcome probabilities

$$\frac{n_{is}}{N} = p_{is} = \mathrm{Tr}(\rho_s M_i), \tag{3}$$

where N is the total number of measurements and n_{is} is the number of outcomes i observed from calibration state ρ_s. To enable full detector tomography, the set of calibration states must be informationally complete. In our experiment, we chose the symmetric and informationally complete tetrahedron states

$$\begin{aligned} |\psi_1\rangle &= |0\rangle\,, \\ |\psi_2\rangle &= \frac{1}{\sqrt{3}}\,|0\rangle + \sqrt{\frac{2}{3}}\,|1\rangle\,, \\ |\psi_3\rangle &= \frac{1}{\sqrt{3}}\,|0\rangle + \sqrt{\frac{2}{3}}e^{i\frac{2\pi}{3}}\,|1\rangle\,, \\ |\psi_4\rangle &= \frac{1}{\sqrt{3}}\,|0\rangle + \sqrt{\frac{2}{3}}e^{i\frac{4\pi}{3}}\,|1\rangle\,. \end{aligned} \tag{4}$$

With the full set of measurement probabilities in Eq. (3), one can estimate the most likely set of operators through a maximum-likelihood estimator, as outlined, for example, in Ref. [32].

C. Readout-error-mitigated quantum-state tomography

With a reconstruction of a noisy, experimentally measured POVM, it is possible to use the POVM to construct a readout-error-mitigated density matrix. The approach we use is described in Ref. [23] and is based on directly integrating the full set of estimated measurement operators into a density matrix estimator. The protocol is specifically based on estimating the density matrix using the adapted likelihood function,

$$\mathcal{L}_{\boldsymbol{M}^{\text{estm}}}(\rho) \propto \Pi_i \text{Tr}(\rho M_i^{\text{estm}})^{n_i} = \left(\Pi_i (p_i)^{\hat{p}_i}\right)^N, \quad (5)$$

with either a maximum-likelihood estimator [33] or a Bayesian-mean estimator [34].

This approach allows for beyond-classical readout-error mitigation as it naturally accommodates coherent errors, and it does not require any channel inversion or numerical optimization while still guaranteeing physical density matrix reconstruction.

D. Multiplexed qubit readout

Multiplexed readout [26] is a central architectural choice for superconducting quantum processors. By coupling multiple resonators to a single transmission line, one can save circuit footprint on the chip. In conventional transmon processors, transmission lines account for a significant chip surface area. As resonators now share their transmission line, their readout pulses are sent through the cryostat on the same microwave line, and they are also amplified and mixed simultaneously after reading out the qubit state. This architectural choice needs fewer amplifiers and other microwave equipment; however, it has a significant drawback: resonators on a single transmission line cannot be too close in frequency, otherwise they will be measured with crosstalk. A readout pulse synthesized for one resonator will also drive a neighboring one, leading to a bias in the derived qubit state. Recently, deep-learning methods have been explored to increase single-shot assignment fidelity in multiplexed readout operations [35].

E. Quantifying correlated readout errors

A major challenge in scaling up quantum hardware is induced correlations and crosstalk, e.g., due to multiplexed qubit readout. A useful tool for characterizing readout correlations was introduced in Refs. [12,18], where, by using reconstructed two-qubit POVMs, asymmetric correlation coefficients $c_{j\to i}$ could be extracted. These coefficients quantify how much the measurement results from qubit i are affected by the state of qubit j. For a more detailed discussion of the definition of the correlation coefficients, see Appendix A.

Since the correlation coefficients are asymmetric, we also define a symmetric version, $c_{i\leftrightarrow j} = (c_{j\to i} + c_{i\to j})/2$. The coefficients can be modified to measure only the classical correlations, represented as $c^{C}_{i\leftrightarrow j}$, which are determined exclusively by the diagonal elements in the measurement operators. The full correlation coefficients contain the effect of both the diagonal and off-diagonal elements, which we call quantum correlation coefficients, represented by $c^{Q}_{i\leftrightarrow j}$.

F. Infidelity

To quantify the quality of a reconstructed density matrix, we use the quantum infidelity, $I(\rho, \sigma)$, as a figure of merit. In the case of pure target states σ, the infidelity with respect to the estimated state ρ is given by

$$I(\rho, \sigma) = 1 - \text{Tr}(\rho\sigma). \quad (6)$$

III. EXPERIMENTAL REALIZATION

A. Experimental setup

We investigate a device with four transmon qubits [36], each coupled individually to a resonator for readout. The device is measured below 15 mK in a dry dilution cryostat. The frequencies of the qubits and resonators, as well as typical coherence times can be found in Appendix B. We define System A as the two-qubit subsystem comprising q0 and q1, System B as the subsystem comprising q1 and q2, and finally System C as the subsystem comprising q2 and q3, as shown in Fig. 1. A description of the readout chain in the context of readout errors is found in Appendix C.

B. Experimental protocol

To reconstruct the POVM that describes the measurement process, we perform detector tomography. We prepare calibration states on the qubits using single-qubit gates. The calibration states are explicitly given in Sec. II B. For the phase offset between the state-preparation and readout control pulses, we use virtual Z gates. The calibration states are measured in a set of different bases corresponding to the projectors onto the eigenbasis of the three Pauli operators X, Y, and Z. The POVM is then reconstructed in postprocessing according to Ref. [23].

Once the detector tomography has been completed, we prepare different quantum states to which we can apply quantum readout-error mitigation. The prepared states are randomly selected from the set of product states. Each single-qubit state is generated by uniformly sampling from the surface of the corresponding Bloch sphere, referred to as Haar random states. Haar random states are created by

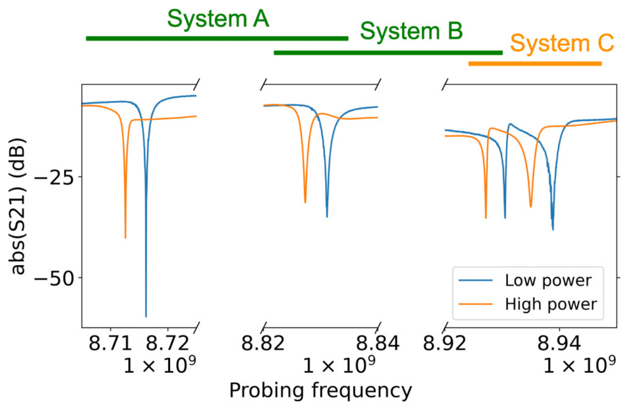


FIG. 1. Resonator spectroscopy of the four-qubit system. At sufficiently low readout powers, each qubit couples coherently to its resonator, enabling dispersive readout. Systems A, B, and C are two-qubit systems of neighboring resonator-qubit systems. An important feature is the closeness of resonators in System C, symbolized by the orange color. From left to right, the corresponding qubits are q0, q1, q2, and q3.

applying a random unitary matrix U, drawn at random from the unitary group $U(2^{n_{\text{qubits}}})$, to the ground state $|0\rangle$ [37].

C. Inducing readout noise

To study the effect of readout noise on detector tomography and its expectation values, we induce noise that does not affect state preparation and takes effect only after the quantum state of interest has been prepared. The dominant noise sources we investigate stem from suboptimal (i.e., weaker or stronger than ideal) readout power of the resonators and insufficient amplification in the readout chain.

The amplitude of the readout pulses can be adjusted individually or simultaneously for multiple qubits. Generally, if the readout amplitude is too low, the measurement may not be fully projective, and reliable state distinguishability is not possible. In the case of very high readout amplitude, we excite states outside the qubit manifold, biasing the data [13].

We sweep the frequency and amplitude of the continuous pump tone of the parametric amplifier. This results in varying amplification and therefore modifies the overlap of basis states in the IQ (in phase and quadrature signal) plane.

IV. READOUT-ERROR-MITIGATED MULTIQUBIT STATE TOMOGRAPHY

We apply multiqubit readout-error mitigation to the system described in the previous section. We use a generalization of the experimental protocol depicted in Fig. 1 of Ref. [23] to multiqubit systems. We study the rate of information extraction from readout-error-mitigated quantum-state tomography experiments. The rate of information extraction serves a proxy for the total readout noise in the system, which includes both correlated and coherent errors. We use this quantity as a quality metric for the readout, which we compare to assignment fidelity. Assignment fidelity is the standard readout quality metric, but it captures only readout errors in the computational basis. In particular, we measure the number of experimental shots required to reach a given reconstruction infidelity threshold in Sec. IV B. We then discuss the optimal distribution of a fixed shot budget between the different stages of the scheme in Sec. IV C. Finally, we apply the protocol to full three-qubit state tomography to assess the scalability of the approach (Sec. IV B 1).

A. Extracting infidelity threshold crossings from Bayesian-mean estimated state tomography

We quantify the rate of information extraction by how many shots are required to reach a certain infidelity threshold in a state reconstruction experiment. To motivate the use of readout-error mitigation instead of unmitigated quantum-state tomography, we investigate the infidelity curves of both cases in our two-qubit experiment. Figure 2 shows four mean infidelity curves, two quantum readout-error mitigated (QREM, blue) and two unmitigated (std. QST, red), as functions of the number of shots used for state tomography. Each curve is averaged over 16 products of single-qubit Haar random states [38]. The crossing points of these curves with a given constant infidelity (0.20, 0.15, or 0.10) are extracted using a Bayesian-mean estimator [23]. This allows one to identify the exact number of shots at which the infidelity curve crosses a specified infidelity threshold value. We note that the unmitigated curves saturate at high infidelity values. In fact, at high levels of noise, they do not even cross these thresholds (green horizontal lines). This makes them inappropriate for precisely benchmarking readout quality. If one uses the readout-error-mitigated infidelity curves (depicted in blue) instead, the infidelity thresholds are crossed in the steady exponential regime (specifically, in the linear segment of the log-log plot) of the curves, making them suitable for benchmarking. Furthermore, infidelity curves are most accurately obtained when employing a Bayesian-mean estimator, highlighting an advantage over maximum-likelihood estimators in the reconstruction process. Bayesian-mean estimators can be efficiently computed for two qubits, after which, maximum-likelihood estimators gain an additional advantage in speed.

B. Readout noise affecting information-extraction rate

To investigate how the rate of quantum information extraction is influenced by readout noise, we perform readout-error-mitigated two-qubit quantum-state tomography experiments on System A (see Fig. 1) with induced

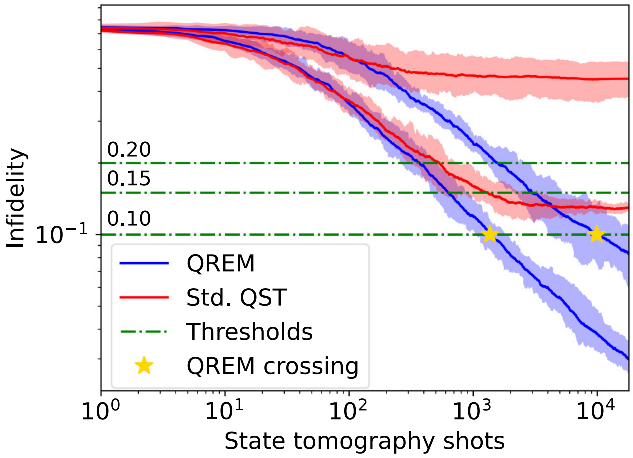


FIG. 2. Mean infidelity curves crossing different threshold values (corresponding to infidelity values of 0.10, 0.15, and 0.20). Two different readout parameters are examined, where blue curves are with and red curves are without quantum readout-error mitigation (QREM). The curves are created using a Bayesian-mean estimator of the reconstructed quantum states. Only readout-error-mitigated infidelities reliably cross the given thresholds, where the lowest threshold was only crossed by the QREM curves. Each curve is averaged over 16 Haar random initial states, and the shaded area indicates the interquartile range.

readout noise. We compare the results with a corresponding analysis using assignment fidelity as a readout quality metric. Since assignment fidelity depends exclusively on computational basis states and does not fully benchmark readout fidelity, our analysis aims to confirm the suitability of using assignment fidelity as a measurement standard.

1. Varying readout power

We vary the readout power for both resonators simultaneously, so that they are equal at all times. This leads to the possible existence of multiple optima for information extraction. The upper figure in Fig. 3 illustrates the number of shots necessary to achieve the infidelity thresholds. The plot shows one optimum for information-extraction rates at 0.48 (arb. units) readout amplitude, illustrated by a red dashed line. This is in agreement with both the product and the sum of single-qubit assignment fidelities (see the bottom panel in Fig. 3). The optimum obtained for infidelity thresholding is statistically significant, whereas the measured assignment fidelities are also consistent with a plateau between readout amplitudes of 0.35 and 0.48. This is due to the relatively large error bars, limited by the number of computational basis states measured. It is expected that repeating the experiments with more shots would also highlight one clear optimum at 0.48 (arb. units) readout power. The general features of the readout parameter landscape are explained by low state distinguishability at low readout powers and state leakage at higher readout powers, as also observed, for example, in Ref. [13]. Both the assignment fidelity and infidelity convergence landscapes are similar and well explained by the above reasoning. The fraction of leaked states at higher readout power is plotted in Appendix D. In summary, the experiment demonstrates that the more comprehensive infidelity thresholding aligns with assignment fidelities at the optimal readout power for two-qubit multiplexed readout. Given the higher cost associated with infidelity thresholding without a corresponding increase in accuracy, assignment fidelity is preferred in this setup.

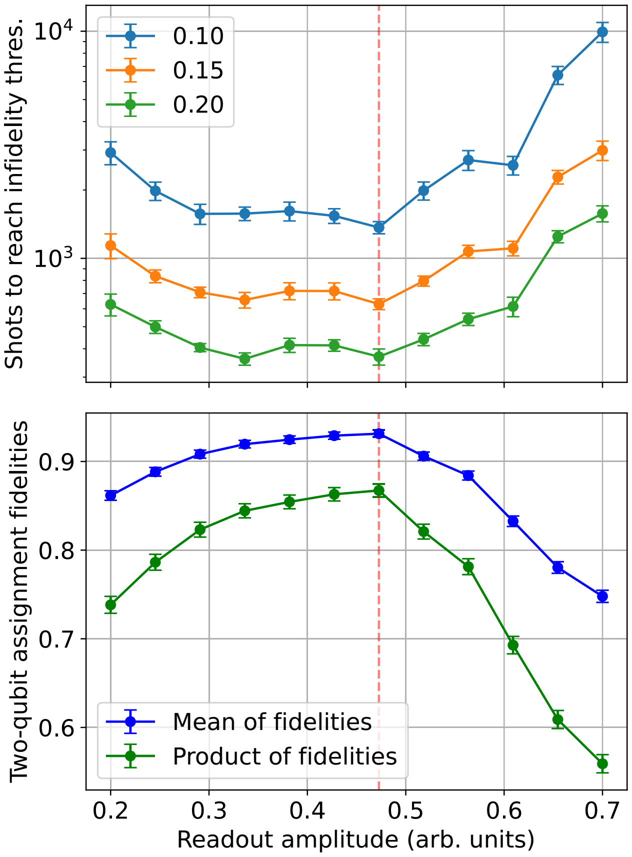


FIG. 3. Comparison of quantum-state-tomography-based readout benchmark with assignment fidelities for varying readout amplitudes. The vertical dashed red line indicates optimal readout amplitude. Top: number of two-qubit state tomography shots required to reach infidelity thresholds of 0.10, 0.15, and 0.20 for various readout amplitudes. The error bars are bootstrapped standard deviations. For very low readout powers, the distinguishability of the basis states is significantly reduced. For very high readout amplitudes, the measurement process excites higher states in the system, which the state classifier cannot correctly take into account. Each point uses 558 000 shots. Bottom: mean and product of single-shot assignment fidelities. Each point uses 2000 shots.

2. Parametric amplification

The resonators of System A are detuned by about 115 MHz from each other; therefore, it is expected that they could have different optimal operating parameters on the shared TWPA. This is due to the nontrivial frequency dependence of the gain, as measured in Ref. [27]. Thus, it is not sufficient to test the TWPA on a single qubit. Performing state tomography on this system (see Fig. 4), we see that information extraction is fastest at around 6.0 dBm pump power. This is in good accordance with the maximum of both the sum and product of single-shot assignment fidelities. Once the TWPA is oversaturated in pump power, a significant slowing of information extraction by about a factor of 6 compared to the optimum is observed (for a more detailed measurement of this saturation, see Appendix E). At low pump powers, the information-extraction rate is slowed because of insufficient distinguishability of the computational basis states at readout, which can also be observed in the corresponding assignment fidelities.

Similar to varying the readout power, both infidelity thresholding and assignment fidelity agree on the optimal pump power.

C. Optimal shot budget distribution

Experiments often have to be limited in run time. This can be due to experimental drift or simply resource sharing. This raises the question of how one should optimally allocate a fixed measurement budget between detector tomography and state tomography to minimize the resulting infidelity in the reconstruction of a quantum state. We define the shot ratio as

$$r = \frac{\text{QDT shots}}{\text{QDT shots} + \text{QST shots}}, \tag{7}$$

i.e., the ratio of the shots used for detector tomography to the total number of shots performed in the reconstruction of a single state.

To find the optimal ratio for our experiment, we performed single- and two-qubit tomography with a fixed number of total shots for different values of r. For a representative convergence of the state reconstruction, an average of the same 40 Haar random states was used for each value of r for single-qubit reconstruction. For two-qubit shot budgeting, ten products of single-qubit Haar random states were used.

The results of the experiments can be seen in Fig. 5. For both single- and two-qubit readout-error-mitigated state tomography, spending half the shot budget on calibration (QDT) is a reasonable choice. In the case of very few shots being spent on calibration, the resulting infidelity is limited by sample fluctuations in the estimated measurement operator. On the upper end, if too few shots remain for state tomography, the experiment will not reach the lowest infidelities. The ideal shot ratio ranges are highlighted in green in Fig. 5. From these plots, we can estimate the optimal value of r, as defined in Eq. (7), to be

$$r_1 = 0.4 \pm 0.2$$

for a single qubit and

$$r_2 = 0.6 \pm 0.3$$

for two-qubit experiments. In Appendix F, we consider a simple theoretical model based on the number of independent parameters to be estimated. Our predicted optima are at larger values than the observed ones, indicating that

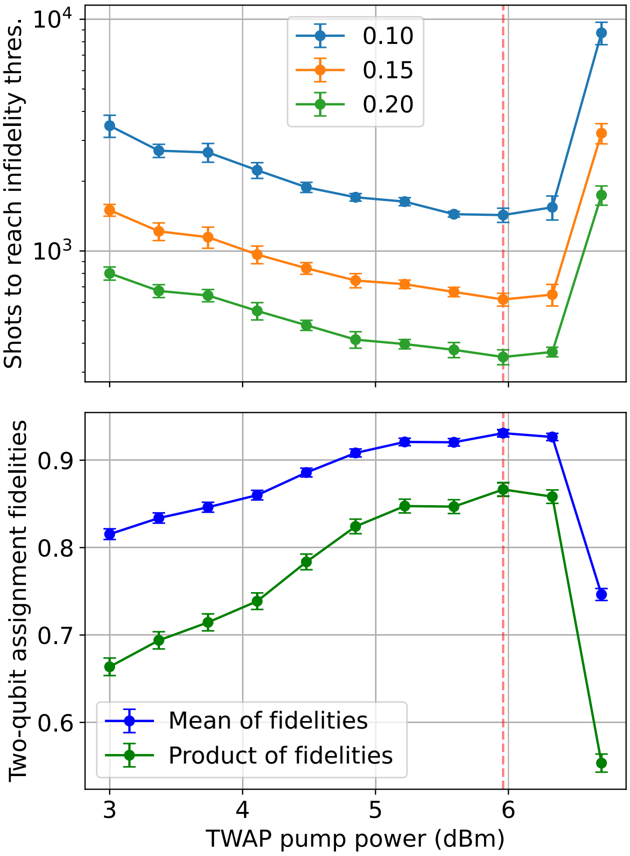


FIG. 4. Comparison of quantum-state-tomography-based readout benchmark with assignment fidelities for varying amplification. The vertical dashed red line indicates optimal pump power. Top: number of shots in state tomography needed to reach 0.10, 0.15, and 0.20 reconstruction infidelity. The error bars show bootstrapped standard deviations. Each point uses 558 000 shots. Bottom: sum and product of assignment fidelities for the performed experiments. State separation increases with pump power up to approximately 6.5 dBm, after which it drops sharply. The error bars show propagated standard deviations from the assignment fidelity. Each point uses 2000 shots.

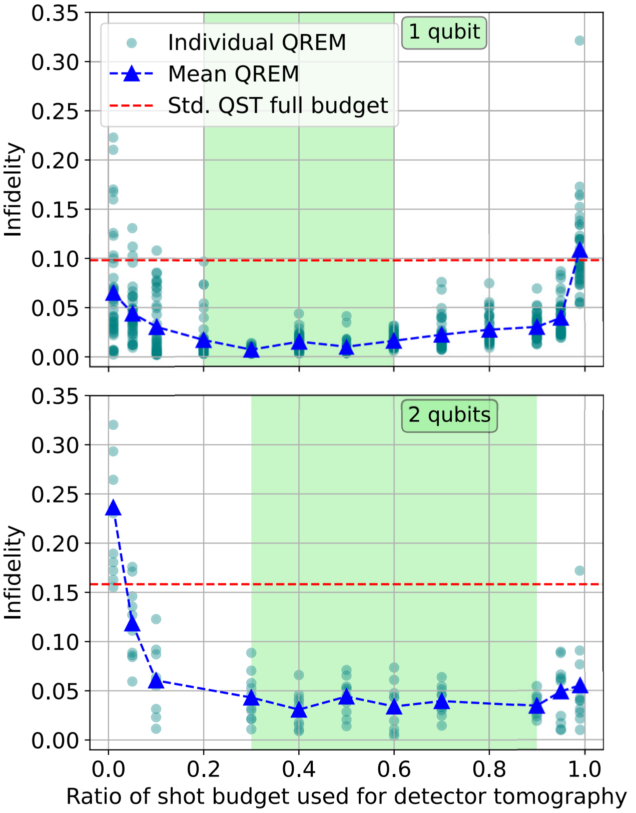


FIG. 5. Reachable infidelities for different distributions of a given shot budget in readout-error-mitigated state reconstruction. The plotted infidelities refer to the final infidelity at the end of quantum-state reconstruction. The blue triangles correspond to the mean infidelities, while the spread in each case is shown with teal circles. The red horizontal line indicates the mean final infidelity at which the full budget was used for standard state tomography. The identified optimal shot ratio is marked in green. Top: single qubit, shot budget of 10 000, averaged over 40 Haar random states. Bottom: two-qubit reconstruction with a shot budget of 40 000, averaged over ten Haar random states.

there is a significant amount of structure, which likely simplifies the statistical estimation task.

D. Three-qubit state tomography

In light of the renewed interest in three-qubit gates, noted for their potentially higher fidelity, shorter duration, and ability to replace multiple two-qubit gates [28,29,39], we perform three-qubit quantum-state tomography with and without readout-error mitigation. This examination aims to evaluate the scalability of the protocol introduced in Ref. [23] to more qubits. The experiment was repeated on the same prepared state with an increasing number of shots per basis measurement in both detector tomography and state tomography. The results are plotted in Fig. 6. Infidelities below 0.01 were reached, with an outlier data point at 12 000 shots per basis measurement likely attributable to measurement drift, which we investigate further in Appendix G. A comparison of the theoretically expected density matrix to the reconstructed one with and without readout-error mitigation can be found in Appendix H Fig. 12. We observe an infidelity improvement of around a factor of 30 when comparing state tomography with and without readout-error mitigation. Therefore, we conclude that the method introduced in Ref. [23] is scalable to three superconducting qubits and is capable of significantly reducing readout errors. This provides a way to benchmark and understand three-qubit gates, even in the presence of coherent readout errors in the system. Complete state reconstruction is intrinsically not a scalable method; however, with the established applicability to three-qubit benchmarking, the technique can be applied to characterize few-qubit subsystems within larger qubit systems, which is typically sufficient for mitigation of correlated errors [12,21,25].

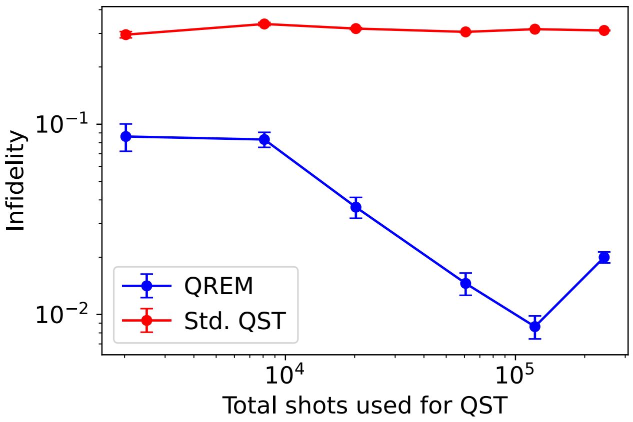


FIG. 6. Reconstruction of a three-qubit density matrix with and without quantum readout-error mitigation (QREM) as a function of the total number of shots used for QST. The error bars indicate the standard deviation of the estimator obtained from 1000 bootstrap resamplings.

V. TWO-QUBIT CORRELATIONS FROM DETECTOR TOMOGRAPHY

As multiplexed readout has become widespread in superconducting qubit devices of growing size, it is important to reliably quantify potential readout correlations that exist due to limited resonance frequency spacing. In this section, we use correlation coefficients extracted from detector tomography to investigate our qubit system. In addition, we analyze the dependence of these correlation coefficients on the amplitude of the readout pulse.

A. Correlation measurements for two-qubit subsystems

To investigate the effect of resonator frequency spacing, we performed detector tomography on Systems A

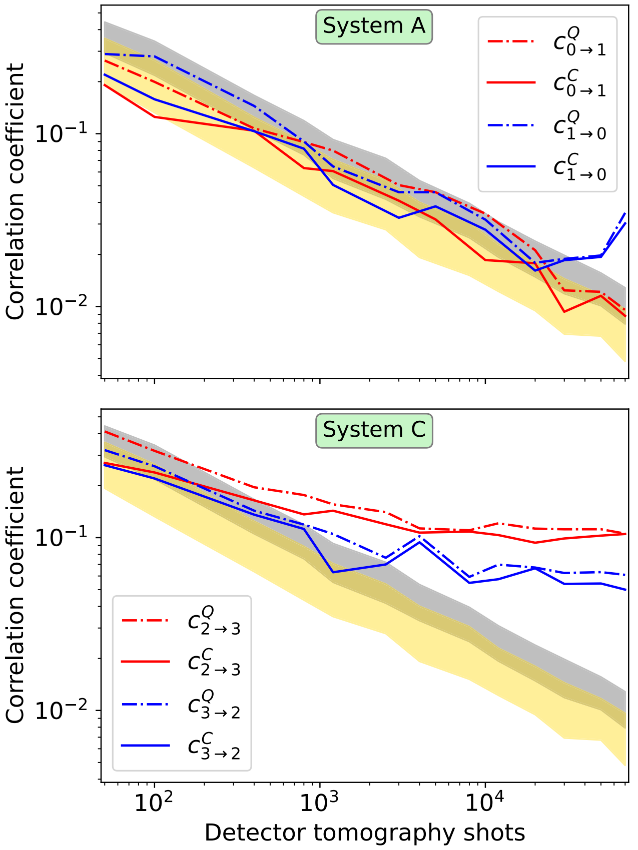


FIG. 7. Correlation coefficients plotted as a function the total number of detector tomography shots used for reconstructing the measurement operators. The silver and gold shaded areas represent the standard deviations of quantum and classical correlation coefficients, respectively, derived from numerical simulations of measurements without correlation. Curves that fall within these regions indicate compatibility with uncorrelated readout. Top: in System A, the readout correlations do not exhibit any readout correlations within the bounds of the experimental constraints. Bottom: System C displays nonzero asymmetric correlations, which suggests that readouts are indeed correlated within the system; however, no coherent correlations were observed within the limitations of the experiment.

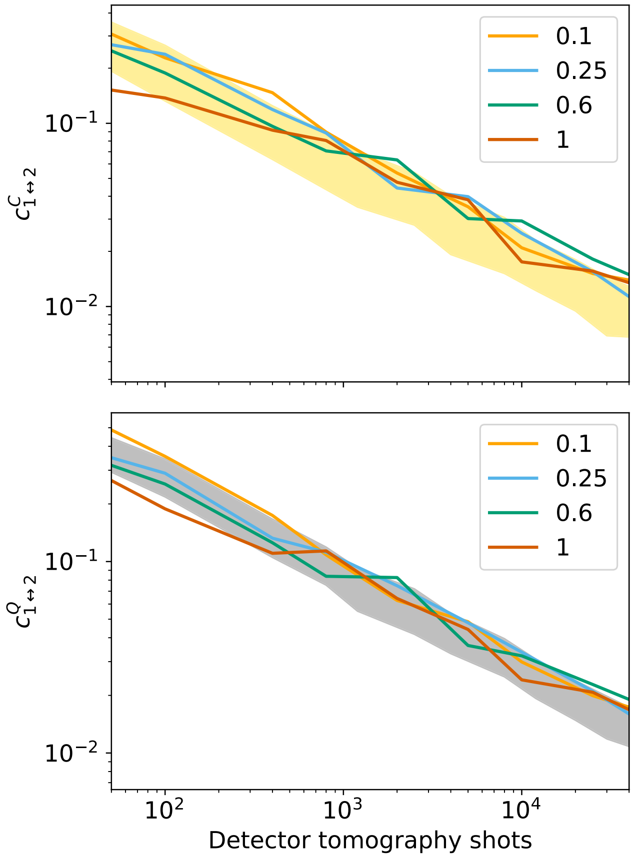


FIG. 8. Correlation coefficients in System A as a function of the number of detector tomography shots for lowered relative readout amplitudes. The silver and gold shaded areas represent the standard deviations of quantum and classical correlation coefficients, respectively, derived from numerical simulations of measurements without correlation. Top: classical correlations decrease independently of the relative readout amplitude, remaining consistent with uncorrelated readout. Bottom: quantum correlations have larger values but are always consistent with uncorrelated readout.

and C (see Sec. III A) and extracted both quantum and classical correlation coefficients as functions of detector tomography shots. The results are plotted in Fig. 7. At low numbers of shots, the correlation coefficients are compatible with uncorrelated POVMs (shaded silver and gold areas), indicating that the coefficients are dominated by sample fluctuations. System A is always compatible with uncorrelated readout, while System C shows strong asymmetric readout correlations. At high numbers of shots for System A, the $0 \rightarrow 1$ correlation coefficients move outside the standard deviation of uncorrelated POVMs, but due to low statistics, we consider this to be consistent with uncorrelated POVMs. These observation are consistent with the observed frequency spacing of resonators in the two systems: $\delta_B = 100$ MHz vs $\delta_C = 8$ MHz. Neither system exhibits significant correlated coherent readout errors. In System C, the correlation coefficients are of the order of 0.1, comparable to those reported for the Rigetti Aspen-M-3 in Ref. [12]. Reference [21] presents numerical simulations of noisy readout using both experimental and synthetic systems. Fitting the probabilistic *i*SWAP channel $\mathcal{E}(M_i) = (1-k)M_i + kU_{i\mathrm{SWAP}}M_iU^\dagger_{i\mathrm{SWAP}}$ to the observed correlation coefficients in System C is consistent with a noise strength of $k \approx 0.1$.

We therefore conclude that the correlation coefficients can be used as a precise tool to test whether resonators are sufficiently far from each other both spatially on the chip and in the frequency domain. As correlations are dependent on *in situ* readout variables such as readout power, it is possible to optimize these parameters to maximize the rate of information extraction while simultaneously minimizing the correlations between different qubit readouts. Our results therefore open the way for a more comprehensive optimization of multiqubit readout.

B. Two-qubit correlation dependence on readout noise strength

To better understand how the correlation coefficients behave under additional noise sources, we investigated how the coefficients change when lowering the distinguishability of computational basis states using readout pulses with varying amplitudes. Lowered state distinguishability can be modeled as depolarizing-type noise on each qubit, which is not expected to impact the correlation coefficients; see Appendix J for more details.

We define the relative readout amplitude as the ratio between the used amplitude of the readout tone and its ideal amplitude. To avoid leakage to higher levels, we restrict ourselves to decreased readout amplitudes, corresponding to relative readout amplitudes not larger than 1. *Ad extremum*, the single-shot assignment fidelity measured with a relative readout amplitude of 0.1 is only about 55% compared to around 95% at a readout ratio of 1.

The results of the experiments can be seen in Fig. 8. Within the bounds of the experiment, both quantum and classical correlations are independent of the strength of the readout noise, even at very high levels of noise, confirming our theoretical expectations.

VI. CONCLUSIONS AND OUTLOOK

We investigated a generalized method for readout-error mitigation based on detector tomography and explored its utility as a benchmarking tool for superconducting qubits. The protocol was applied to various two- and three-qubit state reconstruction experiments with superconducting transmon qubits. We introduced a benchmark for qubit readout quality that goes beyond conventional techniques focused on the fidelity of computational basis state assignments, which primarily detect classical errors. Our approach uses quantum infidelities between prepared states and readout-error-mitigated reconstructed states as a proxy for the information-extraction rate. Using reconstruction infidelity characterizes both coherent and correlated readout errors, making it a more comprehensive benchmark than assignment fidelity. We thus propose using the rate of infidelity threshold crossings as a validation tool for assignment fidelity and as a suitable comprehensive figure of merit for certifying single- and multiqubit readout processes. In particular, when optimizing the readout chain, assignment fidelity fails to properly capture correlated readout errors, while infidelity-thresholding-based benchmarking is able to quantify the correlations present in the measurement process. Therefore, if assignment fidelity falls short, this new figure of merit can be used to find optimal operating parameters for readout chains, such as fine-tuning amplifiers for ideal readout performance. We studied the impact of different readout parameters on the rate of information extraction. In particular, the dependence on the pump power of the parametric amplifier and the amplitude of the readout resonator drive tone were investigated. The assignment fidelity- and infidelity-convergence-based methods produced very similar results in the investigated parameter range, validating the use of assignment fidelity in the investigated system.

We investigated how the budgeting of shots between detector tomography and state tomography influenced the state reconstruction infidelity and thus identified the optimal distribution of shots. To address the increasing interest in two- and three-qubit gates for superconducting qubits, we demonstrated that beyond-classical readout-error mitigation is suitable for these system sizes.

By performing readout correlation measurements, we showed that readout correlations are significant in the weakly detuned system. We observed no correlated coherent readout errors in this system. By adjusting the readout amplitude of the resonators, we showed that even high noise levels do not induce readout correlations, validating correlation coefficients as an additional useful benchmark for the multiplexed readout of superconducting qubits. Our comprehensive benchmark for quantifying qubit readout precision can be applied to test innovative qubit architectures [40] and readout methods [41]. The optimization of readout parameters should not only focus on improving the performance of individual qubits, but also on keeping readout correlations in multiplexed readout to a minimum. It would be interesting to perform optimization experiments on weakly detuned resonator systems to test this limit of qubit readout optimization. Importantly, we encourage the convergence of reconstruction infidelity to be used in upcoming superconducting qubit experiments that aim to push the boundaries of readout accuracy. Furthermore, it would be interesting to compare the infidelity-thresholding approach to a similar approach that uses QREM that only mitigates correlated classical errors, as these approaches require fewer shots for calibration and could be a more economical solution if coherent errors can be neglected.

ACKNOWLEDGMENTS

The authors thank Andy Ding for providing the qubit chip used in the study. We also thank W. Oliver and G. Calusine for providing the parametric amplifier. This work was partially financed by the Baden-Württemberg Stiftung gGmbH.

A.D.G. performed the experiments with supervision from H.R., J.L., and A.U. The theory and numerics were implemented by A.S.A. with supervision from M.G. A.D.G. and A.S.A. performed the data analysis. The first manuscript was prepared by A.D.G. and A.S.A. All authors contributed to the finalization of the manuscript.

The authors declare no competing interests.

DATA AVAILABILITY

The data that support the findings of this article are openly available [42]; embargo periods may apply.

APPENDIX A: DETAILS OF CORRELATION COEFFICIENTS

References [12,18] introduced a metric for quantifying how the state of one qubit impacts the state of another. This is based on efficiently measuring and reconstructing all possible two-qubit POVMs with detector overlapping tomography [18,43]. A readout correlation measure can be obtained as follows. For each two-qubit POVM, trace out half of the POVM conditioned on a state being measured in the traced-out subsystem. Optimize over all possible states present on the traced-out subsystems and select the largest distance between two traced-out POVMs. These correlation coefficients can be written explicitly as

$$c_{j\to i} = \sup_{\rho_j,\sigma_j} ||(M_0^{i,\rho_j} - M_0^{i,\sigma_j})||_2, \tag{A1}$$

where $M_{x_i}^{i,\sigma_j} = \sum_{x_j} \mathrm{Tr}_j\,(M_{x_i x_j}^{ij}\,(\mathbb{1}_i \otimes \sigma_j))$, x_i and x_j label the outcomes on the two-qubit system, and $||A||_2$ refers to the spectral norm (i.e., the largest singular value of A). This corresponds to the correlation coefficients employed in Ref. [12] using the “worst-case” distance, with the difference that in that work, the norm applied is the infinity operator norm, $||A||_\infty = \max_{1\le k\le d}\sum_{l=1}^{d}|a_{kl}|$, where $d = \dim(A)$. The coefficients quantify the impact of the state of qubit j on the effective POVM measured on qubit i. Due to the asymmetry of these coefficients, it is useful to define the symmetric correlation coefficient $c_{i\leftrightarrow j} = (c_{j\to i} + c_{i\to j})/2$.

If the optimization in Eq. (A1) is restricted to only the computational basis states, i.e., products of eigenstates of the σ_z operator $\{|0\rangle\,,|1\rangle\}^{\otimes N}$, the coefficients are called classical correlation coefficients c^C. If the optimization is performed over all pure states [12], the resulting coefficients are called quantum correlation coefficients c^Q. The quantum correlation coefficients cannot be smaller than the classical ones and are only equal if there are no coherent errors present in the readout process (up to statistical fluctuations).

TABLE I. Qubit names and resonator and qubit frequencies of the quantum device. The qubits are named in order from left to right on the physical chip, starting from q0. Qubit and resonator frequencies typically varied up to a few MHz between cooldowns.

Qubit	ω_{res} (GHz)	ω_{01} (GHz)
q0	8.716	4.455
q1	8.831	5.409
q2	8.931	4.127
q3	8.939	4.926

APPENDIX B: QUBIT AND RESONATOR PARAMETERS

Qubit and resonator parameters varied slightly between different cooldowns. Some representative values are listed in Tables I and II.

APPENDIX C: READOUT CHAIN

To understand the origin of readout errors, we provide a brief review of the readout chain used for typical dispersive-shift measurements of superconducting qubits. After reaching the transmission line on the chip, the readout signals couple into the resonators that match their frequency. The frequency of the resonator depends on the collapsed qubit state through the dispersive shift [44]. This signal is amplified first on the mixing chamber stage with a Josephson traveling-wave parametric amplifier (JTWPA), driven by a continuous microwave pump tone, which is also channeled into the remaining chain. Some signal amplitude is lost before and after the JTWPA due to nonsuperconducting cabling and connectors. The resulting signal enters two high-electron-mobility transistor (HEMT) amplifiers, which add noise photons to the signal. External mixers leak sidebands into the signal before it reaches the digitizer and integration module, finally yielding a measured value of 0 or 1.

TABLE II. Typical coherence times and single-shot readout fidelities of the individual qubits. The assignment-fidelity values show the best possible single-qubit readout assignment fidelity. Multiqubit assignment fidelities were observed to be lower in some experiments.

Qubit	Assignment fidelity (%)	T_1 (μs)	T_2 (μs)
q0	96	118	67
q1	93	22	30
q2	92	27	56
q3	80	23	ca. 15

APPENDIX D: LEAKAGE AND ASSIGNMENT FIDELITY

In this section, we discuss the comparison of infidelity thresholding and assignment fidelity for varying readout powers in more detail (see Sec. IV B 1).

We treat the transmons as three-level quantum systems (qutrits) and distinguish three regions in the IQ plane, corresponding to the ground and excited states and the third (leaked) state. We prepare an equal number of $|0\rangle$ and $|1\rangle$ states and apply readout pulses with growing amplitudes to each qubit. The resulting distributions in the IQ plane are grouped into three clusters (ground, excited, and leaked), and a linear sum assignment is performed to relabel the clusters in maximum accordance with the known (but sometimes incorrect) labels 0 or 1. Note that the lowest readout power at which we can reliably perform the leakage estimate in this way is 1.5× the optimal readout power. This is because at lower readout powers, it is more difficult to reliably estimate the leakage ratio, as the dataset becomes very unbalanced, leading to misinterpretation (overfitting). This is a known problem for N-clustering algorithms. The results are shown in Fig. 9. Although q1 is more prone to leakage, its assignment fidelity is less sensitive to this effect at high readout powers. This effect can be attributed to the position of states in the IQ (in phase-quadrature) plane.

APPENDIX E: TWPA POWER SATURATION

Pumping the parametric amplifier with increasing amplitude saturates the amplification, and the signal quality drops off sharply at around 7.3 dBm pump power (see Fig. 10). This measurement was conducted with a different set of microwave cables; therefore, there is a constant offset in power compared to Fig. 4.

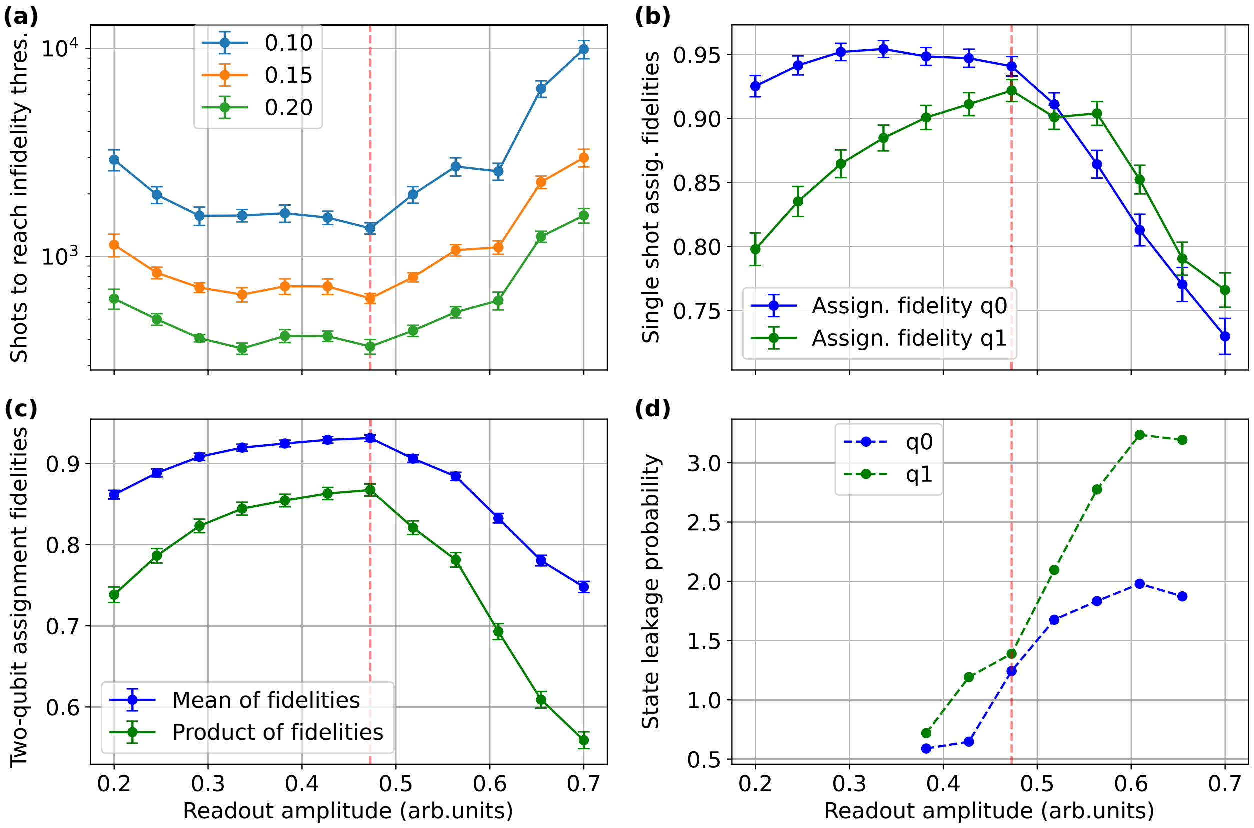


FIG. 9. Comparison of readout-error-mitigated quantum-state-tomography-based readout benchmark with assignment fidelities. The vertical dashed red line indicates the optimal readout amplitude identified in Fig. 3. (a) Number of two-qubit state tomography shots required to reach infidelity thresholds of 0.10, 0.15, and 0.20 for various readout amplitudes. The error bars are bootstrapped standard deviations. For very low readout powers, the distinguishability of the basis states is significantly lowered. For very high readout amplitudes, the measurement process excites higher states in the system, which the state classifier cannot correctly take into account. (b) Single-shot assignment fidelities during the experiments; each point uses 2000 shots. The error bars show one standard deviation. (c) Mean and product of single-shot assignment fidelities from Fig. 3. (d) Probability of measuring a leaked state as a function of readout amplitude.

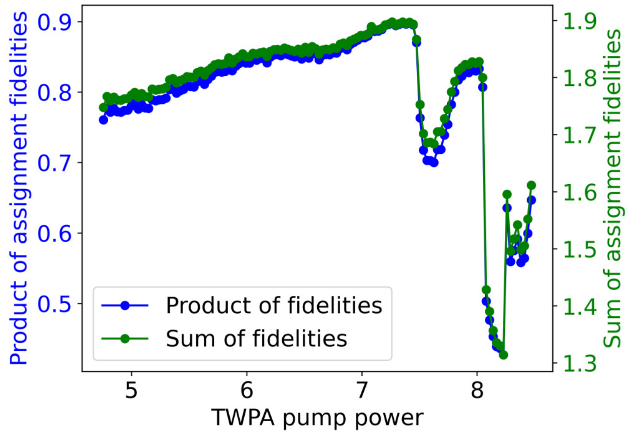


FIG. 10. Sum and product of single-qubit assignment fidelities in System A. A sharp drop in both variables occurs at approximately 7.3 dBm pump power.

APPENDIX F: THEORETICAL SHOT BUDGETING

In a standard setup of detector tomography followed by state tomography, the distribution of shots should be dictated by the number of parameters to estimate in each part. If there is no knowledge of the quantum states to be estimated, then there are $4^n - 1$ parameters for each normalized n-qubit state. An m-outcome POVM has $(m-1)4^n$ parameters that need to be estimated. If the POVM is informationally complete (IC), $m = 4^n$. Inserting this into Eq. (7),

$$r_n = \frac{(m-1)4^n}{(m-1)4^n + N_s(4^n - 1)} \overset{\text{IC}}{=} \frac{4^n}{4^n + N_s}, \tag{F1}$$

where N_s is the number of states to be estimated with the same reconstructed POVM. Compared with the experiment in Sec. IV C, setting $N_s = 1$, we get $r_1 = 0.80$ and $r_2 = 0.94$, significantly higher.

The discrepancy between these two results may arise from the fact that not all the parameters are free and independent in a physical setting. For example, if the reconstructed states are close to pure, there are effectively $2(2^n - 1)$ parameters in the state (since most parameters are close to zero). In addition, the POVM implemented in the experiment is not an IC POVM but a set of projective measurements on the eigenstates of the Pauli operators, each individually normalized.

If the errors at readout are close to classical (i.e., only redistributing the eigenvalues of the POVM elements), the number of parameters in the POVM is $(4^n - 1)2^n$. If we assume close-to-classical errors, the ratio is

$$\begin{aligned} r_n &\overset{\text{mixed}}{=} \frac{2^n}{2^n + N_s}, \\ &\overset{\text{pure}}{=} \frac{4^n + 2^n}{4^n + 2^n + 2N_s} \end{aligned} \tag{F2}$$

which, compared to the experiment ($N_s = 1$), gives us $r_1 = (0.67, 0.75)$ and $r_2 = (0.80, 0.91)$ for (mixed, pure) target states.

APPENDIX G: MEASUREMENT DRIFT

As both QDT and QST scale exponentially in run time with the number of qubits, the experiment has to run self-consistently over a prolonged period. One limitation placed on the length of experiments is the temporal drift of the operational parameters. To quantitatively estimate the timescales and effects of these drifts on our POVMs, we repeatedly perform two-qubit detector tomography over 11 h while tracking the temperature and humidity of the laboratory. Specifically, we reconstruct the measurement operators corresponding to ZZ- and XX-basis-multiplexed readout of System A and plot their coherent errors in Fig. 11. We define coherent errors in Appendix I. While the coherent errors in the ZZ basis are small and can be explained by statistical errors, the XX basis contains significant coherent errors. This can be explained by the additional rotation gates used, which are not present in ZZ-basis readout. Under stable laboratory conditions, errors are observed to be temporally constant for both. Introducing rapid temperature and humidity changes by turning the laboratory air conditioning off, the amount of coherence in the reconstructed measurement operator of the XX basis drifts significantly. As the ZZ-basis coherences are stable, we identify the likely source of coherence drift as originating from the manipulation of the qubit used for XX-basis measurements.

APPENDIX H: THREE-QUBIT QUANTUM-STATE TOMOGRAPHY WITH READOUT-ERROR MITIGATION

The three-qubit density matrix obtained with readout-error mitigation has around an order-of-magnitude smaller error in the worst entry of the unmitigated density matrix (see Fig. 12). The nonvanishing matrix entries have the largest errors both with and without readout-error mitigation. The infidelity of reconstruction was calculated from these density matrices.

APPENDIX I: COHERENT ERRORS

One way to quantify the degree of nonclassicality of readout errors is to measure the distance between the reconstructed POVM and the nearest POVM subject to

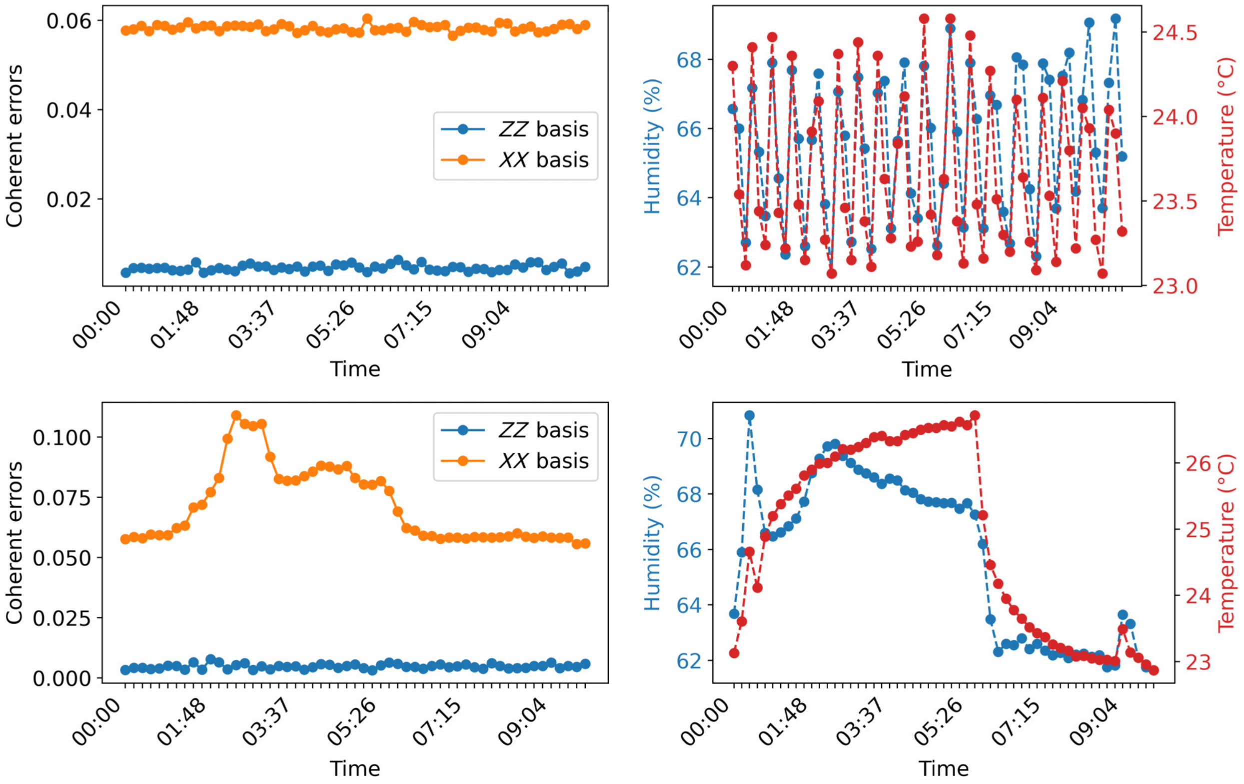


FIG. 11. Stability of coherent readout errors with varied laboratory temperature and humidity. Top left: coherent errors [defined in Eq. (I4)] in the reconstructed measurement operators for *ZZ*- and *XX*-basis readout under stable laboratory conditions. Top right: temperature and humidity values during the corresponding measurement. The rapid oscillations come from the feedback mechanism of the air conditioning. Bottom left: coherent errors in the reconstructed measurement operators for *ZZ*- and *XX*-basis readout under unstable laboratory conditions induced by turning the air conditioning off, then on. Bottom right: temperature and humidity values during the corresponding measurements.

only classical errors. Classical errors are captured by a confusion matrix [17], which involves a statistical redistribution of the probabilities of measurement outcomes in the computational basis. Specifically, it can be described by a matrix C, which gives the relation between an ideal measurement probability p_i of outcome i and the measured frequencies $\tilde{p}_i = n_i/N$,

$$\tilde{p}_i = C_{ij} p_j . \tag{I1}$$

Equation (I1) can be rephrased in terms of a noisy POVM $\tilde{\boldsymbol{M}}$ and an ideal projective measurement $\boldsymbol{P}$

$$\tilde{M}_i = C_{ij} P_i . \tag{I2}$$

If coherent errors are present, then Eq. (I2) gets an additional term Δ_i,

$$\tilde{M}_i = C_{ij} P_i + \Delta_i . \tag{I3}$$

Since every projective measurement operator P_i contains only diagonal entries, we can quantify coherent errors by computing the distance between the reconstructed POVM $\tilde{\boldsymbol{M}}$ and the same POVM with only diagonal entries [17],

$$\epsilon_{\text{coh}} = \text{dist}(\tilde{\boldsymbol{M}}, \mathcal{E}_{\text{dep}}(\tilde{\boldsymbol{M}})), \tag{I4}$$

where $\text{dist}(\cdot,\cdot)$ is a suitable POVM distance and $\mathcal{E}_{\text{dep}}(\tilde{M}_i) = \text{diag}(\tilde{M}_i)$ is the maximally dephasing channel. For coherent errors, we use the average-case distance for POVMs, which is defined as [45]

$$d_{\text{AV}}(\boldsymbol{M}, \boldsymbol{N}) = \frac{1}{2d} \sum_{i=1}^{n} \sqrt{\|M_i - N_i\|_{\text{HS}}^2 + \text{Tr}(M_i - N_i)^2}, \tag{I5}$$

where n is the number of outcomes for the POVMs $\boldsymbol{M}$ and $\boldsymbol{N}$, d is the dimension of the measurement operators, and $\|A\|_{\text{HS}} = \sqrt{\text{Tr}(A^2)}$.

APPENDIX J: EFFECT OF DEPOLARIZING NOISE ON CORRELATION COEFFICIENTS

Many common sources of experimental errors, such as reduced computational-state distinguishability, can be

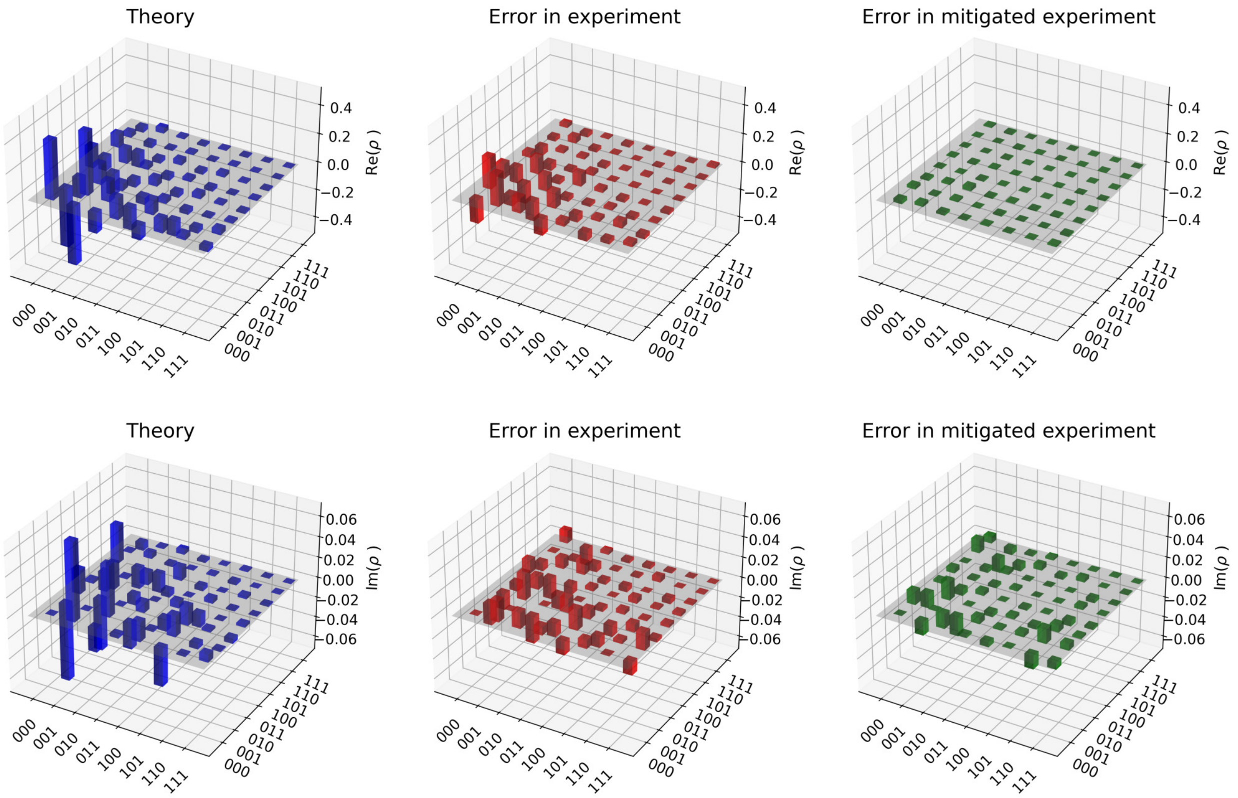


FIG. 12. Comparison of theoretical density matrix of a three-qubit state and its reconstruction with and without readout-error mitigation. The top row shows the real part of the complex eight-dimensional matrices, and the bottom row shows the imaginary part. The quantum states reconstructed with and without readout-error mitigation have infidelities of 0.030 and 0.447, respectively. The different scales for the real and imaginary parts are due to the chosen quantum state. The errors in the imaginary and real parts of the error-mitigated experiments are of the same order of magnitude.

modeled as a depolarizing channel on individual qubits. The single-qubit depolarizing channel can be written as [46]

$$\mathcal{E}_{\text{depol}}(M_i) = \left(1 - \frac{3k}{4}\right) M_i + \frac{k}{4}\left(XM_iX + YM_iY + ZM_iZ\right), \tag{J1}$$

where k denotes the strength of depolarization and X, Y, and Z are the Pauli matrices.

In the experiment discussed in Sec. V B, the distinguishability of the two-qubit computational basis affects each qubit independently. We therefore expect the noise to be separable:

$$\mathcal{E}^2_{\text{depol}}(M_i) = (\mathcal{E}_{\text{depol}} \otimes \mathcal{E}_{\text{depol}})(M_i). \tag{J2}$$

Applying the separable depolarizing channel to the two-qubit measurement operators in the definition of the correlation coefficients, we find that

$$\begin{aligned} M^{i,\sigma_j}_{x_i,\text{depol}} &= \sum_{x_j} \text{Tr}(\mathcal{E}^2_{\text{depol}}(M^{ij}_{x_i x_j})(1 \otimes \sigma_j)) \\ &= \sum_{x_j} \text{Tr}(M^{ij}_{x_i x_j}(\mathbb{1} \otimes \mathcal{E}_{\text{depol}}(\sigma_j))) \\ &= M^{i,\mathcal{E}_{\text{depol}}(\sigma_j)}_{x_i}. \end{aligned} \tag{J3}$$

The correlation coefficients in Eq. (A1) can therefore be written

$$c_{j\to i} = \sup_{\rho_j,\sigma_j} ||(M_0^{i,\mathcal{E}_{\text{depol}}(\rho_j)} - M_0^{i,\mathcal{E}_{\text{depol}}(\sigma_j)})||_2. \tag{J4}$$

The depolarizing channel acts on the optimization states ρ_j and σ_j and effectively reduces the search space for the optimization. Consequently, the correlation coefficients are nonincreasing under single-qubit depolarizing noise.

[1] R. Acharya *et al.*, Suppressing quantum errors by scaling a surface code logical qubit, Nature **614**, 676 (2023).

[2] R. Acharya, *et al.*, Quantum error correction below the surface code threshold, Nature **638**, 920 (2024).

[3] Y. Li, and S. C. Benjamin, Efficient variational quantum simulator incorporating active error minimization, Phys. Rev. X **7**, 024510 (2017).

[4] K. Temme, S. Bravyi, and J. M. Gambetta, Error mitigation for short-depth quantum circuits, Phys. Rev. Lett. **119**, 180509 (2017).

[5] Y. Quek, D. Stilck França, S. Khatri, J. J. Meyer, and J. Eisert, Exponentially tighter bounds on limitations of quantum error mitigation, Nat. Phys. **20**, 1648 (2024).

[6] Z. Li, P. Liu, P. Zhao, Z. Mi, H. Xu, X. Liang, T. Su, W. Sun, G. Xue, J.-N. Zhang, W. Liu, Y. Jin, and H. Yu, Error per single-qubit gate below 10^{-4} in a superconducting qubit, Npj Quantum Inf. **9**, 111 (2023).

[7] F. Yan, P. Krantz, Y. Sung, M. Kjaergaard, D. L. Campbell, T. P. Orlando, S. Gustavsson, and W. D. Oliver, Tunable coupling scheme for implementing high-fidelity two-qubit gates, Phys. Rev. Appl. **10**, 054062 (2018).

[8] M. Werninghaus, D. J. Egger, F. Roy, S. Machnes, F. K. Wilhelm, and S. Filipp, Leakage reduction in fast superconducting qubit gates via optimal control, Npj Quantum Inf. **7**, 14 (2021).

[9] E. Hyyppä, A. Vepsäläinen, M. Papič, C. F. Chan, S. Inel, A. Landra, W. Liu, J. Luus, F. Marxer, C. Ockeloen-Korppi, S. Orbell, B. Tarasinski, and J. Heinsoo, Reducing leakage of single-qubit gates for superconducting quantum processors using analytical control pulse envelopes, PRX Quantum **5**, 030353 (2024).

[10] A. Marton, and J. K. Asboth, Coherent errors and readout errors in the surface code, Quantum **7**, 1116 (2023).

[11] C. D. Wilen, S. Abdullah, N. A. Kurinsky, C. Stanford, L. Cardani, G. D'Imperio, C. Tomei, L. Faoro, L. B. Ioffe, C. H. Liu, A. Opremcak, B. G. Christensen, J. L. DuBois, and R. McDermott, Correlated charge noise and relaxation errors in superconducting qubits, Nature **594**, 369 (2021).

[12] J. Tuziemski, F. B. Maciejewski, J. Majsak, O. Słowik, M. Kotowski, K. Kowalczyk-Murynka, P. Podziemski, and M. Oszmaniec, Efficient reconstruction, benchmarking and validation of cross-talk models in readout noise in near-term quantum devices, arXiv:2311.10661.

[13] T. Walter, P. Kurpiers, S. Gasparinetti, P. Magnard, A. Potočnik, Y. Salathé, M. Pechal, M. Mondal, M. Oppliger, C. Eichler, and A. Wallraff, Rapid high-fidelity single-shot dispersive readout of superconducting qubits, Phys. Rev. Appl. **7**, 054020 (2017).

[14] J. Heinsoo, C. K. Andersen, A. Remm, S. Krinner, T. Walter, Y. Salathé, S. Gasparinetti, J.-C. Besse, A. Potočnik, A. Wallraff, and C. Eichler, Rapid high-fidelity multiplexed readout of superconducting qubits, Phys. Rev. Appl. **10**, 034040 (2018).

[15] J. S. Lundeen, A. Feito, H. Coldenstrodt-Ronge, K. L. Pregnell, C. Silberhorn, T. C. Ralph, J. Eisert, M. B. Plenio, and I. A. Walmsley, Tomography of quantum detectors, Nat. Phys. **5**, 27 (2008).

[16] Y. Chen, M. Farahzad, S. Yoo, and T.-C. Wei, Detector tomography on IBM quantum computers and mitigation of an imperfect measurement, Phys. Rev. A **100**, 052315 (2019).

[17] F. B. Maciejewski, Z. Zimborás, and M. Oszmaniec, Mitigation of readout noise in near-term quantum devices by classical post-processing based on detector tomography, Quantum **4**, 257 (2020).

[18] F. B. Maciejewski, F. Baccari, Z. Zimborás, and M. Oszmaniec, Modeling and mitigation of cross-talk effects in readout noise with applications to the quantum approximate optimization algorithm, Quantum **5**, 464 (2021).

[19] S. Kono, J. Pan, M. Chegnizadeh, X. Wang, A. Youssefi, M. Scigliuzzo, and T. J. Kippenberg, Mechanically induced correlated errors on superconducting qubits with relaxation times exceeding 0.4 ms, Nat. Commun. **15**, 3950 (2024).

[20] X. Li, J. Wang, Y.-Y. Jiang, G.-M. Xue, X. Cai, J. Zhou, M. Gong, Z.-F. Liu, S.-Y. Zheng, D.-K. Ma, M. Chen, W.-J. Sun, S. Yang, F. Yan, Y.-R. Jin, S. P. Zhao, X.-F. Ding, and H.-F. Yu, Cosmic-ray-induced correlated errors in superconducting qubit array, Nat. Commun. **16**, 4677 (2025).

[21] A. S. Aasen, A. Di Giovanni, H. Rotzinger, A. V. Ustinov, and M. Gärttner, Mitigation of correlated readout errors without randomized measurements, Phys. Rev. A **112**, 022601 (2025).

[22] P. C. de Groot, A. F. van Loo, J. Lisenfeld, R. N. Schouten, A. Lupaşcu, C. J. P. M. Harmans, and J. E. Mooij, Low-crosstalk bifurcation detectors for coupled flux qubits, Appl. Phys. Lett. **96**, 123508 (2010).

[23] A. S. Aasen, A. Di Giovanni, H. Rotzinger, A. V. Ustinov, and M. Gärttner, Readout error mitigated quantum state tomography tested on superconducting qubits, Commun. Phys. **7**, 301 (2024).

[24] M. R. Geller, Conditionally rigorous mitigation of multiqubit measurement errors, Phys. Rev. Lett. **127**, 090502 (2021).

[25] M. R. Geller, and M. Sun, Toward efficient correction of multiqubit measurement errors: Pair correlation method, Quantum Sci. Technol. **6**, 025009 (2021).

[26] M. Jerger, S. Poletto, P. Macha, U. Hübner, E. Il'ichev, and A. V. Ustinov, Frequency division multiplexing readout and simultaneous manipulation of an array of flux qubits, Appl. Phys. Lett. **101**, 042604 (2012).

[27] C. Macklin, K. O'Brien, D. Hover, M. E. Schwartz, V. Bolkhovsky, X. Zhang, W. D. Oliver, and I. Siddiqi, A near–quantum-limited Josephson traveling-wave parametric amplifier, Science **350**, 307 (2015).

[28] Y. Kim, A. Morvan, L. B. Nguyen, R. K. Naik, C. Jünger, L. Chen, J. M. Kreikebaum, D. I. Santiago, and I. Siddiqi, High-fidelity three-qubit *i*Toffoli gate for fixed-frequency superconducting qubits, Nat. Phys. **18**, 783 (2022).

[29] T. Itoko, M. Malekakhlagh, N. Kanazawa, and M. Takita, Three-qubit parity gate via simultaneous cross-resonance drives, Phys. Rev. Appl. **21**, 034018 (2024).

[30] X.-L. Li, Z. Tao, K. Yi, K. Luo, L. Zhang, Y. Zhou, S. Liu, T. Yan, Y. Chen, and D. Yu, Hardware-efficient and fast three-qubit gate in superconducting quantum circuits, Front. Phys. **19**, 51205 (2024).

[31] T. Thorbeck, A. Eddins, I. Lauer, D. T. McClure, and M. Carroll, Two-level-system dynamics in a superconducting

qubit due to background ionizing radiation, PRX Quantum **4**, 020356 (2023).

[32] J. Fiurášek, Maximum-likelihood estimation of quantum measurement, Phys. Rev. A **64**, 024102 (2001).

[33] A. I. Lvovsky, Iterative maximum-likelihood reconstruction in quantum homodyne tomography, J. Opt. B: Quantum Semiclassical Opt. **6**, S556 (2004).

[34] G. I. Struchalin, I. A. Pogorelov, S. S. Straupe, K. S. Kravtsov, I. V. Radchenko, and S. P. Kulik, Experimental adaptive quantum tomography of two-qubit states, Phys. Rev. A **93**, 012103 (2016).

[35] B. Lienhard, A. Vepsäläinen, L. C. Govia, C. R. Hoffer, J. Y. Qiu, D. Ristè, M. Ware, D. Kim, R. Winik, A. Melville, B. Niedzielski, J. Yoder, G. J. Ribeill, T. A. Ohki, H. K. Krovi, T. P. Orlando, S. Gustavsson, and W. D. Oliver, Deep-neural-network discrimination of multiplexed superconducting-qubit states, Phys. Rev. Appl. **17**, 014024 (2022).

[36] R. Barends, J. Kelly, A. Megrant, D. Sank, E. Jeffrey, Y. Chen, Y. Yin, B. Chiaro, J. Mutus, C. Neill, P. O'Malley, P. Roushan, J. Wenner, T. C. White, A. N. Cleland, and J. M. Martinis, Coherent Josephson qubit suitable for scalable quantum integrated circuits, Phys. Rev. Lett. **111**, 080502 (2013).

[37] F. Mezzadri, How to generate random matrices from the classical compact groups, arXiv:math-ph/0609050.

[38] K. Życzkowski, K. A. Penson, I. Nechita, and B. Collins, Generating random density matrices, J Math. Phys. **52**, 062201 (2011).

[39] C. W. Warren, J. Fernández-Pendás, S. Ahmed, T. Abad, A. Bengtsson, J. Biznárová, K. Debnath, X. Gu, C. Križan, A. Osman, A. Fadavi Roudsari, P. Delsing, G. Johansson, A. Frisk Kockum, G. Tancredi, and J. Bylander, Extensive characterization and implementation of a family of three-qubit gates at the coherence limit, Npj Quantum Inf. **9**, 44 (2023).

[40] A. Kreuzer, T. Krumrey, H. Tohamy, A. Ionita, H. Rotzinger, and A. V. Ustinov, Stacked Josephson junctions for quantum circuit applications, arXiv:2503.11437.

[41] A. M. Gunyhó, S. Kundu, J. Ma, W. Liu, S. Niemelä, G. Catto, V. Vadimov, V. Vesterinen, P. Singh, Q. Chen, and M. Möttönen, Single-shot readout of a superconducting qubit using a thermal detector, Nat. Electron. **7**, 288 (2024).

[42] A. Aasen, "CREMST.", 2025, https://github.com/AdrianAasen/CREMST

[43] J. Cotler, and F. Wilczek, Quantum overlapping tomography, Phys. Rev. Lett. **124**, 100401 (2020).

[44] G. Zhu, D. G. Ferguson, V. E. Manucharyan, and J. Koch, Circuit QED with fluxonium qubits: Theory of the dispersive regime, Phys. Rev. B **87**, 024510 (2013).

[45] F. B. Maciejewski, Z. Puchała, and M. Oszmaniec, Operational quantum average-case distances, Quantum **7**, 1106 (2023).

[46] M. A. Nielsen, and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2012).

Part III

Adaptive measurement strategies

7

READOUT NOISE IN OPTIMAL MEASUREMENT STRATEGIES

Extracting information from quantum systems is inherently difficult, not only because ideal quantum measurements are destructive, but also because the physical measurements are noisy or constrained. Therefore, it is important to use measurement strategies that are, on average, as informative as possible. The choice of an optimal strategy depends on both the prior assumptions that can be made about the system and the quality metric used to assess performance. General optimal strategies might not be feasible to implement, for example, because they require projective measurements on entangled bases [172]. While optimal and practical measurement strategies are well understood for many circumstances, little is known about how these strategies fare with varying degrees of readout noise. In this chapter, our aim is to clarify one of the underlying information-theoretic mechanisms used to find optimal measurement strategies and to show that this mechanism stops working even with vanishingly small amounts of readout noise.

This chapter is structured as follows. In Section 7.1 the estimator and metric used for optimality is explicitly defined and justified from an experimental point of view. In Section 7.2 we give a brief review of some types of optimal measurement strategies, including strategies with classical and quantum communication. Lastly, in Section 7.3 we investigate how adaptive strategies are impacted by simple readout noise.

7.1 DEFINING THE ESTIMATION PROBLEM

As in Part II, we will use quantum state tomography as our benchmarking framework. The quantum infidelity (Section 2.9.1) between the estimated state $\hat{\rho}$ and the target state ρ_0 will serve as our quality metric. In particular, we will focus on how quantum infidelity scales with the number of measurements used. The rate at which quantum infidelity reduces to zero is closely related to an estimator's ability to reduce the variance of the estimate. Recall from Equation 3.2, that the mean squared error of an estimate has two contributions, the bias and the variance,

$$E[(\hat{O} - \mathrm{Tr}(\rho_0 O))^2] = b_O^2 + \mathrm{var}[\hat{O}^2]. \tag{7.1}$$

In Part II, we have focused on reducing the bias term by introducing readout-error-mitigated estimators. In this part, we investigate how the choice of measurement settings affects the variance of the estimator. The bias is intrinsic to the estimator and remains constant regardless of the number of measurements performed. In contrast, the variance decreases as the number of measurements increases. It is known that for most sensible measurement strategies, the infidelity has a worst-case scaling of $\propto 1/\sqrt{N}$ [1] [173, 174]. According to the Gill-Massar bound, we have an upper limit on the optimal infidelity scaling achievable by any estimator, which is proportional to $1/N$ [175]. Consequently, an estimator is considered asymptotically optimal if it obtains the scaling $1/N$ as N approaches infinity, despite further improvements to the scaling coefficient still being possible.

In most experiments, the target states are pure states and it would therefore seem natural to use a pure-state estimator when trying to estimate the states. There are, however, two major drawbacks in restricting a state estimator to only pure states. Firstly, the space of pure states is not convex, and it is not guaranteed that the likelihood function is monoidal in this space [45]. The nonconvexity of the estimator also makes the estimates incompatible with both uncertainty estimates and error bars. Secondly, since it is impossible to prepare ideal pure states in any experimental setup, the prepared state is always mixed and is therefore not in the estimator support. Due to these limitations, our focus will be exclusively on mixed-state estimators. To ensure that the estimators perform optimally on average, the target states will be randomly chosen from a set of Haar-random pure states. Estimating pure and close-to-pure states in mixed-state estimators is particularly challenging because of the behavior of estimators when target states are close to the boundary of the estimator domain. An intuitive explanation for this is provided in Appendix C.1.

In summary, our investigation will focus on mixed-state estimators with pure target states. The quality measure is the asymptotic scaling of the quantum infidelity in the number of measurements. The quantum infidelity is averaged over Haar-random target states.

7.2 OPTIMAL MEASUREMENT STRATEGIES

In this section, we provide a concise overview of various optimal measurement strategies, each requiring different experimental resources. To begin with, we review protocols that utilize classical and quantum communication in Section 7.2.1 and Section 7.2.2. Following this, we address optimal static methods in Section 7.2.3. A tabulated summary of optimal strategies is presented in Section 7.2.4.

7.2.1 *Collective measurements*

The most general type of measurements that can be performed on an ensemble of quantum states are *collective measurements*. Instead of measuring each sample one after another, the whole ensemble of states is measured together at the same time. This type of measurement can be described by a set of measurement operators that span the collective Hilbert space of all states in the ensemble [176, 177]. This allows for quantum communication between measurements because beyond-classical correlations can be generated between the samples with a collective measurement operator.

---

[1] Do not confuse the $1/\sqrt{N}$ infidelity scaling with the scaling of the variance of a mean observable, such as the mean spin $\langle\sigma_z\rangle$. The standard deviation of independent mean observables scales as $\sigma_{\text{mean}} = \sigma/\sqrt{N}$, where σ is the variance of an individual random variable. For more complicated observables, such as infidelity, the observables are no longer independent and this simple relation no longer holds.

With collective measurements, the single-qubit infidelity scale as

$$\begin{aligned} &1-F=\tfrac{1}{N}+\mathcal{O}\left(\tfrac{1}{N^2}\right) \quad \text{for pure states with pure state estimators [178] and} \\ &1-F=\tfrac{3+2\langle r\rangle}{4N}+\mathcal{O}\left(\tfrac{1}{N^2}\right) \quad \text{for mixed states with mixed state estimators [172],} \end{aligned} \tag{7.2}$$

where $\langle r\rangle$ is the average purity of the prior distribution[2]. Although these estimators have the highest information extraction efficiency, the significant hardware requirements to execute collective measurements have, to the best of our knowledge, never been implemented in practice.

7.2.2 *Adaptive measurement strategies*

Adaptive measurement strategies reduce the complexity from collective measurement by only measuring a single copy of the quantum state a time[3], but allows classical communication between measurements. Adaptive quantum state reconstruction strategies have been adopted in various experimental qubit platforms, including photonic systems [179–182] and superconducting architectures [183]. We will differentiate between two types of adaptive strategies, those that update the measurement setting only once and those that update the measurement setting more often.

Two-step strategies

Single update methods are the most economical of the two approaches, since updating measurement settings can be very slow compared to continuing measurements in the same setting. If the target state is known in advance, the best strategy is to perform a projective measurement in the eigenbasis of the state. This measurement strategy has the potential to exceed the Gill-Massar bound [182], as they create a biased estimator. However, such a biased estimator is of little use when the target state is already known. An unbiased version of this strategy, with an unknown target state, can be created by splitting the estimation process into two steps [175, 184]:

1. Perform an initial state estimate $\hat{\rho}_{\text{init}}$ based on an informationally complete measurement, such as the $\mathcal{M}_{\text{Pauli-6}}$ measurements.

2. Find the eigenstate $|\lambda_1\rangle$ with the largest eigenvalue λ_1 of $\hat{\rho}_{\text{init}}$, and perform the remaining measurements using the generalized measurement $U\mathcal{M}_{\text{Pauli-6}}U^{\dagger}$, where U is a unitary rotation matrix that rotates the all-zero qubit state to this eigenstate of $\hat{\rho}_{\text{init}}$, $U|0\ldots0\rangle = |\lambda_1\rangle$.

According to Refs. [75, 173], it is optimal to split the total measurement budget evenly between the two steps. This protocol was found to be sensitive to the prior used to initialize the estimator [184].

Although these methods were found to be optimal, the eigenbasis of the initial estimate is very often entangled and could cause additional overhead to implement experimentally. The underlying mechanism was found to be related to the singular values of the Fisher information matrix (FIM) [174, 179], which allowed Ref. [179] to develop a two-step strategy that uses only

---

[2] When $\langle r\rangle \to 1$ for mixed states, it does not mirror the pure state limit because mixed state estimators have slightly worse asymptotic scaling relative to pure state estimators when estimating pure states.

[3] Sometimes called *local measurements*.

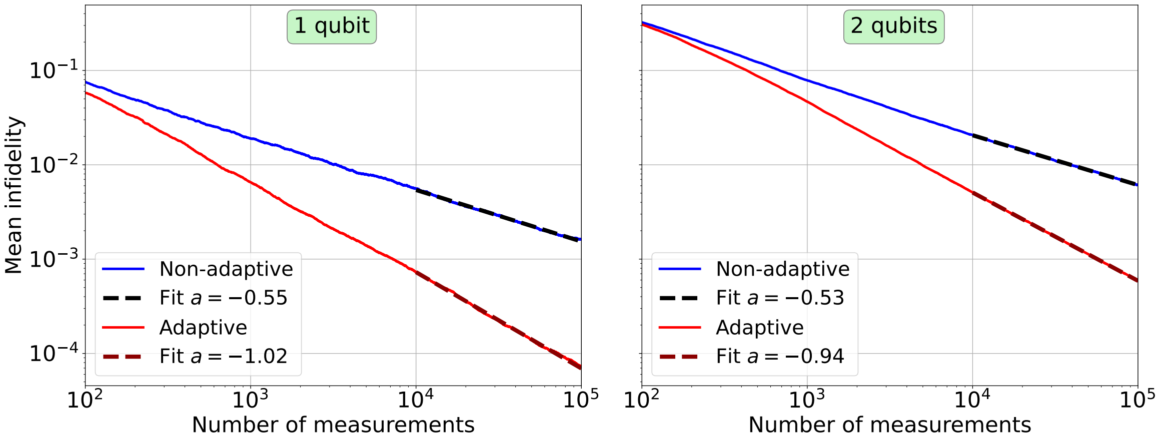


Figure 7.1: Comparison of mean infidelity scaling between adaptive and non-adaptive Bayesian measurement strategies for single- and two-qubit state reconstruction. A power-law curve N^a is fitted between 10^4 and 10^5 measurements. **Left**: Mean infidelity of single-qubit state reconstruction averaged over 100 pure Haar-random single-qubit states. The non-adaptive approach uses static Pauli-6 measurements. **Right:** Mean infidelity of two-qubit state reconstruction averaged over 100 pure Haar-random states. Both adaptive and non-adaptive strategies uses factorized measurement, i.e. $\mathcal{M} = \mathcal{N}_1 \otimes \mathcal{N}_2$. The non-adaptive approach uses static factorized Pauli-6 measurements, i.e. $\mathcal{M} = \mathcal{M}_{\text{Pauli-6}} \otimes \mathcal{M}_{\text{Pauli-6}}$.

the tensor products of single-qubit measurements. This approach is optimal for certain entangled states, though it is not universally optimal for all states.

Multi-step strategies

Multi-step strategies are a form of *adaptive experimental design*, where the estimator continuously adapts the measurement strategy based on prior measurement outcomes D. These strategies use Bayesian inference to continuously integrate measurement outcomes D into a posterior distribution $\pi_f(\rho)$. A new measurement setting α is chosen by maximizing the expected value of a utility function $U(\alpha, \pi_f(\rho))$ based on the current posterior distribution $\pi_f(\rho)$ [184],

$$\alpha_n = \underset{\alpha}{\operatorname{argmax}} \sum_i \int d\rho\, \pi_f(\rho)\, \operatorname{Tr}(\rho M_i^\alpha) U(\alpha, \pi_f(\rho)). \tag{7.3}$$

There are two common approaches to choosing the utility function, optimizing for a specific parameter, e.g. the infidelity, or maximizing information gain. Refs. [185, 186] demonstrated that using the information gain approach significantly simplifies the computation of the optimization objective in Equation 7.3. The optimization can be rewritten as

$$\alpha_n = \underset{\alpha}{\operatorname{argmax}} \left(H\left(\int d\rho\, \pi_f(\rho)\, \operatorname{Tr}(\rho M_i^\alpha) \right) - \int d\rho\, \pi_f(\rho) H(\operatorname{Tr}(\rho M_i^\alpha)) \right), \tag{7.4}$$

where $H(p_i) = -\sum_i p_i \log(p_i)$ denotes the Shannon entropy of the discrete probability distribution p. The advantage of this reformulation is that it relies solely on the current posterior distribution $\pi_f(\rho)$, and requires only the computation of discrete Shannon entropies. The terms in Equation 7.4 can also be intuitively understood: the first term promotes measurements along the measurement axis with the highest uncertainty (high entropy), while the second term favors measurements where the most likely states have low uncertainty (low entropy). Combined, the optimization seeks the measurement setting α that targets regions of high uncertainty in the outcome distribution while remaining consistent with the most probable candidate states. In Figure 7.1, we show that the information-gain strategy achieves optimal asymptotic infidelity scaling for both single- and two-qubit state reconstruction. Notably, in the two-qubit case optimal scaling is reached using only separable measurement operators, whereas the two-step approach can generally not achieve optimal scaling without measurements in entangled bases [174, 179, 185].

7.2.3 *Mutually unbiased bases*

While adaptive strategies are powerful measurement tools, it is worth comparing them with the simplest form of measurements, *static measurements*, which do not use any form of communication. These fixed strategies are strictly worse than adaptive ones and cannot achieve optimal scaling for the estimation problem defined in Section 7.1. Nevertheless, their inherent simplicity and widespread usage make them an important case worth mentioning. The static measurement strategy that performs best for quantum state reconstruction was shown to be *mutually unbiased bases* (MUB) [187]. Two orthonormal bases, $\{v_i\}$ and $\{u_j\}$, are said to be mutually unbiased if all basis vectors have the same mutual overlap,

$$|\langle v_i|u_j\rangle|^2 = \frac{1}{d}, \quad \forall\, i,j \in [1,d]. \tag{7.5}$$

Performing projective measurements on these sets of bases ensures that there is the smallest possible overlap between the different bases, which maximizes the information gain when switching between the bases. It is still unknown whether MUB exist in every Hilbert space dimension, but they are known to exist for dimensions that are powers of prime numbers [187]. Specifically, MUB are known for n-qubit systems, which have Hilbert space dimensions of 2^n.

7.2.4 *Summarizing optimal strategies*

Concluding this brief overview of optimal measurement strategies, we give a tabulated overview, Table 7.1, of the different approaches presented in this chapter, as well as some more recently developed strategies with respect to the estimation problem outlined in Section 7.1.

Method	Meas. type	Communication	Worst-case infidelity scaling	Ref.
Collective measurements	Entangled	Quantum	$\propto 1/N$	[172]
Two-step adaptive	Entangled	Classical	$\propto 1/N$	[173, 179]
Two-step adaptive	Separable	Classical	$\propto 1/\sqrt{N}$	[179]
Multi-step adaptive	Separable	Classical	$\propto 1/N$	[185]
Adaptive linear regression	Separable	Classical	$\propto 1/N$	[182]
Adaptive neural network	Separable	Classical	$\propto 1/N$	[188]
MUB	Separable	None	$\propto 1/\sqrt{N}$	[173]

Table 7.1: Summarizing the different worst-case infidelity scaling of different measurement strategies for the estimation problem defined in Section 7.1. The necessary experimental resources for the different strategies are listed, detailing whether they involve projective measurements on entangled bases or use communication between measurements.

7.3 NOISE IN OPTIMAL MEASUREMENT STRATEGIES

In the review of optimal strategies so far, we have limited ourselves to the cases where measurements are described by idealized projective measurement operators. In practice, measurements will be affected by noise, which could change what measurements are optimal. Other works have investigated readout noise in the context of optimal measurement strategies. The decrease in reconstruction infidelity for static measurement approaches due to noise was examined in Refs. [189, 190]. Ref. [191] developed a new optimal static measurement approach, which considers readout noise modeled as a noisy unitary gate occurring prior to an ideal readout. In this section, we will go a step further and investigate what happens to optimal adaptive strategies when readout noise is introduced.

For these considerations, we have to assume that we have access to the noisy generalized measurement operators and can perform readout-error mitigation. Otherwise, the reconstruction infidelity saturates because the state estimator will converge to a noisy reconstructed state instead of the ideal target state. In this section we assume that we have the exact noisy measurement operators, while in Ref. [D] we also consider practical scenarios where the measurement operators need to be reconstructed.

7.3.1 *Mechanism behind optimality*

What follows is a summary of the comprehensive derivation found in Appendix C.2. To maintain conciseness, we focus solely on the primary quantities relevant to the upcoming sections.

The basis of this analysis is the mechanism derived in Ref. [174], and refined in Ref. [179]. The central result of these works is that the state reconstruction infidelity scales, on average, with the singular values of the FIM $\mathcal{I}$ (defined in Equation 2.63),

$$\langle 1 - F \rangle = \sum_{i=2}^{\nu+1} \frac{1}{\sigma_i}, \tag{7.6}$$

where σ_i are the singular values of the FIM and ν is the number of degrees of freedom in the estimated quantum state. The sum in Equation 7.6 starts from $i = 2$ because the singular value σ_1 is related to the normalization of the state and is therefore fixed. Following our estimation problem outlined in Section 7.1, we only consider full-rank estimators, and therefore $\nu = 4^n - 1$. The asymptotic scaling of the singular values is proportional to N, which means that the expected infidelity scales proportionally to $1/N$ (see Equation C.20 in Appendix C.2.2). Equation 7.6 is only valid if all singular values are non-vanishing [174, 179]. However, the singular values are, in general, only all non-vanishing if the estimator degrees of freedom match the number of degrees of freedom in the target state. For example, if the target state is pure, there are only $2(2^n - 1) < \nu$ non-vanishing singular values. This means that when the target state is pure or close-to-pure, many of the singular values vanish. As a result, the scaling described in Equation 7.6 is only valid for the non-vanishing singular values, and the remaining singular values must be treated separately, resulting in contributions which scale as $1/\sqrt{N}$ [179][4].

---

[4] The suboptimal scaling can be explained intuitively by considering that the estimator is over-parameterized, leading to some eigenvalues of the reconstructed state being zero. Thus, these eigenvalues lie along the edge of the allowable parameter space. In Appendix C.1, we demonstrate that infidelity near the boundary of the parameter space scales with a square-root degradation compared to when the actual value resides within the interior of the parameter space.

The singular values of the FIM are related to the eigenvalues of the estimated density matrix. The rows and columns of $\mathcal{I}$ are proportional to the vector [179]

$$\omega_\gamma^\alpha = \begin{pmatrix} \sqrt{\lambda_1} M_\gamma^\alpha \left|\lambda_1\right\rangle \\ \sqrt{\lambda_2} M_\gamma^\alpha \left|\lambda_2\right\rangle \\ \vdots \\ \sqrt{\lambda_{2^n}} M_\gamma^\alpha \left|\lambda_{2^n}\right\rangle \end{pmatrix}, \quad \mathcal{I} \propto \sum_\alpha \sum_\gamma \frac{1}{p_\gamma^\alpha} \omega_\gamma^\alpha (\omega_\gamma^\alpha)^\dagger, \tag{7.7}$$

where λ_i is the eigenvalue of the target state, with the corresponding eigenstate $|\lambda_i\rangle$, and γ labels the outcome of measurement setting α with the respective probability p_γ^α. From Equation 7.7 we see that if any eigenvalues $\lambda_i = 0$, whole rows and columns in $\mathcal{I}$ will be zero. The FIM effectively behaves as a rank-deficient estimator since the rank of the target state is lower than the estimator rank. To recover optimal scaling, the measurement strategy should ensure that the estimator behaves as a full-rank estimator. The zero rows and columns of $\mathcal{I}$ can be amplified by a factor $1/p_\gamma^\alpha$ by choosing a measurement setting $\mathcal{M}^\alpha$ such that the entries of ω_γ^α are all approximately equal[5]

$$||\sqrt{\lambda_1} M_\gamma^\alpha \left|\lambda_1\right\rangle ||_2 \approx ||\sqrt{\lambda_2} M_\gamma^\alpha \left|\lambda_2\right\rangle ||_2 \approx \cdots \approx ||\sqrt{\lambda_{2^n}} M_\gamma^\alpha \left|\lambda_{2^n}\right\rangle ||_2, \tag{7.8}$$

where the norm is defined $|| \left|\psi\right\rangle ||_2 = \sqrt{\langle\psi|\psi\rangle}$. If there are small eigenvalues λ_i, the corresponding terms in Equation 7.8 become vanishingly small. Therefore, for all terms to be equalized, one should choose a measurement setting $\mathcal{M}^\alpha$ that has measurement operators that are orthogonal to all eigenstates with non-zero eigenvalues simultaneously. Noting that

$$p_\gamma^\alpha = \sum_i ||\sqrt{\lambda_i} M_\gamma^\alpha \left|\lambda_i\right\rangle ||_2^2, \tag{7.9}$$

if Equation 7.8 holds, the vanishing rows in Equation 7.7 are effectively canceled out by the $1/p_\gamma$ factor, resulting in a matrix with a new non-zero singular value. This process can be repeated with a new linearly independent measurement operator M_γ^α that satisfies Equation 7.8, each increasing the number of non-zero singular values of $\mathcal{I}$ by one.

We can summarize the above discussion succinctly. To recover optimal asymptotic scaling, the measurement strategy must ensure that all vanishing singular values of the Fisher information matrix become non-zero. This is achieved by selecting measurement operators that are orthogonal to the eigenvectors associated with non-zero eigenvalues of the target state. In other words, one needs to find measurement operators of the form

$$M_\gamma^\alpha = \tau_\gamma^\alpha \left|\psi_\gamma^\alpha\right\rangle \left\langle\psi_\gamma^\alpha\right| \text{ such that } \left\langle\psi_\gamma^\alpha \middle| \lambda_k\right\rangle = 0 \text{ for all } k \text{ with } \lambda_k \neq 0, \tag{7.10}$$

and the number of such measurement operators required is given by [179]

$$N_\mathcal{M} = \begin{cases} 4^n - 1 - 2(2^n - 1), & \text{for pure target states,} \\ 4^n - 1 - \left(R2^{n+1} - R^2 - 1\right), & \text{for target states of rank } R, \end{cases} \tag{7.11}$$

---

[5] If an eigenvalue λ_i is exactly zero, then one needs to find measurements where p_γ^α is also exactly zero for all non-zero eigenvalues λ_γ. The zero rows in Equation 7.7 is then canceled out by the $1/p_\gamma^\alpha$ factors.

which is equivalent to the condition stated in Equation 7.8.

The two adaptive approaches discussed in Section 7.2.2 achieve optimal infidelity scaling in slightly different ways. The two-step adaptive strategy achieves equalization of the terms in Equation 7.8 by realigning the measurement operators with the non-zero eigenstates of an intermediate state estimate $\hat{\rho}$ [174, 179]. These intermediate eigenstates are often entangled, which means that optimal scaling cannot always be achieved with separable measurements [179]. Multi-step adaptive Bayesian strategies are not directly based on the discovery of orthogonal measurement operators, but continuously update the measurement strategy based on the uncertainty of the posterior distribution $\pi_f(\rho)$ [180, 185]. Summing over the different measurement settings α in Equation 7.7 is likely to cause $\mathcal{I}$ to be full-rank.

7.3.2 *Noise in optimal measurement strategies*

The main question we aim to address in this chapter is whether it is possible to retrieve optimal infidelity scaling in the presence of readout noise by applying readout-error mitigation. In this brief analysis, we will only focus on depolarizing readout noise, given by the depolarizing channel in Equation 2.34, but we expect our results to generalize to other more complicated noise models.

Two-step strategies

In the case of two-step strategies, it is already possible to make a statement about the recoverability of optimal asymptotic scaling. Using the readout-error-mitigated likelihood function in Equation 12 of Chapter 4 and rederiving the vector ω_γ^α, the ideal measurement operators in Equation 7.7 are replaced by their depolarized counterparts

$$M_\gamma^\alpha \to \mathcal{E}_{\text{depol}}(M_\gamma^\alpha) = \tilde{M}_\gamma^\alpha. \tag{7.12}$$

The corresponding outcome probabilities are then lower-bounded by the depolarizing strength p,

$$p_\gamma^\alpha \geq \tau_\gamma^\alpha \frac{p}{2^n}, \tag{7.13}$$

which ensures that the prefactor $1/p_\gamma^\alpha$ in Equation 7.7 remains finite for all measurement settings α. Consequently, the zero rows and columns of the FIM can no longer be canceled[6]. It follows that no two-step approach can recover optimal asymptotic infidelity scaling in the presence of depolarizing readout noise. The detailed derivation of Equation 7.13 is found in Appendix C.2.

Multi-step strategies

The fact that two-step approaches cannot reach optimal scaling does not immediately rule out the usefulness of alternative approaches, such as multi-step methods based on Bayesian strategies. Unlike Fisher information-based methods, Bayesian strategies involve optimizing a utility function that may not be directly linked to the FIM. To investigate Bayesian methods with noise, we conduct numerical simulations using the information gain approach described in Refs. [180, 185]. Similarly to the two-step approach, we focus on readout-error-mitigated states

---

[6] Note that choosing $\tau_\gamma \to 0$ does not help since $\frac{1}{p_\gamma^\alpha}\omega_\gamma^\alpha(\omega_\gamma^\alpha)^\dagger \propto \tau_\gamma^\alpha$.

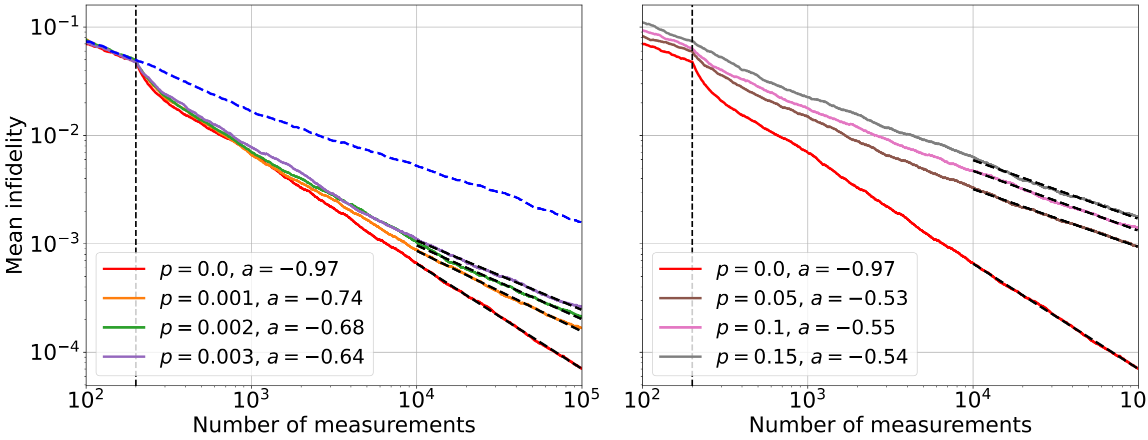


Figure 7.2: Mean infidelity for single-qubit readout-error-mitigated quantum state tomography, averaged over 100 Haar-random initial states. Each curve corresponds to depolarizing readout noise of strength p, applied according to Equation 2.34. A black dashed power-law curve N^a is fitted between 10^4 and 10^5 measurements. The vertical dashed line indicate when adaptive measurements were selected. In the readout-error-mitigation protocol we have used the exact noisy measurement operators. **Left:** A very small amount of depolarizing noise is applied. Initially most curves follows the $1/N$ scaling, but once reconstruction accuracy get high enough, a splitting occurs and the ideal scaling is lost. The dashed blue curve is the non-adaptive curve from Figure 7.1. **Right:** Stronger depolarizing noise. Loss of optimal asymptotic scaling causes dramatic loss in accuracy for larger number of measurements compared to the noiseless case. While increasing noise even further causes only minor loss in reconstruction accuracy.

using the protocol in Ref. [A], because unmitigated methods exhibit only transient scaling and saturate due to a non-vanishing bias term.

In Figure 7.2, we show the averaged infidelity for readout-error-mitigated quantum states for various levels of depolarizing readout noise. Introducing even a minimal level of depolarizing noise changes the asymptotic scaling from $1/N$ to $\approx 1/\sqrt{N}$. At first, the low-noise cases follow the optimal infidelity curve, but once the reconstruction accuracy is large enough, a split occurs between noiseless and noisy curves. This shows that introducing any amount of readout noise disproportionally reduces the quality of the state reconstruction compared to the noiseless case. This result is significant because, for most realizations of quantum systems, there is a small amount of noise that often manifests itself as depolarizing noise. This raises the question whether these adaptive measurement strategies can be effectively implemented in practice. It is notable that the use of adaptive strategies does improve the prefactor of the scaling behavior, which shows that adaptively updating the measurement does improve the overall reconstruction accuracy by a constant factor. Further analysis of the asymptotic scaling behavior is presented in Ref. [D].

Part IV

Summary

8

CONCLUSION AND OUTLOOK

In this thesis we have presented two distinct approaches to manage readout errors in quantum systems. In this closing chapter, we give some concluding remarks and highlight some open questions and directions for future research.

8.1 READOUT-ERROR MITIGATION

In Part II, we have investigated readout errors, how to model them, and how to mitigate their impact in information retrieval tasks. Across the three works presented, we have developed a comprehensive framework for both characterization and mitigation of beyond-classical readout errors and validated our protocol on a superconducting qubit experiment. The core contributions can be grouped into two main areas. First, the protocol was designed to be capable of mitigating coherent and correlated errors while only using single-qubit measurements, which makes the protocol simple and experimentally friendly. Secondly, on the experimental front we had direct control over the hardware, which allowed us to probe experimental parameters and induce noise which are typically not accessible in other cloud-based platforms.

In Chapter 4 we first introduced a hardware-agnostic readout-error-mitigation protocol capable of mitigating any readout error within the qubit manifold. The protocol was tested on a superconducting qubit where some experimental noise sources were deliberately amplified. The REMST protocol consistently improved reconstruction infidelity for all tested cases. We identified which noise sources the protocol worked particularly well for, such as reduced readout amplification and lowered readout resonator populations, seeing state reconstruction infidelity improvements of a factor up to 30. As a tradeoff for such comprehensive noise handling, the protocol scales exponentially in the number of qubits, making it infeasible for system sizes much larger than five qubits.

In Chapter 5, we addressed the main scalability limitation of the REMST protocol. The correlated REMST protocol systematically mitigates correlated readout errors for arbitrary multi-qubit observables in large qubit systems. The protocol builds on previous works [36, 161] by including a deterministic implementation of overlapping detector tomography and providing a simple and heuristic-free clustering algorithm. We demonstrated that the protocol works on realistic noise extracted directly from a four-qubit superconducting system, and artificial strong coherent and correlated noise models. For experimental readout errors, the protocol yielded an approximate twofold reduction in the MSE of error-mitigated observables compared to a naive application of the protocol introduced in Chapter 4.

In Chapter 6, we investigated the use of state reconstruction infidelity as a proxy for readout quality. Using the REMST protocol developed in Chapter 4, it became possible to use the increased variance of the error-mitigated estimators to infer the total strength of the readout noise in the system. We proposed to use the number of measurements required to reach a given infidelity threshold as a more comprehensive readout calibration tool and compared it to the standard readout quality metric of assignment fidelity. When calibrating the readout of the superconducting qubit experiment, the two metrics agreed on the optimal readout amplitude, indicating that there are predominantly classical readout errors in the investigated system. In addition to testing the quality metric, three-qubit readout-error mitigation was performed, showing an improvement in infidelity by an order of magnitude. Correlated readout noise was also investigated between the different qubits on the chip, where the proposed correlation coefficients identified a pair of qubits to have correlated readout, which could be attributed to the readout resonators being too close in frequency. It was also verified experimentally that introducing depolarizing noise did not induce correlated errors between the qubits.

Combined, these works form the basis for a comprehensive readout error management framework. The proposed protocols provide valuable tools for improving readout calibration in quantum systems, a topic of ongoing investigation [192–194]. Our work focuses specifically on superconducting qubit systems, which exhibit particularly large readout errors [36, 132, 195] that can significantly impact error correction codes [196, 197]. With increasing interest in multiplexed qubit readout, where correlated readout errors occur naturally, correlated REMST could prove valuable [159, 198–200]. The tools investigated in Chapter 6 could be used to collectively tune the readout parameters for multiple qubits simultaneously, rather than individually optimizing the readout parameters [F]. Beyond this, an obvious next step is to implement the correlated REMST protocol in larger experiments and use the protocol to extract readout-error-mitigated observables for relevant problems, such as determining the critical exponents in phase transitions [167–169]. It would also be interesting to apply these protocols to other platforms such as photonic systems, which do not immediately follow the standard POVM formalism [201, 202]. A potential application could be to mitigate the effect of the lens used to collect photons after scattering off a sample in a quantum polarimeter [203]. Here, one could potentially correct state preparation errors separately by calibrating the calibration states without the collection lens.

8.2 ADAPTIVE MEASUREMENT STRATEGIES WITH NOISE

In Part III and Chapter 7 we investigated how readout noise affects adaptive measurement strategies for quantum state reconstruction. We focused specifically on pure state reconstruction in mixed state estimators, which is a particularly challenging estimation problem [45, 180]. Adaptive measurement strategies have the potential to improve the asymptotic scaling of reconstruction infidelity from $1/\sqrt{N}$ to $1/N$. We showed that with a very small amount of depolarizing readout noise, the benefit gained from performing adaptive measurements vanishes. This breakdown is derived analytically and demonstrated numerically for multi-step strategies. This result puts into question the viability of adaptive strategies in practical state reconstruction tasks, as noise can never be entirely removed. Our investigation was solely focused on depolarizing noise, which is sufficiently general to cover most relevant cases. However, it is possible that device-specific noise models might reveal methods capable of maintaining better asymptotic scaling despite noise. It is still unclear whether these results extend to more general measurement strategies, such as collective measurements, that use quantum communication [172]. The next step would be to verify this property experimentally, as current adaptive experiments do not perform readout-error mitigation, allowing only the extraction of transient scaling [179–181]. With these negative results for adaptive measurement strategies, it would be interesting to see if the implementation of measurement strategies that are aware of hardware noise could recover some of the benefit of adaptive measurement [191]. Furthermore, it would be interesting to investigate whether it is possible to derive a direct connection between the Fisher-information adaptive strategy [174, 179] and the Bayesian utility function approach [180, 185]. Further research on some of these questions is detailed in Ref. [D].

> "*Lastly, she pictured to herself how this same little sister of hers would, in the after-time, be herself a grown woman; and how she would keep, through all her riper years, the simple and loving heart of her childhood: and how she would gather about her other little children, and make their eyes bright and eager with many a strange tale, perhaps even with the dream of Wonderland of long-ago: and how she would feel with all their simple sorrows, and find a pleasure in all their simple joys, remembering her own child-life, and the happy summer days.*"
>
> – Lewis Carroll, Alice in Wonderland

A

PROJECTION-VALUED MEASURE

A.1 ORTHONORMALITY

In this appendix, we aim to demonstrate that if Equation 2.17 holds for a set of projectors $P_i \in \mathcal{P}$, then each P_i is orthonormal and that $|\mathcal{P}| = \dim(\mathcal{H})$.

We start by rewriting the projectors $P_i \in \mathcal{P}$ in terms of their eigenbasis Equation 2.17,

$$\sum_i |e_i\rangle \langle e_i| = \mathbb{1}, \tag{A.1}$$

where $|e_i\rangle \langle e_i| = P_i$. Consider the inner product $\langle e_j | e_k \rangle$,

$$\langle e_j | e_k \rangle = \sum_i \langle e_j | e_i \rangle \langle e_i | e_k \rangle \tag{A.2}$$

This can be reformulated in terms of the Gram matrix $G_{ij} = \langle e_j | e_k \rangle$, which then states

$$G = G^2, \tag{A.3}$$

which is only possible if the Gram matrix itself is a projector. The Gram matrix must have full rank because the basis vectors span the full Hilbert space, leaving us with the only solution $G = \mathbb{1}$, which means that

$$\langle e_j | e_k \rangle = \delta_{jk} \tag{A.4}$$

and

$$P_j P_k = P_j \delta_{jk}. \tag{A.5}$$

Since each operator is orthonormal and the rank of the full sum of projectors is equal to the dimension of the Hilbert space, the number of elements in the set must be equal to the dimension of the Hilbert space, $|\mathcal{P}| = \dim(\mathcal{H})$.

B

QUANTUM STATE RECONSTRUCTION

B.1 ITERATIVE MAXIMUM LIKELIHOOD ESTIMATION

This appendix outlines the iterative maximum likelihood estimation (IMLE) method used in Refs. [B, C] and originally introduced in Ref. [80]. IMLE, like all other MLE approaches, is based on finding the quantum state ρ that maximized the likelihood function

$$\mathcal{L}(\rho) = \Pi_i \operatorname{Tr}(\rho M_i)^{f_i}, \tag{B.1}$$

where $f_i = n_i/N$ denotes the relative frequency of outcome i, where n_i is the count of observations for outcome i, N is the total number of measurements, and M_i is the measurement operator associated with outcome i. To find the state $\hat{\rho}$ that is most likely to have produced the observed outcomes, one can define a fixed-point equation

$$\hat{\rho}^{k+1} = R(\hat{\rho}^k)\hat{\rho}^k R(\hat{\rho}^k), \tag{B.2}$$

where the iterative operator is defined

$$R(\hat{\rho}^k) = \sum_i \frac{f_i}{\operatorname{Tr}(\hat{\rho}^k M_i)} M_i. \tag{B.3}$$

The update in Equation B.2 is applied iteratively until it converges, that is, such that $\hat{\rho}^{k+1} \approx \hat{\rho}^k \approx \hat{\rho}^\infty$, at which point you have the final state estimate,

$$\hat{\rho}_{\text{MLE}} = \hat{\rho}^\infty \approx \hat{\rho}^k. \tag{B.4}$$

B.2 DERIVING THE BAYESIAN UPDATE MECHANISM

The goal of this section is to derive the Bayesian update mechanism in Equation 2.51 by iteratively including the measurement of a data set D into a prior distribution π_0 by repeated application of Baye's law.

To start with, consider the scenario with an initial data set $D_0 = \{o_0\}$. We want to obtain an updated probability distribution of some test state ρ given the new data set, written as a conditional probability $P(\rho|D_0)$. Applying Baye's rule, we have

$$P(\rho|D_0) = \frac{P(D_0|\rho)P(\rho)}{P(D_0)}. \tag{B.5}$$

$P(\rho)$ can be understood as the prior distribution π_0, which is based on some already known information on what states it is possible to produce without any measurements being made. In an completely agnostic estimation setup, π_0 should be some kind of uniform distribution in Hilbert space. $P(D_0)$ is a normalization constant necessary for the right-hand side of Equation B.5 to be normalized. To simplify the notation, the normalization will be omitted as it does not depend on the test state ρ. $P(D_0|\rho)$ is the probability that the data set D_0 occurs given a specific test state ρ. Using the Born rule (Equation 2.38), this probability becomes

$$P(D_0|\rho) = \mathrm{Tr}(\rho M_{o_0}). \tag{B.6}$$

Incorporating Equation B.6 into Equation B.5, we have

$$P(D_0|\rho) \propto \mathrm{Tr}(\rho M_{o_0})\pi_0, \tag{B.7}$$

where the proportionality comes from the neglected normalization constant. Relabeling $P(D_0|\rho)$ to π_1, we can include a new data point $D_1 = \{o_1, o_0\}$ and repeat the previous process to obtain

$$P(D_1|\rho) \propto \mathrm{Tr}(\rho M_{o_1})\pi_1 \propto \mathrm{Tr}(\rho M_{o_1})\,\mathrm{Tr}(\rho M_{o_0})\pi_0. \tag{B.8}$$

Recognizing the pattern of the likelihood function $\mathcal{L}(\rho)$ in Equation 2.48, we can include the complete data set $D = \{o_{f-1}, \ldots, o_1, o_0\}$, which yields the Bayesian update rule for the posterior distribution π_f,

$$\pi_f = P(D|\rho) \propto \mathcal{L}(\rho)\pi_0. \tag{B.9}$$

B.3 IMPLEMENTATION DETAILS OF BAYESIAN MEAN ESTIMATION

In this appendix, we outline the numerical implementation of BME used in Refs. [A, C, D]. The implementation is primarily based on the structure suggested in Refs. [45, 180]. The core idea of BME is to represent the likelihood function $\mathcal{L}(\rho)$ as an integrable posterior probability distribution $\pi_f(\rho)$, so that a mean estimate can be performed,

$$\hat{\rho}_{\mathrm{BME}} = \int d\rho\, \rho \pi_f(\rho). \tag{B.10}$$

This is made computationally tractable by introducing a particle swarm, where the posterior distribution $\pi_f(\rho)$ is discretized and approximated by a set of particles ρ_i and associated weights w_i,

$$\pi_f(\rho) \approx \sum_i \rho_i w_i. \tag{B.11}$$

This particle swarm representation forms the basis for implementing the Bayesian update cycle, which incorporates a data set $D = \{o_{f-1}, \ldots, o_1, o_0\}$ into the prior π_0.

B.3.1 *The Bayesian update cycle*

The prior distribution $\pi_0(\rho)$ is initialized by drawing n_{part} random particle states ρ_i from the Hilbert-Schmidt distribution of mixed states [204] and the weights equalized $w_i = 1/n_{\mathrm{part}}$. The data is integrated one by one through the Bayesian update rule in Appendix B.2, expressed in discretized form as

$$w_i^{n+1} = \frac{w_i^n \operatorname{Tr}(\rho_i O_{n+1})}{\sum_j w_j^n \operatorname{Tr}\big(\rho_j O_{n+1}\big)}, \tag{B.12}$$

where O_i is the measurement operator corresponding to outcome o_i.

This update procedure is repeated until the majority of the weight is contained in only a few particles. When this happens, the approximation becomes unreliable, and the swarm must be resampled to better approximate the actual posterior distribution. To evaluate whether the distribution is too concentrated, we have used the following quantity,

$$s = \frac{1}{\sum_i w_i^2}, \tag{B.13}$$

where resampling occurs if $s < s_{\text{thres}} = \tau \times n_{\text{part}}$, where τ is a threshold scaling. If this threshold is crossed, the distribution is resampled using the Metropolis-Hastings (MH) algorithm.

B.3.2 *Metropolis-Hastings resampling*

Resampling refines the particle swarm to better approximate the posterior distribution during updates. The key idea is to perturb existing particles so that they cluster more densely near regions of high posterior probability. The resampling follows the procedure from Appendix C of Ref. [180], summarized here.

Randomly select a particle ρ from the particle swarm, where the probability of selecting a particle ρ_i is w_i. Then uniformly perturb the particle $\rho \to \tilde{\rho}$, and accept the perturbation with probability p using the following rule,

$$p = \begin{cases} 1 & \text{when } \mathcal{L}(\tilde{\rho}) \geq \mathcal{L}(\rho), \\ \mathcal{L}(\tilde{\rho})/\mathcal{L}(\rho), & \text{when } \mathcal{L}(\tilde{\rho}) < \mathcal{L}(\rho). \end{cases} \tag{B.14}$$

If accepted, the particle state ρ is updated to the perturbed state $\tilde{\rho}$. To achieve a uniform perturbation, we use a purification scheme in which the states are embedded in a larger Hilbert space $\rho_{d\times d} \to |\psi\rangle_{d^2}$, where d is the dimension of the original Hilbert space. We will drop the index in the following. In the purified space, the perturbed state vector $|\tilde{\psi}\rangle$ is obtained as follows,

$$|\tilde{\psi}\rangle = a\,|\psi\rangle + b\frac{|g\rangle - |\psi\rangle\,\langle\psi|g\rangle}{||\,|g\rangle - |\psi\rangle\,\langle\psi|g\rangle\,||}, \tag{B.15}$$

where $a = 1 - c^2/2$, $b = \sqrt{1-a^2}$, c is drawn at random from the normal distribution $\mathcal{N}(0,\sigma)$, where σ is proportional to the width of the current posterior distribution, $\sigma = 2k\sum_i(1 - F(\rho_i, \bar{\rho}))w_i$, where $\bar{\rho} = \sum_i w_i \rho_i$ and k is a variance-strength modifier. $|g\rangle$ is a state vector with real and imaginary entries sampled from a normal distribution $\mathcal{N}(0,1)$. The purified perturbed state $|\tilde{\psi}\rangle$ is then traced back down to the original Hilbert space, where the step is accepted or rejected. The perturbation is then repeated n_{MH} more times to the same state, before finally being inserted into a new particle swarm, where all weights are equalized $w_i = 1/n_{\text{part}}$. This resampling procedure is repeated until n_{part} new particles are inserted into the new particle swarm.

Parameter	1 qubit	2 qubits
n_{part}	100	1000
τ	0.1	0.05
k	0.4	0.5
n_{MH}	50	75

Table B.1: BME parameters that differ between one- and two-qubit cases.

Explicit parameters used for the one- and two-qubit BME implementations are provided in Table B.1.

C

OPTIMAL MEASUREMENT STRATEGIES

C.1 PURE STATES IN MIXED STATE ESTIMATION

Estimating pure states proves to be more challenging than estimating mixed states when using mixed state estimators. Given the critical role that pure states play in quantum technologies and information theory, it is essential to understand their distinct behavior.

Here we examine the issue using two single-qubit states $\rho = \frac{1}{2}(\mathbb{1} + \vec{r} \cdot \vec{\sigma})$ and $\tau = \frac{1}{2}(\mathbb{1} + \vec{t} \cdot \vec{\sigma})$. Using the Bloch representation, the fidelity in Equation 2.59 can be rewritten as

$$F(\rho, \tau) = \frac{1}{2}\left(1 + \vec{r} \cdot \vec{t} + \sqrt{1 - |\vec{r}|^2}\sqrt{1 - |\vec{t}|^2}\right), \tag{C.1}$$

where we have used the identity

$$\det(\vec{r} \cdot \vec{\sigma}) = -|\vec{r}|^2. \tag{C.2}$$

In this example, we denote $\vec{r} = (s_x, s_y, s_x)$ as the Bloch vector of the target state, where s_i is the ith spin, and $\vec{t} = (\langle \sigma_x \rangle, \langle \sigma_z \rangle, \langle \sigma_z \rangle)$ as the estimated Bloch vector. The expected deviation between these two Bloch vectors can be computed,

$$\begin{aligned} \langle |\vec{r} - \vec{t}|^2 \rangle &= \left\langle \sum_i \left(\langle \sigma_i \rangle - s_i\right)^2 \right\rangle \\ &= \sum_i \mathrm{Var}\left(\langle \sigma_i \rangle\right) \\ &= \frac{1}{N} \sum_i \mathrm{Var}(\sigma_i) \\ &:= \epsilon^2. \end{aligned} \tag{C.3}$$

To simplify the discussion, let us assume that the vectors $\vec{r}$ and $\vec{t}$ share a common axis, allowing us to express their norms in relation to each other by

$$|\vec{t}| = |\vec{r}| \pm \epsilon. \tag{C.4}$$

Initially, consider the scenario where the target state ρ is a mixed state. In this case, we can express $|\vec{r}| = 1 - \alpha$, where α is sufficiently large such that $1 > \alpha \gg \epsilon$. Inserting this and Equation C.4 into Equation C.1, and performing a Taylor expansion around $\epsilon = 0$, we get

$$1 - F(\rho, \sigma) = \frac{\epsilon^2}{2(\alpha - \alpha^2)} + O(\epsilon^3). \tag{C.5}$$

Now consider the case where the target qubit is pure, $|\vec{r}| = 1$, the estimate is restricted to be of the form $|\vec{t}| = 1 - \epsilon$. Inserting this into Equation C.1 we end up with

$$1 - F(\rho, \sigma) = \frac{\epsilon}{2}. \tag{C.6}$$

Thus, we conclude that in the estimation of pure states or states close to the outer shell of the Bloch sphere, the scaling is expected to be square-root worse compared to mixed states, provided that the samples are statistically independent. Since we estimate a mean quantity, the statistical error will be $\epsilon = \sigma_{\text{mean}} = \frac{\sigma}{\sqrt{N}}$ for N samples. We therefore get

$$1 - F(\rho, \sigma) = \begin{cases} \frac{1}{2\sqrt{N}} & \text{for pure states ,} \\ \propto \frac{1}{N} & \text{for mixed states.} \end{cases} \tag{C.7}$$

C.2 SCALING BEHAVIOR OF RANK DEFICIENT ESTIMATORS

In this section, we derive the connecting steps between the results presented in Section 7.3 and the original work, Ref. [172], and the refined explanation found in Ref. [179]. Many of the steps were clarified by Ref. [75].

Our goal is to derive the conditions that must be satisfied by the generalized measurement $\mathcal{M} = \{M_\gamma\}$ so that the singular values σ_i of the *Fisher information matrix* (FIM) scale optimally. When the singular values are not vanishingly small, they are inversely proportional to the reconstruction accuracy, Equation 7.6, restated here for convenience,

$$\langle 1 - F \rangle = \sum_{i=2}^{\nu+1} \frac{1}{\sigma_i}. \tag{C.8}$$

Specifically, we want to find conditions on M_γ such that all singular values are non-vanishing and scale as $\sigma_i \propto N$, where N is the total number of measurements. We stress that Equation C.8 is only valid for the singular values that are non-vanishing. If there are vanishing singular values, they must be treated separately while the non-vanishing singular values follow Equation C.8. When the singular values are equal to zero, they generically acquire a scaling proportional to $1/\sqrt{N}$ because they are part of the boundary of the parameter space [179], as in Appendix C.1.

C.2.1 *Defining notation*

Starting out, we need to define the notation for our specific problem. The target states ρ are pure states with rank 1 in a d dimensional Hilbert space. The target state can be written in terms of an eigendecomposition,

$$\rho = \sum_{k=1}^{d} \lambda_k \left|\lambda_k\right\rangle \left\langle\lambda_k\right|, \tag{C.9}$$

where $\{|\lambda_k\rangle\}$ forms a basis in the d-dimensional Hilbert space. We still want to use full-rank estimators, meaning that the rank of the estimated state will be d, and therefore there are $d^2 - 1$ parameters that must be estimated, and $\nu = d^2 - 1$ in Equation C.8. We will purify this system by introducing an auxiliary Hilbert space of the same dimension, totaling $\mathcal{H}_d \otimes \mathcal{H}_d$.

This purification allows us to represent the density matrices as state vectors. In addition to the purification, we will embed the complex purified state vector and measurement operators into a space with all real entries, which means that we can write

$$\begin{aligned} \rho_{d\times d} &\overset{\text{Purification}}{\rightarrow} \sum_{k=1}^{d} \sqrt{\lambda_k}\,|k\rangle_d \otimes |\lambda_k\rangle_d = |\Psi\rangle_{d^2} \overset{\text{Real}}{\rightarrow} v = \begin{pmatrix} \mathcal{R}(|\Psi\rangle) \\ \mathcal{I}(|\Psi\rangle) \end{pmatrix}, \\ M_{\gamma,d\times d} &\overset{\text{Purification}}{\rightarrow} \mathbb{1}_{d\times d} \otimes M_{\gamma,d\times d} = A_\gamma \overset{\text{Real}}{\rightarrow} O_\gamma = \begin{pmatrix} \mathcal{R}(A_\gamma) & -\mathcal{I}(A_\gamma) \\ \mathcal{I}(A_\gamma) & \mathcal{R}(A_\gamma) \end{pmatrix}. \end{aligned} \tag{C.10}$$

Using this notation, we can rewrite the Born rule, Equation 2.38, as

$$p_\gamma = \mathrm{Tr}(M_\gamma \rho) = v^T O_\gamma v, \tag{C.11}$$

where written out explicitly,

$$|\Psi\rangle = \begin{pmatrix} \sqrt{\lambda_1}\,|\lambda_1\rangle \\ \vdots \\ \sqrt{\lambda_d}\,|\lambda_d\rangle \end{pmatrix} = \begin{pmatrix} \sqrt{\lambda_1}\,\langle 1|\lambda_1\rangle \\ \vdots \\ \sqrt{\lambda_1}\,\langle d|\lambda_1\rangle \\ \sqrt{\lambda_2}\,\langle 1|\lambda_2\rangle \\ \vdots \\ \sqrt{\lambda_d}\,\langle d|\lambda_d\rangle \end{pmatrix} \quad \text{and} \quad A_\gamma = \mathrm{diag}(M_\gamma). \tag{C.12}$$

Using this notation, we can compute the FIM.

C.2.2 *Fisher information matrix for a multinomial distribution*

Our measurement follows the multinomial distribution, and the likelihood functions can be written in the purified notation as[1]

$$\mathcal{L}(v; n_\gamma) \propto \Pi_\gamma (v^T O_\gamma v)^{n_\gamma}. \tag{C.13}$$

The score function is therefore (dropping n_γ as an argument)

$$s(v) = \ln(\mathcal{L}(v)) = \sum_\gamma n_\gamma \ln\left(v^T O_\gamma v\right). \tag{C.14}$$

We can compute the second formulation of the FIM in Equation 2.63, recalling that $p_\gamma = v^T O_\gamma v$,

$$\begin{aligned} \frac{\ln(\mathcal{L}(v))}{\partial v_i} &= \sum_\gamma \frac{n_\gamma}{p_\gamma} \frac{\partial p_\gamma}{\partial v_i}, \\ \frac{\partial^2 \ln(\mathcal{L}(v))}{\partial v_i \partial v_j} &= -\sum_\gamma \frac{n_\gamma}{p_\gamma^2} \frac{\partial p_\gamma}{\partial v_i} \frac{\partial p_\gamma}{\partial v_j} + \sum_\gamma \frac{n_\gamma}{p_\gamma} \frac{\partial^2 p_\gamma}{\partial v_i \partial v_j}, \end{aligned} \tag{C.15}$$

---

[1] To simplify notation, we limit ourselves to just one measurement setting, which leads us to omit the index α that was previously used in Section 7.3.1.

where

$$\frac{\partial p_\gamma}{\partial v_i} = 2\sum_a (O_\gamma)_{ia} v_a. \tag{C.16}$$

Combining everything, we get the overall result

$$(\mathcal{I}(v))_{ij} = \left\langle \sum_\gamma \frac{4n_\gamma}{p_\gamma^2}(O_\gamma v)_i(v^T O_\gamma)_j - \sum_\gamma \frac{n_\gamma}{p_\gamma}\frac{\partial^2 p_\gamma}{\partial v_i \partial v_j} \right\rangle. \tag{C.17}$$

We can compute the expectation value noting that for a multinomial distribution we have $\langle n_\gamma \rangle = Np_\gamma$, giving us

$$(\mathcal{I}(v))_{ij} = N\left(\sum_\gamma \frac{4}{p_\gamma}(O_\gamma v)_i(v^T O_\gamma)_j - \sum_\gamma \frac{\partial^2 p_\gamma}{\partial v_i \partial v_j} \right). \tag{C.18}$$

The last term vanishes because the probabilities have to be conserved $\sum_\gamma p_\gamma = 1$, which means that any derivative must vanish $\sum_\gamma \partial p_\gamma / \partial v_i = 0$. This leads us to the final version of the FIM,

$$\mathcal{I}(v) = N\sum_\gamma \frac{4}{p_\gamma} O_\gamma v v^T O_\gamma, \tag{C.19}$$

and in the complex form

$$\mathcal{I}(|\Psi\rangle) = N\sum_\gamma \frac{4}{p_\gamma} A_\gamma |\Psi\rangle\langle\Psi| A_\gamma. \tag{C.20}$$

C.2.3 *Rank-deficient target states*

With the FIMs in Equation C.20 and Equation C.19, it is clear that the singular values σ_i are generally proportional to N. However, this is not the case when the target state is rank deficient compared to the estimator rank. When talking about the rank of the FIM, we refer to the formulation in Equation C.19. The full FIM in Equation C.19 has rank $2d^2$. Half of these dimensions come from purification and are not physically relevant. The remaining ranks are related to the physical DoF of the estimated state resulting in $d^2 - 1$ relevant dimensions, subtracting one for normalization. In the estimation problem outlined in Section 7.1, the target states are pure and therefore have $2(d-1)$ DoF. The rank deficiency of the estimator is the difference between these two [179],

$$\text{rank deficiency} = d^2 - 1 - 2(d-1) \overset{d=2^n}{=} 4^n - 1 - 2(2^n - 1). \tag{C.21}$$

To better understand what happens, we will explicitly consider a single-qubit example. The target state can be written as

$$\rho = \lambda_1 |\lambda_1\rangle\langle\lambda_1| + \lambda_2 |\lambda_2\rangle\langle\lambda_2|, \tag{C.22}$$

where $\lambda_1 = 1$ and $\lambda_2 = 0$. This results in the terms of Equation C.20 being written as

$$A_\gamma |\Psi\rangle\langle\Psi| A_\gamma = \begin{pmatrix} (\mathbb{1}\otimes M_\gamma)|\lambda_1\rangle\langle\lambda_1|(\mathbb{1}\otimes M_\gamma) & 0 \\ 0 & 0 \end{pmatrix}, \tag{C.23}$$

for all γ in $\mathcal{I}(|\Psi\rangle)$. Therefore, due to the rank-deficient target state, the FIM has zero singular values when all $p_\gamma \neq 0$. As a result, Equation C.8 becomes invalid for vanishing singular values, and reconstruction infidelity is influenced by contributions that scale suboptimally as $1/\sqrt{N}$.

C.2.4 *Optimal adaptive measurement strategies*

The way adaptive strategies recover optimal asymptotic scaling is by increasing the non-zero singular values of $\mathcal{I}(|\Psi\rangle)$ by finding measurement operators M_γ that satisfy $\mathrm{Tr}(M_\gamma |\lambda_k\rangle\langle\lambda_k|) = 0$ for state vectors $|\lambda_k\rangle$ with non-zero eigenvalues, $\lambda_k \neq 0$. This effectively amplifies all zero singular values by a factor $1/p_\gamma$. To understand how this comes about, consider a measurement operator of the form $M_\gamma = \tau_\gamma |\psi_\gamma\rangle\langle\psi_\gamma|$, where τ_γ denotes the trace of the operator, $\mathrm{Tr}(M_\gamma) = \tau_\gamma$. Using this form, we can write the FIM starting from

$$A_\gamma |\Psi\rangle = \begin{pmatrix} \sqrt{\lambda_1} M_\gamma |\lambda_1\rangle \\ \sqrt{\lambda_2} M_\gamma |\lambda_2\rangle \end{pmatrix} = \tau_\gamma \begin{pmatrix} \sqrt{\lambda_1} \langle\psi_\gamma|\lambda_1\rangle |\psi_\gamma\rangle \\ \sqrt{\lambda_2} \langle\psi_\gamma|\lambda_2\rangle |\psi_\gamma\rangle \end{pmatrix} = \sqrt{\tau_\gamma} \begin{pmatrix} g_{1,\gamma} |\psi_\gamma\rangle \\ g_{2,\gamma} |\psi_\gamma\rangle \end{pmatrix}, \tag{C.24}$$

where we have defined new factors

$$g_{i,\gamma}^2 = g_{i,\gamma} g_{i,\gamma}^* = \tau_\gamma \lambda_i |\langle\psi_\gamma|\lambda_i\rangle|^2. \tag{C.25}$$

This allows us to rewrite the outcome probabilities

$$p_\gamma = \sum_i g_{i,\gamma}^2 = \langle\Psi| A_\gamma |\Psi\rangle, \tag{C.26}$$

and the components of the FIM

$$A_\gamma |\Psi\rangle\langle\Psi| A_\gamma = \tau_\gamma \begin{pmatrix} g_{1,\gamma}^2 |\psi_\gamma\rangle\langle\psi_\gamma| & g_{1,\gamma} g_{2,\gamma}^* |\psi_\gamma\rangle\langle\psi_\gamma| \\ g_{1,\gamma}^* g_{2,\gamma} |\psi_\gamma\rangle\langle\psi_\gamma| & g_{2,\gamma}^2 |\psi_\gamma\rangle\langle\psi_\gamma| \end{pmatrix}. \tag{C.27}$$

If a measurement operator M_γ were found such that

$$g_{1,\gamma}^2 \approx g_{2,\gamma}^2 := g_\gamma^2, \tag{C.28}$$

then

$$\frac{1}{p_\gamma} A_\gamma |\Psi\rangle\langle\Psi| A_\gamma = \frac{\tau_\gamma g_\gamma^2}{\sum_i g_\gamma^2} \begin{pmatrix} |\psi_\gamma\rangle\langle\psi_\gamma| & |\psi_\gamma\rangle\langle\psi_\gamma| \\ |\psi_\gamma\rangle\langle\psi_\gamma| & |\psi_\gamma\rangle\langle\psi_\gamma| \end{pmatrix} \propto \begin{pmatrix} |\psi_\gamma\rangle\langle\psi_\gamma| & |\psi_\gamma\rangle\langle\psi_\gamma| \\ |\psi_\gamma\rangle\langle\psi_\gamma| & |\psi_\gamma\rangle\langle\psi_\gamma| \end{pmatrix}. \tag{C.29}$$

This matrix introduces a new nonzero singular value to the FIM $\mathcal{I}(v)$. For a single-qubit, the rank deficiency (Equation C.21) is $4^n - 1 - 2(2^n - 1) \overset{n=1}{=} 1$, and a single M_γ that satisfies Equation C.28 is sufficient to recover the optimal asymptotic scaling.

What exactly does the condition in Equation C.28 tell us about the measurement operator? Written out explicitly, we have

$$\sqrt{\lambda_1} |\langle\psi_\gamma|\lambda_1\rangle|^2 = \sqrt{\lambda_2} |\langle\psi_\gamma|\lambda_2\rangle|^2, \tag{C.30}$$

and inserting the eigenvalues,

$$|\langle\psi_\gamma|\lambda_1\rangle|^2 = 0. \tag{C.31}$$

We therefore need to find a measurement operator that is orthogonal to the eigenvector corresponding to the non-zero eigenvalue of the target state. This condition extends naturally to the higher-dimensional cases. In general, the requirement for adaptive measurements to achieve optimal scaling is that the measurement operators must be orthogonal to all eigenvectors corresponding to non-zero eigenvalues of the target state. Specifically, the condition is

$$M_\gamma = \tau_\gamma |\psi_\gamma\rangle\langle\psi_\gamma| \text{ such that } \langle\psi_\gamma|\lambda_k\rangle = 0 \text{ for all } k \text{ with } \lambda_k \neq 0, \tag{C.32}$$

and the number of such measurement operators required is given by [179]

$$N_{\mathcal{M}} = \begin{cases} 4^n - 1 - 2(2^n - 1), & \text{for pure target states,} \\ 4^n - 1 - \left(R2^{n+1} - R^2 - 1\right), & \text{for target states of rank } R. \end{cases} \tag{C.33}$$

C.3 OPTIMAL ADAPTIVE MEASUREMENT STRATEGIES WITH NOISE

The condition described in Equation C.32 explains how using multiple measurement settings can recover optimal asymptotic infidelity scaling for rank-deficient sates. We can now investigate how readout noise influences the optimality of adaptive strategies. We will consider the depolarizing channel introduced in Equation 2.34, which is restated here for convenience,

$$\mathcal{E}_{\text{depol}}(A) = p\frac{\mathbb{1}}{2^n}\operatorname{Tr}(A) + (1-p)A. \tag{C.34}$$

Applied to the measurement operators of the form $M_\gamma = \tau_\gamma |\psi_\gamma\rangle\langle\psi_\gamma|$ yields

$$\tilde{M}_\gamma = \mathcal{E}_{\text{depol}}(M_\gamma) = p\frac{\mathbb{1}}{2^n}\tau_\gamma + (1-p)\tau_\gamma |\psi_\gamma\rangle\langle\psi_\gamma|. \tag{C.35}$$

Inserting Equation C.35 into the single-qubit example in Equation C.24, we have

$$A_\gamma|\Psi\rangle = \tau_\gamma \begin{pmatrix} \sqrt{\lambda_1}(\frac{p}{2}|\lambda_1\rangle + (1-p)\langle\psi_\gamma|\lambda_1\rangle|\psi_\gamma\rangle) \\ \sqrt{\lambda_2}(\frac{p}{2}|\lambda_2\rangle + (1-p)\langle\psi_\gamma|\lambda_2\rangle|\psi_\gamma\rangle) \end{pmatrix} \tag{C.36}$$

where we have used $\mathbb{1} = \sum_i |\lambda_i\rangle\langle\lambda_i|$. The probability in Equation C.26 becomes

$$p_\gamma = \langle\Psi|A_\gamma|\Psi\rangle = \tau_\gamma\left(\lambda_1\left(\frac{p}{2} + (1-p)|\langle\psi_\gamma|\lambda_1\rangle|^2\right) + \lambda_2\left(\frac{p}{2} + (1-p)|\langle\psi_\gamma|\lambda_2\rangle|^2\right)\right). \tag{C.37}$$

Since all terms are positive, we can set a lower bound on the probability

$$p_\gamma \geq \tau_\gamma\frac{p}{2}(\lambda_1 + \lambda_2) = \tau_\gamma\frac{p}{2}. \tag{C.38}$$

The minimum is reached when $\langle\psi_\gamma|\lambda_1\rangle = 0$ if $\lambda_2 = 0$, which matches the condition found in Equation C.32. Since p_γ does not approach zero, the zero rows and columns in Equation C.27 cannot be amplified if $\lambda_2 \to 0$.

More generally, the entries of the FIM with depolarized readout can be written as

$$(\mathcal{I}(|\Psi\rangle))_{ij} = N\sum_{\gamma}\frac{4}{p_\gamma}(A_\gamma|\Psi\rangle)_i(\langle\Psi|A_\gamma)_j = N\sum_{\gamma}\frac{4}{p_\gamma}\tau_\gamma^2\sqrt{\lambda_i\lambda_j}a_{ij,\gamma} \tag{C.39}$$

where $i, j \in \{1, \dots, 2^n\}$ and

$$a_{ij,\gamma} = \left(\frac{p}{2^n}|\lambda_i\rangle + (1-p)\langle\psi_\gamma|\lambda_i\rangle|\psi_\gamma\rangle\right)\left(\frac{p}{2^n}\langle\lambda_j| + (1-p)\langle\lambda_j|\psi_\gamma\rangle\langle\psi_\gamma|\right). \tag{C.40}$$

The probabilities are similarly lower bounded in the n-qubit case,

$$p_\gamma = \tau_\gamma\frac{p}{2^n} + \tau_\gamma(1-p)\sum_{i=1}^{2^n}\lambda_i|\langle\psi_\gamma|\lambda_i\rangle|^2 \geq \tau_\gamma\frac{p}{2^n}. \tag{C.41}$$

Hence, readout noise introduces a finite lower bound on all p_γ, preventing the prefactor $4/p_\gamma$ in Equation C.39 from “blowing up” to recover rows and columns associated with vanishing eigenvalues, regardless of the chosen measurement operator M_γ. Note that choosing a $\tau_\gamma \to 0$ will not help because $\frac{1}{p_\gamma}A_\gamma|\Psi\rangle\langle\Psi|A_\gamma \propto \tau_\gamma$.

BIBLIOGRAPHY

[1] M. Planck, ‘Über das Gesetz der Energieverteilung im Normalspectrum,’ Annalen der Physik **309**, 553–563 (1901).

[2] A. Einstein, ‘Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt,’ Annalen der Physik **322**, 132–148 (1905).

[3] N. Bohr, ‘On the constitution of atoms and molecules,’ The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **26**, 1–25 (1913).

[4] W. Gerlach and O. Stern, ‘Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld,’ Zeitschrift für Physik **9**, 349–352 (1922).

[5] A. H. Compton, ‘A quantum theory of the scattering of x-rays by light elements,’ Physical Review **21**, 483–502 (1923).

[6] E. Schrödinger, ‘An undulatory theory of the mechanics of atoms and molecules,’ Physical Review **28**, 1049–1070 (1926).

[7] W. Heisenberg, ‘Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen,’ Zeitschrift für Physik **33**, 879–893 (1925).

[8] A. Aspect, P. Grangier, and G. Roger, ‘Experimental tests of realistic local theories via Bell’s theorem,’ Physical Review Letters **47**, 460–463 (1981).

[9] A. Aspect, P. Grangier, and G. Roger, ‘Experimental realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: a new violation of Bell’s inequalities,’ Physical Review Letters **49**, 91–94 (1982).

[10] A. Aspect, J. Dalibard, and G. Roger, ‘Experimental test of Bell’s inequalities using time-varying analyzers,’ Physical Review Letters **49**, 1804–1807 (1982).

[11] P. Grangier, G. Roger, and A. Aspect, ‘Experimental evidence for a photon anticorrelation effect on a beam splitter: a new light on single-photon interferences,’ Europhysics Letters **1**, 173–179 (1986).

[12] A. Aspect, *Einstein and the quantum revolutions* (University of Chicago Press, Chicago, 2024),

[13] Y. I. Manin, ‘Computable and non-computable,’ Sovetskoe Radio, 128 (1980).

[14] P. Benioff, ‘The computer as a physical system: a microscopic quantum mechanical Hamiltonian model of computers as represented by Turing machines,’ Journal of Statistical Physics **22**, 563–591 (1980).

[15] R. P. Feynman, ‘Simulating physics with computers,’ International Journal of Theoretical Physics **21**, 467–488 (1982).

[16] D. Deutsch, ‘Quantum theory, the Church-Turing principle and the universal quantum computer,’ Proceedings of the Royal Society of London A **400**, 97–117 (1985).

[17] R. Blatt and C. F. Roos, ‘Quantum simulations with trapped ions,’ Nature Physics **8**, 277–284 (2012).

[18] C. Gross and I. Bloch, ‘Quantum simulations with ultracold atoms in optical lattices,’ Science **357**, 995–1001 (2017).

[19] M. Morgado and S. Whitlock, ‘Quantum simulation and computing with Rydberg-interacting qubits,’ AVS Quantum Science **3**, 023501 (2021).

[20] M. Saffman, ‘Quantum computing with atomic qubits and Rydberg interactions: progress and challenges,’ Journal of Physics B **49**, 202001 (2016).

[21] A. Browaeys and T. Lahaye, ‘Many-body physics with individually controlled Rydberg atoms,’ Nature Physics **16**, 132–142 (2020).

[22] I. M. Georgescu, S. Ashhab, and F. Nori, ‘Quantum simulation,’ Reviews of Modern Physics **86**, 153–185 (2014).

[23] E. Altman et al., ‘Quantum simulators: architectures and opportunities,’ PRX Quantum **2**, 017003 (2021).

[24] A. I. Ekimov and A. A. Onushchenko, ‘Quantum size effect in three-dimensional microscopic semiconductor crystals,’ JETP Letters **118**, S15–S17 (2023).

[25] A. Ekimov, A. Efros, and A. Onushchenko, ‘Quantum size effect in semiconductor microcrystals,’ Solid State Communications **56**, 921–924 (1985).

[26] J. Preskill, ‘Quantum computing in the NISQ era and beyond,’ Quantum **2**, 79 (2018).

[27] S. Ebadi et al., ‘Quantum optimization of maximum independent set using Rydberg atom arrays,’ Science **376**, 1209–1215 (2022).

[28] R. Acharya et al., ‘Quantum error correction below the surface code threshold,’ Nature **638**, 920–926 (2024).

[29] V. V. Sivak, A. Eickbusch, B. Royer, S. Singh, I. Tsioutsios, S. Ganjam, A. Miano, B. L. Brock, A. Z. Ding, L. Frunzio, S. M. Girvin, R. J. Schoelkopf, and M. H. Devoret, ‘Real-time quantum error correction beyond break-even,’ Nature **616**, 50–55 (2023).

[30] P. Wang, C.-Y. Luan, M. Qiao, M. Um, J. Zhang, Y. Wang, X. Yuan, M. Gu, J. Zhang, and K. Kim, ‘Single ion qubit with estimated coherence time exceeding one hour,’ Nature Communications **12**, 233 (2021).

[31] F.-M. L. Régent, *Awesome quantum computing experiments: benchmarking experimental progress towards fault-tolerant quantum computation*, 2025, arXiv:`2507.03678 [quant-ph]`.

[32] Y.-Y. Jiang, C. Deng, H. Fan, B.-Y. Li, L. Sun, X.-S. Tan, W. Wang, G.-M. Xue, F. Yan, H.-F. Yu, Y.-S. Zhang, Y.-R. Zhang, and C.-L. Zou, ‘Advancements in superconducting quantum computing,’ National Science Review **12**, 4677 (2025).

[33] Z. Zimborás et al., *Myths around quantum computation before full fault tolerance: what no-go theorems rule out and what they don't*, 2025, arXiv:`2501.05694 [quant-ph]`.

[34] J. Eisert and J. Preskill, *Mind the gaps: the fraught road to quantum advantage*, 2025, arXiv:`2510.19928 [quant-ph]`.

[35] Z. Cai, R. Babbush, S. C. Benjamin, S. Endo, W. J. Huggins, Y. Li, J. R. McClean, and T. E. O' Brien, ‘Quantum error mitigation,’ Reviews of Modern Physics **95**, 045005 (2023).

[36] J. Tuziemski, F. B. Maciejewski, J. Majsak, O. Słowik, M. Kotowski, K. Kowalczyk-Murynka, P. Podziemski, and M. Oszmaniec, *Efficient reconstruction, benchmarking and validation of cross-talk models in readout noise in near-term quantum devices*, 2023, arXiv:`2311.10661 [quant-ph]`.

[37] Z. Zhou, R. Sitler, Y. Oda, K. Schultz, and G. Quiroz, ‘Quantum crosstalk robust quantum control,’ Physical Review Letters **131**, 210802 (2023).

[38] M. R. Geller and M. Sun, ‘Toward efficient correction of multiqubit measurement errors: pair correlation method,’ Quantum Science and Technology **6**, 025009 (2021).

[39] M. Sarovar, T. Proctor, K. Rudinger, K. Young, E. Nielsen, and R. Blume-Kohout, ‘Detecting crosstalk errors in quantum information processors,’ Quantum **4**, 321 (2020).

[40] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information* (Cambridge University Press, June 2012),

[41] J. Preskill, ‘Lecture notes for physics 229: quantum information and computation,’ California institute of technology **16**, 1–8 (1998).

[42] R. de Wolf, *Quantum computing: lecture notes*, 2023, arXiv:`1907.09415 [quant-ph]`.

[43] J. J. Sakurai and J. Napolitano, *Modern quantum mechanics* (Cambridge University Press, Sept. 2017),

[44] A. N. Jordan and I. A. Siddiqi, *Quantum measurement: theory and practice* (Cambridge University Press, 2024).

[45] R. Blume-Kohout, ‘Optimal, reliable estimation of quantum states,’ New Journal of Physics **12**, 043034 (2010).

[46] S. Bravyi and A. Kitaev, ‘Universal quantum computation with ideal Clifford gates and noisy ancillas,’ Physical Review A **71**, 022316 (2005).

[47] V. Veitch, S. A. Hamed Mousavian, D. Gottesman, and J. Emerson, ‘The resource theory of stabilizer quantum computation,’ New Journal of Physics **16**, 013009 (2014).

[48] C. H. Bennett and S. J. Wiesner, ‘Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states,’ Physical Review Letters **69**, 2881–2884 (1992).

[49] C. H. Bennett and G. Brassard, ‘Quantum cryptography: public key distribution and coin tossing,’ Theoretical Computer Science **560**, 7–11 (2014).

[50] A. K. Ekert, ‘Quantum cryptography based on Bell's theorem,’ Physical Review Letters **67**, 661–663 (1991).

[51] A. Y. Kitaev, ‘Quantum computations: algorithms and error correction,’ Russian Mathematical Surveys **52**, 1191–1249 (1997).

[52] A. W. Harrow, B. Recht, and I. L. Chuang, ‘Efficient discrete approximations of quantum gates,’ Journal of Mathematical Physics **43**, 4445–4451 (2002).

[53] M. J. Bremner, C. M. Dawson, J. L. Dodd, A. Gilchrist, A. W. Harrow, D. Mortimer, M. A. Nielsen, and T. J. Osborne, ‘Practical scheme for quantum computation with any two-qubit entangling gate,’ Physical Review Letters **89**, 247902 (2002).

[54] *Native gates and operations | IBM Quantum Documentation*, [Accessed 22.04.2025], `https://quantum.cloud.ibm.com/docs/en/guides/native-gates#native-gates-on-platform` (visited on 04/22/2025).

[55] Y. Kim, A. Morvan, L. B. Nguyen, R. K. Naik, C. Jünger, L. Chen, J. M. Kreikebaum, D. I. Santiago, and I. Siddiqi, ‘High-fidelity three-qubit iToffoli gate for fixed-frequency superconducting qubits,’ Nature Physics **18**, 783–788 (2022).

[56] C. W. Warren et al., ‘Extensive characterization and implementation of a family of three-qubit gates at the coherence limit,’ npj Quantum Information **9**, 44 (2023).

[57] T. Itoko, M. Malekakhlagh, N. Kanazawa, and M. Takita, ‘Three-qubit parity gate via simultaneous cross-resonance drives,’ Physical Review Applied **21**, 034018 (2024).

[58] J. von Neumann, *Mathematical foundations of quantum mechanics: new edition*, edited by R. T. Beyer and N. A. Wheeler (Princeton University Press, 2018),

[59] T. E. Phipps and J. B. Taylor, ‘The magnetic moment of the hydrogen atom,’ Physical Review **29**, 309–320 (1927).

[60] R. Frisch and E. Segrè, ‘Über die Einstellung der Richtungsquantelung. II,’ Zeitschrift für Physik **80**, 610–616 (1933).

[61] G. Lüders, ‘Über die Zustandsänderung durch den Meßprozeß,’ Annalen der Physik **443**, 322–328 (1950).

[62] K. A. Kirkpatrick, *Translation of Lueders' "Uber die Zustandsanderung durch den Messprozess"*, 2006, arXiv:`quant-ph/0403007 [quant-ph]`.

[63] A. Peres, ‘Unperformed experiments have no results,’ American Journal of Physics **46**, 745–747 (1978).

[64] S. W. Hawking, *The nature of space and time*, 1994, arXiv:`hep-th/9409195 [hep-th]`.

[65] W. F. Stinespring, ‘Positive functions on c*-algebras,’ Proceedings of the American Mathematical Society **6**, 211–216 (1955).

[66] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan, 'Completely positive dynamical semigroups of n-level systems,' Journal of Mathematical Physics **17**, 821–825 (1976).

[67] A. S. Holevo and V. Giovannetti, 'Quantum channels and their entropic characteristics,' Reports on Progress in Physics **75**, 046001 (2012).

[68] E. Magesan, 'Depolarizing behavior of quantum channels in higher dimensions,' Quantum Information and Computation **11**, 466–484 (2011).

[69] G. Zauner, 'Quantum designs: foundations of a noncommutative design theory,' International Journal of Quantum Information **09**, 445–507 (2011).

[70] M. Appleby, S. T. Flammia, and G. S. Kopp, *A constructive approach to Zauner's conjecture via the Stark conjectures*, 2025, arXiv:`2501.03970 [math.NT]`.

[71] M. B. Mensky, 'Quantum restrictions for continuous observation of an oscillator,' Physical Review D **20**, 384–387 (1979).

[72] V. P. Belavkin, 'Quantum filtering of Markov signals with white quantum noise,' Quantum Communications and Measurement, 381–391 (1995).

[73] E. Toninelli, B. Ndagano, A. Vallés, B. Sephton, I. Nape, A. Ambrosio, F. Capasso, M. J. Padgett, and A. Forbes, 'Concepts in quantum state tomography and classical implementation with intense light: a tutorial,' Advances in Optics and Photonics **11**, 67–134 (2019).

[74] A. I. Lvovsky and M. G. Raymer, 'Continuous-variable optical quantum-state tomography,' Reviews of Modern Physics **81**, 299–332 (2009).

[75] E. Stock, 'Adaptive quantum state tomography,' Bachelor thesis, **`https://github.com/3ti3nn3/tomography`** (2023).

[76] Z. Hradil, 'Quantum-state estimation,' Physical Review A **55**, R1561–R1564 (1997).

[77] A.-L. Cholesky, 'Note sur une méthode de résolution des équations normales provenant de l'application de la méthode des moindres carrés à un système d'équations linéaires en nombre inférieur à celui des inconnues (procédé du commandant cholesky),' Bulletin Géodésique **2**, 67–77 (1924).

[78] J. Shang, Z. Zhang, and H. K. Ng, 'Superfast maximum-likelihood reconstruction for quantum tomography,' Physical Review A **95**, 062336 (2017).

[79] Z. Hou, H.-S. Zhong, Y. Tian, D. Dong, B. Qi, L. Li, Y. Wang, F. Nori, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, 'Full reconstruction of a 14-qubit state within four hours,' New Journal of Physics **18**, 083036 (2016).

[80] A. I. Lvovsky, 'Iterative maximum-likelihood reconstruction in quantum homodyne tomography,' Journal of Optics B: Quantum and Semiclassical Optics **6**, S556 (2004).

[81] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, 'Equation of state calculations by fast computing machines,' The Journal of Chemical Physics **21**, 1087–1092 (1953).

[82] A. Arrasmith, A. Patterson, A. Boughton, and M. Paini, *Development and demonstration of an efficient readout error mitigation technique for use in NISQ algorithms*, 2023, arXiv:`2303.17741 [quant-ph]`.

[83] G. Torlai and R. G. Melko, 'Machine-learning quantum states in the NISQ era,' Annual Review of Condensed Matter Physics **11**, 325–344 (2020).

[84] T. Schmale, M. Reh, and M. Gärttner, 'Efficient quantum state tomography with convolutional neural networks,' npj Quantum Information **8**, 115 (2022).

[85] M. Neugebauer, L. Fischer, A. Jäger, S. Czischek, S. Jochim, M. Weidemüller, and M. Gärttner, 'Neural-network quantum state tomography in a two-qubit experiment,' Physical Review A **102**, 042604 (2020).

[86] M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, and Y.-K. Liu, 'Efficient quantum state tomography,' Nature Communications **1**, 149 (2010).

[87] T. Baumgratz, D. Gross, M. Cramer, and M. B. Plenio, 'Scalable reconstruction of density matrices,' Physical Review Letters **111**, 020401 (2013).

[88] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, 'Quantum state tomography via compressed sensing,' Physical Review Letters **105**, 150401 (2010).

[89] S. Aaronson, 'Shadow tomography of quantum states,' in Proceedings of the 50th annual acm sigact symposium on theory of computing, STOC'18 (June 2018),

[90] H.-Y. Huang, R. Kueng, and J. Preskill, 'Predicting many properties of a quantum system from very few measurements,' Nature Physics **16**, 1050–1057 (2020).

[91] J. Cotler and F. Wilczek, 'Quantum overlapping tomography,' Physical Review Letters **124**, 100401 (2020).

[92] G. García-Pérez, M. A. C. Rossi, B. Sokolov, E.-M. Borrelli, and S. Maniscalco, 'Pairwise tomography networks for many-body quantum systems,' Physical Review Research **2**, 023393 (2020).

[93] X. Bonet-Monroig, R. Babbush, and T. E. O'Brien, 'Nearly optimal measurement scheduling for partial tomography of quantum states,' Physical Review X **10**, 031064 (2020).

[94] R. A. Walker II and C. J. Colbourn, 'Perfect hash families: constructions and existence,' Journal of Mathematical Cryptology **1**, (2007).

[95] R. Dougherty, *Hash families and applications to t-restrictions* (Arizona State University, 2019).

[96] J. Torres-Jimenez and I. Izquierdo-Marquez, 'Survey of covering arrays,' IEEE Xplore, 20–27 (2013).

[97] C. J. Colbourn, 'Combinatorial aspects of covering arrays,' Le Matematiche **59**, 125–172 (2004).

[98] K. Hansenne, R. Qu, L. T. Weinbrenner, C. de Gois, H. Wang, Y. Ming, Z. Yang, P. Horodecki, W. Gao, and O. Gühne, 'Optimal overlapping tomography,' Physical Review Letters **135**, 060801 (2025).

[99] J. S. Lundeen, A. Feito, H. Coldenstrodt-Ronge, K. L. Pregnell, C. Silberhorn, T. C. Ralph, J. Eisert, M. B. Plenio, and I. A. Walmsley, 'Tomography of quantum detectors,' Nature Physics **5**, 27–30 (2008).

[100] A. Feito, J. S. Lundeen, H. Coldenstrodt-Ronge, J. Eisert, M. B. Plenio, and I. A. Walmsley, 'Measuring measurement: theory and practice,' New Journal of Physics **11**, 093038 (2009).

[101] J. Fiurášek, 'Maximum-likelihood estimation of quantum measurement,' Physical Review A **64**, 024102 (2001).

[102] W. Lewin, *8.01 physics I: classical mechanics, fall 1999*, `http://ocw.mit.edu/courses/physics/8-01-physics-i-classical-mechanics-fall-1999`, 1999.

[103] C. A. Fuchs and C. M. Caves, 'Ensemble-dependent bounds for accessible information in quantum mechanics,' Physical Review Letters **73**, 3047–3050 (1994).

[104] H. Cramér, *Mathematical methods of statistics*, Vol. 9 (Princeton university press, 1999).

[105] C. R. Rao et al., 'Information and the accuracy attainable in the estimation of statistical parameters,' Bulletin of Calcutta Mathematical Society **37**, 81–91 (1945).

[106] M. Fréchet and M. Frechet, 'Sur l' extension de certaines evaluations statistiques au cas de petits echantillons,' Review of the International Statistical Institute **11**, 182 (1943).

[107] A. C. Aitken and H. Silverstone, 'On the estimation of statistical parameters,' Proceedings of the Royal Society of Edinburgh. Section A. Mathematical and Physical Sciences **61**, 186–194 (1942).

[108] F. Jelezko, T. Gaebel, I. Popa, A. Gruber, and J. Wrachtrup, 'Observation of coherent oscillations in a single electron spin,' Physical Review Letters **92**, 076401 (2004).

[109] M. V. G. Dutt, L. Childress, L. Jiang, E. Togan, J. Maze, F. Jelezko, A. S. Zibrov, P. R. Hemmer, and M. D. Lukin, 'Quantum register based on individual electronic and nuclear spin qubits in diamond,' Science **316**, 1312–1316 (2007).

[110] E. Knill, R. Laflamme, and G. J. Milburn, 'A scheme for efficient quantum computation with linear optics,' Nature **409**, 46–52 (2001).

[111] J. L. O'Brien, G. J. Pryde, A. G. White, T. C. Ralph, and D. Branning, 'Demonstration of an all-optical quantum controlled-not gate,' Nature **426**, 264–267 (2003).

[112] D. Jaksch, J. I. Cirac, P. Zoller, S. L. Rolston, R. Côté, and M. D. Lukin, 'Fast quantum gates for neutral atoms,' Physical Review Letters **85**, 2208–2211 (2000).

[113] J. I. Cirac and P. Zoller, 'Quantum computations with cold trapped ions,' Physical Review Letters **74**, 4091–4094 (1995).

[114] C. Monroe, D. M. Meekhof, B. E. King, W. M. Itano, and D. J. Wineland, 'Demonstration of a fundamental quantum logic gate,' Physical Review Letters **75**, 4714–4717 (1995).

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

[116] D. Vion, A. Aassime, A. Cottet, P. Joyez, H. Pothier, C. Urbina, D. Esteve, and M. H. Devoret, 'Manipulating the quantum state of an electrical circuit,' Science **296**, 886–889 (2002).

[117] A. Blais, R.-S. Huang, A. Wallraff, S. M. Girvin, and R. J. Schoelkopf, 'Cavity quantum electrodynamics for superconducting electrical circuits: an architecture for quantum computation,' Physical Review A **69**, 062320 (2004).

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

[119] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, 'A quantum engineer's guide to superconducting qubits,' Applied Physics Reviews **6**, 021318 (2019).

[120] B. Josephson, 'Possible new effects in superconductive tunnelling,' Physics Letters **1**, 251–253 (1962).

[121] B. D. Josephson, 'The discovery of tunnelling supercurrents,' Reviews of Modern Physics **46**, 251–254 (1974).

[122] S. M. Meißner, A. Seiler, J. Lisenfeld, A. V. Ustinov, and G. Weiss, 'Probing individual tunneling fluctuators with coherently controlled tunneling systems,' Physical Review B **97**, 180505 (2018).

[123] S. Simbierowicz, M. Borrelli, V. Monarkha, V. Nuutinen, and R. E. Lake, 'Inherent thermal-noise problem in addressing qubits,' PRX Quantum **5**, 030302 (2024).

[124] W. K. Wootters and W. H. Zurek, 'A single quantum cannot be cloned,' Nature **299**, 802–803 (1982).

[125] D. Dieks, 'Communication by EPR devices,' Physics Letters A **92**, 271–272 (1982).

[126] J. L. Park, 'The concept of transition in quantum mechanics,' Foundations of Physics **1**, 23–33 (1970).

[127] P. W. Shor, 'Scheme for reducing decoherence in quantum computer memory,' Physical Review A **52**, R2493–R2496 (1995).

[128] J. Roffe, 'Quantum error correction: an introductory guide,' Contemporary Physics **60**, 226–245 (2019).

[129] A. Kitaev, 'Fault-tolerant quantum computation by anyons,' Annals of Physics **303**, 2–30 (2003).

[130] S. B. Bravyi and A. Y. Kitaev, *Quantum codes on a lattice with boundary*, 1998, arXiv:`quant-ph/9811052 [quant-ph]`.

[131] A. M. Stephens, 'Fault-tolerant thresholds for quantum error correction with the surface code,' Physical Review A **89**, 022321 (2014).

[132] H. Ali, J. Marques, O. Crawford, J. Majaniemi, M. Serra-Peralta, D. Byfield, B. Varbanov, B. M. Terhal, L. DiCarlo, and E. T. Campbell, 'Reducing the error rate of a superconducting logical qubit using analog readout information,' Physical Review Applied **22**, 044031 (2024).

[133] R. Acharya et al., 'Suppressing quantum errors by scaling a surface code logical qubit,' Nature **614**, 676–681 (2023).

[134] Google Research, *Making quantum error correction work*, [Accessed 08.05.2025], `https://research.google/blog/making-quantum-error-correction-work/`.

[135] K. Kraus, *States, effects, and operations: fundamental notions of quantum theory*, Vol. 190, Lecture Notes in Physics, See Eq. (2.25), pp. 27–28 for the adjoint (Heisenberg-picture) definition. (Springer-Verlag, Berlin, Heidelberg, 1983),

[136] J. Preskill, *Lecture notes for physics 219: quantum computation*, `https://www.preskill.caltech.edu/notes.html`, See Section 3.2.3, Eqs. (3.47)–(3.49) for the Heisenberg-picture formulation of quantum channels., 2015.

[137] S. T. Merkel, J. M. Gambetta, J. A. Smolin, S. Poletto, A. D. Córcoles, B. R. Johnson, C. A. Ryan, and M. Steffen, 'Self-consistent quantum process tomography,' Physical Review A **87**, 062119 (2013).

[138] B. I. Bantysh and Y. I. Bogdanov, 'Quantum tomography of noisy ion-based qudits,' Laser Physics Letters **18**, 015203 (2020).

[139] J. Lin, J. J. Wallman, I. Hincks, and R. Laflamme, 'Independent state and measurement characterization for quantum computers,' Physical Review Research **3**, 033285 (2021).

[140] K. Q. Tuan, H. Q. Nguyen, and L. B. Ho, 'Direct state measurements under state-preparation-and-measurement errors,' Quantum Information Processing **20**, 197 (2021).

[141] A. Jayakumar, S. Chessa, C. Coffrin, A. Y. Lokhov, M. Vuffray, and S. Misra, 'Universal framework for simultaneous tomography of quantum states and SPAM noise,' Quantum **8**, 1426 (2024).

[142] H. Yu and T.-C. Wei, 'Efficient separate quantification of state preparation errors and measurement errors on quantum computers and their mitigation,' Quantum **9**, 1724 (2025).

[143] F. B. Maciejewski, Z. Zimborás, and M. Oszmaniec, 'Mitigation of readout noise in near-term quantum devices by classical post-processing based on detector tomography,' Quantum **4**, 257 (2020).

[144] S. Bravyi, S. Sheldon, A. Kandala, D. C. Mckay, and J. M. Gambetta, 'Mitigating measurement errors in multiqubit experiments,' Physical Review A **103**, 042605 (2021).

[145] Y. Chen, M. Farahzad, S. Yoo, and T.-C. Wei, 'Detector tomography on IBM quantum computers and mitigation of an imperfect measurement,' Physical Review A **100**, 052315 (2019).

[146] B. Nachman, M. Urbanek, W. A. de Jong, and C. W. Bauer, 'Unfolding quantum computer readout noise,' npj Quantum Information **6**, 84 (2020).

[147] B. Pokharel, S. Srinivasan, G. Quiroz, and B. Boots, 'Scalable measurement error mitigation via iterative bayesian unfolding,' Physical Review Research **6**, 013187 (2024).

[148] M. R. Geller, 'Rigorous measurement error correction,' Quantum Science and Technology **5**, 03LT01 (2020).

[149] B. Yang, R. Raymond, and S. Uno, 'Efficient quantum readout-error mitigation for sparse measurement outcomes of near-term quantum devices,' Physical Review A **106**, 012423 (2022).

[150] R. Hicks, B. Kobrin, C. W. Bauer, and B. Nachman, 'Active readout-error mitigation,' Physical Review A **105**, 012419 (2022).

[151] C. Ding, X.-Y. Xu, Y.-F. Niu, S. Zhang, W.-S. Bao, and H.-L. Huang, 'Noise-resistant quantum state compression readout,' Science China Physics, Mechanics & Astronomy **66**, 230311 (2023).

[152] H.-L. Huang and C. Ding, 'Quantum state compression shadow,' The European Physical Journal Special Topics, (2025).

[153] A. Javadi-Abhari, M. Treinish, K. Krsulich, C. J. Wood, J. Lishman, J. Gacon, S. Martiel, P. D. Nation, L. S. Bishop, A. W. Cross, B. R. Johnson, and J. M. Gambetta, *Quantum computing with Qiskit*, 2024, arXiv:`2405.08810 [quant-ph]`.

[154] Cirq Developers, 'Cirq,' Zenoob, (2025).

[155] V. Bergholm et al., *Pennylane: automatic differentiation of hybrid quantum-classical computations*, 2022, arXiv:`1811.04968 [quant-ph]`.

[156] Nvidia, *Cuda Quantum*, [Accessed 08.05.2025], `https://nvidia.github.io/cuda-quantum/latest/index.html`.

[157] Amazon Web Services, *Amazon Braket*, [Accessed 08.05.2025], `https://aws.amazon.com/braket/`.

[158] R. LaRose et al., 'Mitiq: a software package for error mitigation on noisy quantum computers,' Quantum **6**, 774 (2022).

[159] J. Heinsoo, C. K. Andersen, A. Remm, S. Krinner, T. Walter, Y. Salathé, S. Gasparinetti, J.-C. Besse, A. Potočnik, A. Wallraff, and C. Eichler, 'Rapid high-fidelity multiplexed readout of superconducting qubits,' Physical Review Applied **10**, 034040 (2018).

[160] B. Nachman and M. R. Geller, *Categorizing readout error correlations on near term quantum computers*, 2021, arXiv:`2104.04607 [quant-ph]`.

[161] F. B. Maciejewski, F. Baccari, Z. Zimborás, and M. Oszmaniec, 'Modeling and mitigation of cross-talk effects in readout noise with applications to the quantum approximate optimization algorithm,' Quantum **5**, 464 (2021).

[162] S. Chen, W. Yu, P. Zeng, and S. T. Flammia, 'Robust shadow estimation,' PRX Quantum **2**, 030348 (2021).

[163] H.-Y. Hu, A. Gu, S. Majumder, H. Ren, Y. Zhang, D. S. Wang, Y.-Z. You, Z. Minev, S. F. Yelin, and A. Seif, 'Demonstration of robust and efficient quantum property learning with shallow shadows,' Nature Communications **16**, 2943 (2025).

[164] E. Onorati, J. Kitzinger, J. Helsen, M. Ioannou, A. H. Werner, I. Roth, and J. Eisert, *Noise-mitigated randomized measurements and self-calibrating shadow estimation*, 2024, arXiv:`2403.04751 [quant-ph]`.

[165] E. Nielsen, J. K. Gamble, K. Rudinger, T. Scholten, K. Young, and R. Blume-Kohout, 'Gate set tomography,' Quantum **5**, 557 (2021).

[166] R. Blume-Kohout, J. K. Gamble, E. Nielsen, J. Mizrahi, J. D. Sterk, and P. Maunz, *Robust, self-consistent, closed-form tomography of quantum logic gates on a trapped ion qubit*, 2013, arXiv:`1310.4492 [quant-ph]`.

[167] A. Keesling, A. Omran, H. Levine, H. Bernien, H. Pichler, S. Choi, R. Samajdar, S. Schwartz, P. Silvi, S. Sachdev, P. Zoller, M. Endres, M. Greiner, V. Vuletić, and M. D. Lukin, 'Quantum Kibble – Zurek mechanism and critical dynamics on a programmable Rydberg simulator,' Nature **568**, 207–211 (2019).

[168] J. Zhang, S. H. Cantú, F. Liu, A. Bylinskii, B. Braverman, F. Huber, J. Amato-Grill, A. Lukin, N. Gemelke, A. Keesling, S.-T. Wang, Y. Meurice, and S.-W. Tsai, 'Probing quantum floating phases in Rydberg atom arrays,' Nature Communications **16**, 712 (2025).

[169] A. Schuckert, O. Katz, L. Feng, E. Crane, A. De, M. Hafezi, A. V. Gorshkov, and C. Monroe, 'Observation of a finite-energy phase transition in a one-dimensional quantum simulator,' Nature Physics **21**, 374–379 (2025).

[170] P. Richerme, Z.-X. Gong, A. Lee, C. Senko, J. Smith, M. Foss-Feig, S. Michalakis, A. V. Gorshkov, and C. Monroe, 'Non-local propagation of correlations in quantum systems with long-range interactions,' Nature **511**, 198–201 (2014).

[171] P. Jurcevic, B. P. Lanyon, P. Hauke, C. Hempel, P. Zoller, R. Blatt, and C. F. Roos, 'Quasiparticle engineering and entanglement propagation in a quantum many-body system,' Nature **511**, 202–205 (2014).

[172] E. Bagan, M. A. Ballester, R. D. Gill, A. Monras, and R. Muñoz-Tapia, 'Optimal full estimation of qubit mixed states,' Physical Review A **73**, 032301 (2006).

[173] D. H. Mahler, L. A. Rozema, A. Darabi, C. Ferrie, R. Blume-Kohout, and A. M. Steinberg, 'Adaptive quantum state tomography improves accuracy quadratically,' Physical Review Letters **111**, 183601 (2013).

[174] Y. I. Bogdanov, 'Unified statistical method for reconstructing quantum states by purification,' Journal of Experimental and Theoretical Physics **108**, 928–935 (2009).

[175] R. D. Gill and S. Massar, 'State estimation for large ensembles,' Physical Review A **61**, 042312 (2000).

[176] S. Massar and S. Popescu, 'Optimal extraction of information from finite quantum ensembles,' Physical Review Letters **74**, 1259–1263 (1995).

[177] A. Peres and W. K. Wootters, 'Optimal detection of quantum information,' Physical Review Letters **66**, 1119–1122 (1991).

[178] E. Bagan, A. Monras, and R. Muñoz-Tapia, 'Comprehensive analysis of quantum pure-state estimation for two-level systems,' Physical Review A **71**, 062318 (2005).

[179] G. I. Struchalin, E. V. Kovlakov, S. S. Straupe, and S. P. Kulik, 'Adaptive quantum tomography of high-dimensional bipartite systems,' Physical Review A **98**, 032330 (2018).

[180] G. I. Struchalin, I. A. Pogorelov, S. S. Straupe, K. S. Kravtsov, I. V. Radchenko, and S. P. Kulik, 'Experimental adaptive quantum tomography of two-qubit states,' Physical Review A **93**, 012103 (2016).

[181] K. S. Kravtsov, S. S. Straupe, I. V. Radchenko, N. M. T. Houlsby, F. Huszár, and S. P. Kulik, 'Experimental adaptive Bayesian tomography,' Physical Review A **87**, 062122 (2013).

[182] B. Qi, Z. Hou, Y. Wang, D. Dong, H.-S. Zhong, L. Li, G.-Y. Xiang, H. M. Wiseman, C.-F. Li, and G.-C. Guo, 'Adaptive quantum state tomography via linear regression estimation: theory and two-qubit experiment,' npj Quantum Information **3**, 19 (2017).

[183] H. Hwang, J. Choi, and E. Kim, 'Adaptive quantum tomography in an indistinct measurement system with superconducting circuits,' Physical Review Applied **20**, 064007 (2023).

[184] S. S. Straupe, 'Adaptive quantum tomography,' JETP Letters **104**, 510–522 (2016).

[185] F. Huszár and N. M. T. Houlsby, 'Adaptive Bayesian quantum tomography,' Physical Review A **85**, 052120 (2012).

[186] M. K. Patra, 'Quantum state determination: estimates for information gain and some exact calculations,' Journal of Physics A: Mathematical and Theoretical **40**, 10887–10902 (2007).

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

[188] Y. Quek, S. Fort, and H. K. Ng, 'Adaptive quantum state tomography with neural networks,' npj Quantum Information **7**, 105 (2021).

[189] Y. I. Bogdanov, B. I. Bantysh, N. A. Bogdanova, A. B. Kvasnyy, and V. F. Lukichev, 'Quantum states tomography with noisy measurement channels,' International Conference on Micro- and Nano-Electronics 2016 **10224**, edited by V. F. Lukichev and K. V. Rudenko, 102242O (2016).

[190] Y. I. Bogdanov, N. A. Bogdanova, Y. A. Kuznetsov, K. B. Koksharov, and V. F. Lukichev, 'Precise tomography of qudits,' Russian Microelectronics **52**, 135–143 (2023).

[191] V. N. Ivanova-Rohling, N. Rohling, and G. Burkard, 'Optimal quantum state tomography with noisy gates,' EPJ Quantum Technology **10**, 25 (2023).

[192] B. Hu, J. Sinclair, E. Bytyqi, M. Chong, A. Rudelis, J. Ramette, Z. Vendeiro, and V. Vuleti ć, 'Site-selective cavity readout and classical error correction of a 5-bit atomic register,' Physical Review Letters **134**, 120801 (2025).

[193] B. Grinkemeyer, E. Guardado-Sanchez, I. Dimitrova, D. Shchepanovich, G. E. Mandopoulou, J. Borregaard, V. Vuletić, and M. D. Lukin, 'Error-detected quantum operations with neutral atoms mediated by an optical cavity,' Science **387**, 1301–1305 (2025).

[194] P. Boross, D. Svastits, G. Egri, and A. Pályi, 'Multi-shot readout error benchmark of the nitrogen-vacancy center's electronic qubit,' Journal of Applied Physics **137**, 214401 (2025).

[195] Y. Sunada, K. Yuki, Z. Wang, T. Miyamura, J. Ilves, K. Matsuura, P. A. Spring, S. Tamate, S. Kono, and Y. Nakamura, 'Photon-noise-tolerant dispersive readout of a superconducting qubit using a nonlinear Purcell filter,' PRX Quantum **5**, 010307 (2024).

[196] Á. Márton and J. K. Asbóth, 'Coherent errors and readout errors in the surface code,' Quantum **7**, 1116 (2023).

[197] Y. Zheng and K. Kanno, 'Minimizing readout-induced noise for early fault-tolerant quantum computers,' Physical Review Research **6**, 023129 (2024).

[198] M. Jerger, S. Poletto, P. Macha, U. Hübner, E. Il' ichev, and A. V. Ustinov, 'Frequency division multiplexing readout and simultaneous manipulation of an array of flux qubits,' Applied Physics Letters **101**, 042604 (2012).

[199] C. D. Wilen, S. Abdullah, N. A. Kurinsky, C. Stanford, L. Cardani, G. D' Imperio, C. Tomei, L. Faoro, L. B. Ioffe, C. H. Liu, A. Opremcak, B. G. Christensen, J. L. DuBois, and R. McDermott, 'Correlated charge noise and relaxation errors in superconducting qubits,' Nature **594**, 369–373 (2021).

[200] T. Walter, P. Kurpiers, S. Gasparinetti, P. Magnard, A. Potočnik, Y. Salathé, M. Pechal, M. Mondal, M. Oppliger, C. Eichler, and A. Wallraff, 'Rapid high-fidelity single-shot dispersive readout of superconducting qubits,' Physical Review Applied **7**, 054020 (2017).

[201] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, 'Measurement of qubits,' Physical Review A **64**, 052312 (2001).

[202] R. Fazili, P. S. Chauhan, U. Chandrashekara, A. Kržič, R. Gómez, M. Gräfe, and F. Steinlechner, 'Simple but efficient polarization-entangled photon sources,' Journal of the Optical Society of America B **41**, 2692 (2024).

[203] A. Z. Goldberg, 'Quantum polarimetry,' in *Progress in optics* (Elsevier, 2022), pp. 185–274,

[204] K. Życzkowski, K. A. Penson, I. Nechita, and B. Collins, 'Generating random density matrices,' Journal of Mathematical Physics **52**, 062201 (2011).