Applied theory for practical silicon-based quantum computing: from hardware to applications



Hamza Jnane

St-Cross College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2024

To my family

Acknowledgements

The past few years have been an incredible and exciting journey, both personally and academically. I am usually reluctant to express my feelings, but it is important for me to take a moment to thank people without whom this would not have been possible:

Simon Benjamin for being the best supervisor I could have ever hoped for. Your patience, the care you show to your students, and your scientific brilliance continue to amaze me. I am also grateful for your availability, even when I shifted to unusual working hours. I have learned so much from you and look forward to continuing to learn more in the future.

Bálint Koczor, for acting as my co-supervisor on multiple occasions. Thank you for helping me navigate the intricacies of the academic world and for being so supportive. I can confidently say that I was fortunate to work with a brilliant future leader of this field. I wish you all the best with your nascent group, and I am certain your future students will appreciate you as much as I appreciated Simon.

Current and past members of the Qtech Theory Group. *Gregory Boyd, Zhenyu Cai, Armands Strikis* for our Belgian adventures. *Hans Han Song Chan, Daniel Marti Dafcik, Arthur Rattew*, although we did not manage to see each other as much as I would have liked. *Christian Binder, Cica Gustiani, Tyson Jones, Richard Meister* thank you for our many office discussions and debates. *Tim Chan* and *Shaun Pexton* for our discussions on food and music – you will not make me change my mind regarding multiple orders of the same dish! *Adrian Chapman* for being an amazing collaborator, and I look forward to see you in person soon. *Minjun Jeon* for being both a fantastic office mate and flatmate. Finally, *Adam Siegel*, for being the most unexpected friendship I could have imagined. Despite our apparent differences, I am grateful to have found someone who shares a similar view on the importance of beauty in this world. While I still question your music taste, I am thankful for all your recommendations. Finally, thanks for Formosan which helped me go through Oxford's depressing winters.

The Quantum Motion team. This section is too small to fully express my gratitude towards the incredible people I met there. I would like to thank *James Palles-Dimmock* and *John Morton* for giving me the opportunity to join the company even if I lacked experience in the field, *Andrew Fisher* for your help on various occasions, *Fernando Gonzalez-Zalba* for being an excellent collaborator and giving me the chance to work on more experimentally-focused projects. Finally, a huge thanks to: *Jagadeesh Bondela, Jacob Chittock-Wood, Virginia Ciriano-Tejel, Isobel*

Clarke, Julio Figuero Quntana, Chris Gamlath, Sophia Garofeanu, David Ibberson, Constance Lainé, Ross Leon, Christine Ng, Patrycja Ozguven, Sofia Patomaki, Frederic Schlattner, Jack Turner, and *James Williams* for all the pub/conference trips, table tennis games and fun I had during the last few years.

Abdelaziz Benzidia, my high school mathematics teacher, for showing me how fun mathematics can be.

Ola Sidorkiewicz and *Morategi Kale* for being incredible flatmates. Thank you for teaching me so much about your respective fields.

My friends outside of academia *Nassim Amer-Ouali, Ahmed Boualam, Youssef Fahim, Saad Fassi-Fihri, Khaled Grira* and *Samuel Halimi*. Although we could not meet as often as I would have liked, thank you for our adventures around the world which helped me escape from Oxford's monotony.

Lastly, I want to thank my family for keeping me grounded and more especially my parents for their unwavering support and the sacrifices they have made for me to be able to write these words. I hope you are proud, this is just the start.

Abstract

Over the past few years, quantum computing has seen rapid development. With qubit technologies such as ions, superconducting circuits and neutral atoms showing early signs of logical qubits, we are getting closer to building a fault-tolerant quantum computer. However, the quest for the optimal qubit remains open. Although less mature, silicon-based spin qubits show significant potential. This thesis presents a bottom-up exploration of silicon-based quantum computers, from hardware to potential applications.

Starting with the explicit modelling of electrons' spins as qubits, I introduce a tool to study two-qubit gates and their robustness against various noise sources, which can inform better chip design. Assuming scalable qubit integration, single-qubit operations are expected to be achieved through global control, which inevitably introduces crosstalk. To tackle this, I propose techniques to mitigate this effect, assuming reliable two-qubit operations and electron shuttling.

With qubit operations addressed, I then explore architectures. While the compact size of spin qubits could allow for millions of qubits on a single chip, I explore a more practical approach: the multicore architecture. This modular design yields more practical quantum computers through either parallelisation of measurements when unconnected or by employing powerful error mitigation techniques using interlinks, as studied here.

Finally, I investigate potential applications of such quantum computers using classical shadows, a method to extract multiple properties of a quantum state at a low measurement cost. In the absence of error correction, near-term quantum computers will be crippled by errors as explored in the first part of this thesis, leading to useless classical shadows. By combining error mitigation techniques with classical shadows, I demonstrate how to reliably and efficiently extract noiseless properties from noisy quantum computers, albeit at an increased measurement cost. Additionally, I show how to further reduce the measurement cost by deliberately introducing bias into the classical shadow estimator, enhancing the practicality of this approach.

Contents

1 Introduction **1**

1.1 Introduction to quantum computing 2

1.1.1 Mathematical formalism 2

1.1.2 What is a quantum computer? 3

1.2 Introduction to silicon-based quantum computing 4

1.2.1 A scalable platform 4

1.2.2 Qubit states and initialisation 5

1.2.3 Qubit operations 8

1.2.4 Coherence times 12

1.2.5 Qubit measurements 13

2 Ab initio modelling of quantum dot qubits **16**

2.1 Introduction 17

2.2 Model 19

2.2.1 Hamiltonian 19

2.2.2 Grid-based simulation 20

2.3 Exchange coupling 24

2.3.1 State preparation 24

2.3.2 Numerical results 26

2.4 Exchange gate dynamics 27

2.4.1 Qubit states 27

2.4.2 Numerical results 29

2.5 Impact of charge noise on exchange gate 31

2.5.1 Static charge 32

2.5.2 Fluctuating charge 33

2.6 Robust pulse sequences 36

2.6.1 Types of errors 36

2.6.2 Robust pulse sequence 38

2.6.3 Numerical results 40

2.7 Future directions and conclusion 42

3 Harnessing electron transport for global spin qubit control 44
3.1 Introduction . . . 45
3.2 Preliminaries . . . 46
3.2.1 Theoretical model . . . 46
3.2.2 Problem statement . . . 47
3.2.3 Clustering . . . 48
3.2.4 Workflow . . . 50
3.2.5 Metrics . . . 51
3.3 Effective frequencies using SWAP gates . . . 52
3.3.1 Two-qubit case . . . 53
3.3.2 n-qubit case . . . 55
3.3.3 Performance on a 2×N array . . . 57
3.4 Effective frequencies using shuttling . . . 59
3.4.1 Single-qubit case . . . 59
3.4.2 n-qubit case . . . 61
3.4.3 Performance on a 2×N array . . . 62
3.5 Discussion . . . 65

4 Multicore quantum computing 67
4.1 Introduction . . . 68
4.2 Linking distant spin qubits using a cavity . . . 69
4.3 Generating Bell Pairs using cavities . . . 72
4.3.1 The model . . . 72
4.3.2 Bell pair generation . . . 74
4.3.3 Entanglement purification . . . 78
4.4 Linking distant spins using shuttling . . . 83
4.4.1 Overview of electron shuttling . . . 83
4.4.2 Distributing Bell pairs using shuttling . . . 84
4.5 Comparison of entanglement distribution modes . . . 84
4.6 Applications . . . 86
4.6.1 Error mitigation . . . 86
4.6.2 Other applications . . . 90
4.7 Discussion . . . 91

5 Introduction to quantum error mitigation and classical shadows 93
5.1 Quantum error mitigation . . . 94
5.1.1 Probabilistic error cancellation . . . 94
5.1.2 Error extrapolation . . . 97
5.1.3 Symmetry verification . . . 99
5.2 Classical shadows . . . 100
5.2.1 Classical shadows with idealised measurements . . . 101
5.2.2 Classical shadows with noisy measurements . . . 102

6 Quantum error mitigated classical shadows 104
6.1 Introduction . . . 105
6.2 Probabilistic error cancelled shadows . . . 106
6.2.1 Protocol . . . 107
6.2.2 Rigorous performance guarantees . . . 108
6.2.3 Post-processing algorithms computing expectation values . . . 111
6.3 Further error mitigation techniques . . . 113
6.3.1 Error boosted shadows . . . 113
6.3.2 Symmetry verified shadows . . . 114
6.4 Applications . . . 115
6.4.1 Ground-state preparation . . . 117
6.4.2 Error mitigated estimation of entanglement entropies . . . 122
6.4.3 Further applications . . . 123
6.5 Discussion . . . 124

7 Biased estimator channels for classical shadows 126
7.1 Introduction . . . 127
7.2 Biased estimator channel . . . 128
7.2.1 Classical shadows review . . . 128
7.2.2 Biased local estimator channel . . . 129
7.2.3 Optimal biasing strategy . . . 129
7.3 Analysis of the worst- and best-case scenarios . . . 131
7.3.1 Best-case scenario . . . 131
7.3.2 Worst-case scenario . . . 132
7.4 Practical application . . . 135
7.5 Discussion . . . 138

8 Conclusion 139

Appendices

A Appendix for Ab-initio modeling of quantum dot qubits 144
A.1 Device specifications . . . 144
A.2 Abrupt transition . . . 145
A.3 Charge trapped outside of the dot line . . . 145
A.4 Fluctuating charges . . . 146
A.5 Convergence versus basis size . . . 148

B Appendices for Harnessing transport for global spin qubit control 150
B.1 Binning . . . 150

C Appendices for Multicore quantum computing **152**
C.1 Parameter regime 152
C.2 Lindblad master equation 154
C.3 Time evolution 155
C.4 State prior to measurement 155
C.5 Alternative purification protocol 157
C.6 Average generation rate 157

D Appendices for Introduction to quantum error mitigation and classical shadows **159**
D.1 Cost of error mitigation 159
D.2 Details on probabilistic error cancellation 160
D.3 Details on classical shadows 161

E Appendices for Quantum error mitigated shadows **162**
E.1 Details on probabilistic error cancelled shadows 163
E.1.1 Unbiased estimator 163
E.1.2 Shadow norm 163
E.1.3 Performance guarantees 165
E.2 Details on classical post-processing algorithms 170
E.2.1 Expectation value of local Paulis 170
E.2.2 Improved algorithm using light cones 171
E.2.3 Local purities 171

F Appendices for Biased Estimator for Classical Shadows **173**
F.1 Derivation of the optimal biasing point 173
F.2 SNR when using classical shadows 174
F.3 Critical biasing 175

1

Introduction

This introduction connects the mathematical formalism of quantum computing, presented in Section 1.1, to a specific hardware implementation: silicon spin qubits. In Section 1.2, I illustrate how spin qubits in silicon offer a promising platform for building scalable quantum computers. The purpose of this presentation is to describe how the abstract quantum operations performed throughout the thesis can be realised in practice.

Contents

1.1 Introduction to quantum computing 2

1.1.1 Mathematical formalism 2

1.1.2 What is a quantum computer? 3

1.2 Introduction to silicon-based quantum computing . . . 4

1.2.1 A scalable platform 4

1.2.2 Qubit states and initialisation 5

1.2.3 Qubit operations 8

1.2.4 Coherence times 12

1.2.5 Qubit measurements 13

1.1 Introduction to quantum computing

1.1.1 Mathematical formalism

A qubit, or quantum bit, serves as the fundamental building block of quantum computing, analogous to the bit in classical computing. Using Dirac (or the ket-bra) notation, a qubit is abstractly described by a vector (or ket) $|\psi\rangle$ defined as,

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle\,, \tag{1.1}$$

living in a two-dimensional complex Hilbert space $\mathcal{H}$ spanned by the basis $\{|0\rangle\,, |1\rangle\}$ where $\alpha, \beta \in \mathbb{C}$ such that $|\alpha|^2 + |\beta|^2 = 1$. $|\psi\rangle$ is also called a state vector. To each ket $|\phi\rangle$ we can associate the linear form, or bra, $\langle\phi|$ defined as,

$$\langle\phi| : \mathcal{H} \to \mathbb{C} \tag{1.2}$$

$$|\psi\rangle \mapsto \langle\phi|\psi\rangle \tag{1.3}$$

with $\langle\cdot|\cdot\rangle$ the inner product in $\mathcal{H}$. Given a state $|\psi\rangle$, the probability to measure the qubit in the state $|\phi\rangle$ is given by applying the bra $\langle\phi|$ to $|\psi\rangle$ and computing the squared absolute value of the result. For instance, for $|\psi\rangle$ as defined in Eq. (1.1), the probability of finding the qubit in the state $|0\rangle$ is given by,

$$P(|0\rangle) = |\,\langle 0|\psi\rangle\,|^2 = |\alpha|^2. \tag{1.4}$$

When considering a system composed of multiple qubits, the space of multi-qubit states is spanned by the tensor product of individual single-qubit basis states. For instance, the state of two unentangled qubits, with states $|\psi\rangle$ and $|\phi\rangle$, is given by $|\psi\rangle \otimes |\phi\rangle$ or $|\psi\rangle\,|\phi\rangle$ for short.

The goal of quantum information processing is to develop algorithms acting on qubits to solve intractable problems when using classical techniques. Various formalisms exist for describing these algorithms [1–3], here we focus on the most used: the gate-based formalism. In this approach, a quantum algorithm is described by a quantum circuit composed of gates. In the absence of noise, each gate is represented by a unitary matrix U_i acting on one or multiple qubits, and one can compactly represent a circuit by the unitary $U = \bigotimes_i U_i$.

In the presence of noise, this formalism is not sufficient to describe a quantum system. For instance, the environment could measure a qubit without us knowing the outcome of the measurement, or, one might not be able to prepare a state perfectly but knows the probabilities $\{p_n\}_n$ to prepare one of the normalized states $\{|\psi_n\rangle\}_n$. However, this framework does not consider classical uncertainty. To include these ideas, the standard approach is to introduce the concept of density operator (or density matrix). We call density operator, the Hermitian operator representing a statistical mixture of state vectors, given by,

$$\rho = \sum_n p_n \left|\psi_n\right\rangle\!\left\langle\psi_n\right|, \tag{1.5}$$

Supposing that one p_i is equal to one, then $\rho = |\psi_i\rangle\langle\psi_i|$. We call these type of states pure states. They can be represented by a single state vector, indicating that the density matrix approach encapsulates the state vector formalism. Density operators are also used to study open quantum systems, which will be relevant in the study of interactions between a quantum system and its environment (see e.g. Chapter 4). A valid density operator must follow the following requirements,

- $\mathrm{Tr}(\rho) = 1$ assuring that the $\{p_i\}$ sum to one.
- The matrix ρ is positive semi-definite, meaning that the $\{p_i\}$ are non-negative.

Quantum operations on density matrices (or quantum channels) are more general than for pure states, they are given by completely positive trace-preserving or CPTP maps (see [1] for more information). This is the kind of operation I will focus on in Chapter 5 and Chapter 6.

1.1.2 What is a quantum computer?

Quantum computers are devices that leverage the principles of quantum mechanics to achieve an algorithmic speed up for certain problems. Although the theoretical formalism is well established, building quantum computers that demonstrably show an advantage over their classical counterparts remains a formidable challenge [4–10]. In the spirit of giving a more practical perspective, DiVincenzo proposed five criteria that a device must meet to successfully run quantum algorithms [11]:

1. It must be scalable and composed of well-characterised qubits
2. It should allow for the reliable initialisation to a known quantum state
3. Qubits must have long coherence times compared to gate operations
4. It should enable the implementation of a universal set of quantum gates
5. It must be capable of measuring the state of qubits

Various technologies fulfil these requirements such as superconducting qubits [12, 13], ion traps [14], neutral atoms [15], photons [16, 17], and semiconductor qubits [18, 19]. It is worth noting that the first three are the most advanced, and on the track to show the reliable implementation of logical qubits [8, 10, 20]

In this thesis, I will focus on one of the emerging platforms for the realisation of quantum computers: semiconductor spin qubits [19, 21, 22]. Specifically, I will focus on silicon-based spin qubits, which offer multiple advantages as explained below. Note that I will not explore charge qubits in depth because of their small coherence times [18]. However, I will briefly mention them in Chapter 4 in the context of linking distant spins to microcavities.

1.2 Introduction to silicon-based quantum computing

In this section, I will review each of DiVincenzo's criteria and explain how silicon spin qubits meet them.

1.2.1 A scalable platform

Silicon spin qubits are a strong candidate for qubit implementation due to their small size, fast operations speeds, and long coherence times. However, their most significant advantage over other technologies is their potential compatibility with existing industry standards [23–28], making them ideal for building scalable quantum computers.

1.2.2 Qubit states and initialisation

As their name suggests, silicon spin qubits are encoded in the spin of one or more particles in silicon. There are multiple flavours of spin qubits [19], each with its own advantages and drawbacks. Qubits can be encoded in the spin of donor impurities [22], in the spin of an electron [21], or even in the spin of a collection of electrons [19, 29–33]. In this manuscript, I will focus on the simplest kind of spin qubit, the so-called Loss-DiVincenzo (LD) qubits, which are encoded in the spin of a single electron. Before discussing how one can reliably initialise the system to a known quantum state, it is important to properly define LD qubits.

Single qubit

By applying a magnetic field to a single electron, the spin degeneracy is lifted, allowing the electron's spin states in the direction of the field to be associated with the qubit as follows,

$$\begin{aligned} |\downarrow\rangle &\Longleftrightarrow |0\rangle \\ |\uparrow\rangle &\Longleftrightarrow |1\rangle\,. \end{aligned} \tag{1.6}$$

To use the spin of a single electron as a qubit, we first need to isolate the electron. In a semiconductor material, this is achieved by confining the electron in a small region of the material. While there are multiple ways of trapping an electron in a semiconductor, such as using the Coulomb potential of dopant atoms for donor spin qubits or a combination of etching and electrostatic gating for FinFETs [19], the focus of this thesis is on quantum dots generated using a combination of layered semiconductor heterostructures and electrostatic gating [18, 19, 34]. The layered heterostructure provides confinement in the z-direction, while electrodes confine the electron within the xy-plane. The most common heterostructures are Si/SiGe and Si/SiO_2 that could be easily manufactured using current industry standards [23–28].

When dealing with a single qubit, the electron's charge is often not relevant, only the spin states matter as the electron is confined. However, when defining the

states of multiple qubits, the charge must be considered to respect the antisymmetry requirement for a system composed of indistinguishable electrons.

Extension to multiple qubits

I will start by defining the qubit states of two electrons before extending the definition to multiple qubits. Since we are dealing with indistinguishable particles, the state of two qubits cannot be simply represented by the tensor product of the previously defined single qubit states. For instance, if we had $|01\rangle \coloneqq |{\downarrow}{\uparrow}\rangle$, the system will not be antisymmetric under particle exchange as $\mathrm{SWAP}\,|01\rangle = \mathrm{SWAP}\,|{\downarrow}{\uparrow}\rangle = |{\uparrow}{\downarrow}\rangle$, where SWAP is the operator exchanging the electrons. To satisfy the antisymmetry requirement, we need a more nuanced definition of multi-qubit states that involves the charge of the electrons.

Suppose we have two electrons confined in tunnel-coupled quantum dots, known as a double quantum dot (DQD), with each electron occupying a single dot. To simplify, I will associate the fact that an electron is confined in the left or right dot with the charge states $|\mathrm{L}\rangle$ and $|\mathrm{R}\rangle$ respectively.

The state of a qubit should ultimately be defined by what an experimentalist can measure, i.e., the spin state of an electron found in a *specific* dot. For instance, consider the *qubit* state $|01\rangle$. This notation means that upon measurement, I would find the qubit state $|0\rangle$ in the left dot, and $|1\rangle$ in the right dot. As electrons are indistinguishable, this state should contain contributions from both electrons. There are two scenarios that could lead to this state: the first electron is found on the left dot with a spin state $|{\downarrow}\rangle$ while the second is in the right dot with spin state $|{\uparrow}\rangle$, and vice versa. Both scenarios occur with equal probability, leading to the following definition,

$$|01\rangle \coloneqq \frac{1}{\sqrt{2}} \left(|\mathrm{LR}\rangle\, |{\downarrow}{\uparrow}\rangle - |\mathrm{RL}\rangle\, |{\uparrow}{\downarrow}\rangle \right), \tag{1.7}$$

where $|\mathrm{LR}\rangle$ means that the first electron is on the left while the second is on the right. Using this definition, it is straightforward to see that the state is now

antisymmetric with respect to particle exchange. Following a similar reasoning the remaining qubit states are given by,

$$\begin{aligned}
|00\rangle &\coloneqq \frac{1}{\sqrt{2}}\left(|\mathrm{LR}\rangle - |\mathrm{RL}\rangle\right)|{\downarrow\downarrow}\rangle\,,\\
|10\rangle &\coloneqq \frac{1}{\sqrt{2}}\left(|\mathrm{LR}\rangle\,|{\uparrow\downarrow}\rangle - |\mathrm{RL}\rangle\,|{\downarrow\uparrow}\rangle\right),\\
|11\rangle &\coloneqq \frac{1}{\sqrt{2}}\left(|\mathrm{LR}\rangle - |\mathrm{RL}\rangle\right)|{\uparrow\uparrow}\rangle\,.
\end{aligned} \tag{1.8}$$

To generalise to more qubits, I just need to ensure that antisymmetry is enforced in the state's definition. To do so, I need to introduce the antisymmetrising operator $\mathcal{A}$, which, given a wavefunction of N identical fermions, returns the antisymmetric wavefunction under particle exchange. For N electrons, it is defined as,

$$\mathcal{A} \coloneqq \frac{1}{\sqrt{N!}} \sum_{P\sim\pi\in\mathbb{S}_N} (-1)^{\pi} P, \tag{1.9}$$

where $\mathbb{S}_N$ is the symmetric group (defined as the set of all permutations of $\{1, 2, \ldots, N\}$), P is the linear operator associated with the permutation π and $(-1)^{\pi}$ denotes the parity of π. The general N-qubit state $|q_1 q_2 \ldots q_N\rangle$ is given by,

$$|q_1 q_2 \ldots q_N\rangle \coloneqq \mathcal{A}\left(|c_1^1 c_2^2 \ldots c_N^N\rangle\,|s_1^1 s_2^2 \ldots s_N^N\rangle\right), \tag{1.10}$$

with $|q_i\rangle \in \{|0\rangle\,, |1\rangle\}$ the qubit state of the ith dot, $|c_i^j\rangle$ representing the charge state of the jth electron when it is located in the ith dot, and $|s_i^j\rangle$ the corresponding spin state. Note that,

$$\forall\ i, j\ \ |s_i^j\rangle = \begin{cases} |{\downarrow}\rangle & \text{if } |q_i\rangle = |0\rangle \\ |{\uparrow}\rangle & \text{if } |q_i\rangle = |1\rangle \end{cases}. \tag{1.11}$$

Using Eq. (1.9) and Eq. (1.10), we obtain:

$$|q_1 q_2 \ldots q_N\rangle \coloneqq \frac{1}{\sqrt{N!}} \sum_{\pi\in\mathbb{S}_N} (-1)^{\pi}\,|c_{\pi(1)}^1 c_{\pi(2)}^2 \ldots c_{\pi(N)}^N\rangle\,|s_{\pi(1)}^1 s_{\pi(2)}^2 \ldots s_{\pi(N)}^N\rangle\,. \tag{1.12}$$

Initialisation to a fiducial state

One can then prepare the fiducial N-qubit state $|00\ldots0\rangle$,

$$|00\ldots0\rangle \coloneqq \frac{1}{\sqrt{N}} \sum_{\pi\in\mathbb{S}_N} (-1)^{\pi}\,|c_{\pi(1)}^1 c_{\pi(2)}^2 \ldots c_{\pi(N)}^N\rangle\,|{\downarrow\downarrow}\ldots{\downarrow}\rangle\,. \tag{1.13}$$

This is performed by pulling qubits one at a time from an electron reservoir using spin-dependent tunnelling as explained in [19].

1.2.3 Qubit operations

In this subsection, I explain how single- and two-qubit gates are performed on LD qubits and show that the corresponding gate set is universal.

Single-qubit gates

The spin degree of freedom hardly couples to electrical perturbations, due to silicon's small spin-orbit coupling. While this allows for long coherence times, it also means that it is hard to electrically control a qubit state. Ideally, electrical control is preferred since it is easier to apply a specific voltage to a quantum dot using gate electrodes rather than localising a magnetic field to a specific qubit.

By artificially enhancing the spin-orbit coupling, which is achievable using engineered magnetic field gradients, one can electrically control the state of a qubit [35, 36]. This method is called Electron Dipole Spin Resonance (EDSR). While this technique has been successfully demonstrated for a small number of qubits, it seems clear that it is not scalable as it would require engineering of magnets on top of millions of qubits [37, 38].

An alternative approach relies on Electron Spin Resonance (ESR), where an off-chip global magnetic field is generated to manipulate qubits [39]. This method has issues such as crosstalk, which I address in Chapter 3.

The Hamiltonian associated to a single qubit in the presence of a fixed magnetic field (used to lift the spin degeneracy) $\vec{B}_0 = (0, 0, B_0)$ in the z-direction is given by,

$$H = \mu_B g \vec{B}_0 \cdot \vec{S} = \frac{1}{2}\mu_B g B_0 \sigma_z = \frac{\omega_0}{2}\sigma_z, \tag{1.14}$$

where $\hbar = 1$ and $\omega_0 \coloneqq \mu_B g B_0$, with μ_B Bohr's magneton and g the electron's g-factor. In this case, note that ω_0 is also equal to the Zeeman splitting E_z. By adding a time-dependent global field $\vec{B}_1 = (B_1 \cos(\omega t + \phi), 0, 0)$ the Hamiltonian of the system in the laboratory frame becomes,

$$H(t) = \frac{\omega_0}{2}\sigma_z + \Omega \cos(\omega t + \phi)\sigma_x, \tag{1.15}$$

with $\Omega \coloneqq \frac{1}{2}\mu_B g B_1$. To simplify the study of the dynamics, we usually work in a rotating frame at the drive's frequency ω. To obtain the Hamiltonian in this new reference frame, I perform the following transformation,

$$H(t) \rightarrow H^{'}(t) = U(t)H(t)U^{\dagger}(t) + i\frac{dU}{dt}(t)U^{\dagger}(t), \tag{1.16}$$

with

$$U(t) = \exp\left(i\frac{\omega}{2}\sigma_z t\right). \tag{1.17}$$

The Hamiltonian in the new reference frame becomes time-independent and now reads,

$$H^{'} = \frac{\delta\omega}{2}\sigma_z + \frac{\Omega}{2}\left(\cos(\phi)\sigma_x + \sin(\phi)\sigma_y\right), \tag{1.18}$$

where I used the rotating wave approximation to neglect fast-oscillating terms and $\delta\omega = \omega_0 - \omega$ being the detuning between the qubit's and drive frequencies. Thus when driving on resonance at ω_0, the Hamiltonian simplifies to,

$$H^{'} = \frac{\Omega}{2}\left(\cos(\phi)\sigma_x + \sin(\phi)\sigma_y\right). \tag{1.19}$$

By examining Eq. (1.19), we see that depending on the drive's phase, we can control rotations around two orthogonal axes, enabling the implementation of any single-qubit gate. Z gates can also be performed directly by modifying the qubit's frequency ω_0 through Stark shift (i.e. electrically modifying the electron's g-factor) or implemented virtually under certain conditions [40, 41].

Suppose now that $\omega \neq \omega_0$ and $\phi = 0$. One can show that the eigenstates of the Hamiltonian are given by,

$$\begin{aligned} |\psi_-\rangle &= -\sin\left(\frac{\theta}{2}\right)|0\rangle + \cos\left(\frac{\theta}{2}\right)|1\rangle \\ |\psi_+\rangle &= \cos\left(\frac{\theta}{2}\right)|0\rangle + \sin\left(\frac{\theta}{2}\right)|1\rangle \end{aligned} \tag{1.20}$$

with $\theta = \arctan(\Omega/\delta\omega)$ and respective energies $E_\pm = \pm\frac{1}{2}\sqrt{\delta\omega^2 + \Omega^2}$. Suppose that I start from the state $|\psi(0)\rangle = |0\rangle$. The time-evolved state in the reference

frame $|\psi^{'}(t)\rangle$ is given by,

$$\begin{aligned}|\psi^{'}(t)\rangle &= \exp\Big(-iH^{'}t\Big)\,|\psi(0)\rangle \\ &= \exp\Big(-iH^{'}t\Big)\left(-\sin\left(\frac{\theta}{2}\right)|\psi_{-}\rangle + \cos\left(\frac{\theta}{2}\right)|\psi_{+}\rangle\right) \\ &= -\exp(-iE_{-}t)\sin\left(\frac{\theta}{2}\right)|\psi_{-}\rangle + \exp(-iE_{+}t)\cos\left(\frac{\theta}{2}\right)|\psi_{+}\rangle\,. \end{aligned} \tag{1.21}$$

I can now compute the evolution of the probability to find the qubit in the state $|1\rangle$ over time,

$$\begin{aligned}P_{|1\rangle}(t) &= \Big|\langle 1|\psi^{'}(t)\rangle\Big|^2 \\ &= \left|\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right)\left(\exp(-iE_{+}t) - \exp(-iE_{-}t)\right)\right|^2 \\ &= \left|\frac{\sin(\theta)}{2}\left(\exp(-i(E_{+}-E_{-})t) - 1\right)\right|^2 \\ &= \Big|\sin(\theta)\sin\Big(\sqrt{\delta\omega^2+\Omega^2}\,t\Big)\Big|^2 \\ &= \left|\frac{\Omega}{\sqrt{\delta\omega^2+\Omega^2}}\sin\Big(\sqrt{\delta\omega^2+\Omega^2}\,t\Big)\right|^2 \\ &= \frac{\Omega^2}{\delta\omega^2+\Omega^2}\sin^2(\sqrt{\delta\omega^2+\Omega^2}\,t), \end{aligned} \tag{1.22}$$

this is the Rabi formula which I will refer to in Chapter 3. Note that the frequency $\sqrt{\delta\omega^2+\Omega^2}$ is called the Rabi frequency and characterises the speed of single-qubit gates.

Single-qubit fidelities above 99.9% have been achieved experimentally [42–46] with Rabi frequencies below 10 MHz.

Two-qubit gates

The natural way to perform a two-qubit gate for silicon spin qubits in gate-defined quantum dots is to leverage the intrinsic exchange interaction between neighbouring qubits [19, 47–49]. For two qubits in a DQD, this interaction emerges from the combination of the antisymmetry requirement and the Coulomb repulsion between the electrons, as I explain in Chapter 2. In this subsection, it suffices to know

that the Hamiltonian describing the interaction between two qubits in a DQD is given by the Heisenberg Hamiltonian [19],

$$H = J\vec{S}_1 \cdot \vec{S}_2, \tag{1.23}$$

with J the exchange coupling and $\vec{S}_i = \hbar/2\,(\sigma_x, \sigma_y, \sigma_z)$ the spin operator for each electron. While evaluating J is hard in practice, I present how to compute it in Chapter 2 for realistic devices. In the basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$, the Hamiltonian is given by (up to a constant),

$$H = \frac{J}{2}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{1.24}$$

where I set $\hbar = 1$. This matrix represents an operator that swaps the spin states. Evolving a state under this Hamiltonian will then yield a SWAP gate if the interaction is stopped at $T_{\text{SWAP}} = \pi/J$. But most importantly, one can perform a $\sqrt{\text{SWAP}}$ entangling gate by stopping the interaction at $T_{\sqrt{\text{SWAP}}} = \pi/(2J)$. Exchange couplings on the order of hundreds of MHz should be achievable, allowing two-qubits gates to be completed in tens of nanoseconds.

In practice, the two dots will typically have two different g-factors, leading to the following modification of the Hamiltonian [47, 50],

$$H = \frac{1}{2}\begin{pmatrix} -E_z & 0 & 0 & 0 \\ 0 & \Delta E_z - J & J & 0 \\ 0 & J & -\Delta E_z - J & 0 \\ 0 & 0 & 0 & E_z \end{pmatrix}. \tag{1.25}$$

with E_z and ΔE_z the sum and difference in Zeeman splitting between the two dots, respectively. Using this Hamiltonian, one can only generate a $\sqrt{\text{SWAP}}$ gate in the limit $J \gg \Delta E_z$. However, it has been proven than one can use this Hamiltonian to perform a CZ gate through direct evolution [47, 51] or CZ/CNOT gates using additional pulses [48, 52–54].

Two-qubit gates fidelities above 99.5% have been achieved in practice [44–46] with exchange couplings on the order of tens of MHz.

A universal set of gates

To implement any quantum circuit, a quantum computer must be capable of performing a universal set of gates, meaning any gate should be decomposable into this set. Barenco et al. [55] demonstrated that a set containing the CNOT gate along with the ability to perform every single-qubit gate is universal. As I explained previously, any single-qubit gates can be realised using ESR, and a CNOT can be achieved through the combination of a CZ and Hadamard gates or directly by using microwave pulses. Thus, silicon-based quantum computers fulfil DiVincenzo's fourth criterion.

1.2.4 Coherence times

To characterise the lifetime of a qubit's state, it is important to introduce different timescales. First, there is the relaxation time T_1, which corresponds to the characteristic time of the relaxation process where an electron transitions from its excited state to its ground state. Or in practice, the electron relaxes to a state determined statistically by the system's temperature. In our case, this would correspond to a spin relaxing from the state $|\uparrow\rangle$ to $|\downarrow\rangle$. Next, there is the dephasing time T_2^*, which corresponds to the process where the phase of a qubit is altered by the environment. In this subsection, I summarise the different noise processes impacting these timescales for silicon spin qubits.

Spin relaxation occurs when the qubit exchanges energy with the environment, typically due to phonon emission in our case. T_1 times on the order of seconds have been reported in Si/SiO_2 and Si/SiGe devices [24, 56–58], and this will usually not be the limiting timescale.

By encoding qubits in the spins of electrons, they become sensitive to their magnetic environment. More precisely, nuclear spins in the semiconductor interact with qubits through the hyperfine interaction [19, 51, 59] leading to static dephasing, which affects T_2^*. In practice, T_2^* is often the limiting timescale, but careful material selection can alleviate this issue. For instance, while all isotopes of Ga and As have a non-zero nuclear spins, only 4.7% of Si isotopes have non-zero nuclear

spins. By purifying silicon and thus reducing the number of non-zero nuclear spins, T_2^* can be drastically improved from tens of nanoseconds [60–62] to hundreds of microseconds [48, 63, 64].

Another way to increase the coherence time is to use dynamic decoupling (DD) techniques, which involve the repetition of pulses to compensate for the noise introduced by the environment [65, 66]. Amongst these techniques are the well-known Hahn echo and Carr-Purcell-Meiboom-Gill (CPMG) pulses for spin systems. When using DD, the timescale generally used is T_2 instead of T_2^*. While T_2 times on the order of milliseconds were obtained when using Hahn echo [63, 67, 68], one can achieve coherence times of tens of milliseconds when using CPMG sequences [48, 63]. For a continuously updated survey of the different performance metrics in gate-defined quantum dots see [69].

In conclusion, spins of electrons in gate-defined quantum dots exhibit coherence times to gate speeds ratio that are large enough to perform quantum algorithms.

1.2.5 Qubit measurements

To fulfil DiVincenzo's last criterion, one must be able to measure state of the qubits individually. However, due to the the small magnetic dipole of an electron's spin in silicon, it is challenging to directly measure the spin state of a single electron. In contrast, measuring the charge of a single electron is easier using charge sensors such as quantum point contacts [70], single-electron transistors [71] or single-electron boxes [72]. By employing spin-to-charge conversion mechanisms – where the electron's displacement depends on its spin state – one can readout the spin state of a single electron via its charge degree of freedom. In this subsection, I present two different methods to perform readout of the spin of a single electron: Elzerman and Pauli spin blockade readout. While the former is the historical way of performing measurements, the latter is at the core of modern readout techniques. Although this subsection focuses on readout using charge sensors in the vicinity of the dots, more advanced techniques such as gate-based readout are worth exploring

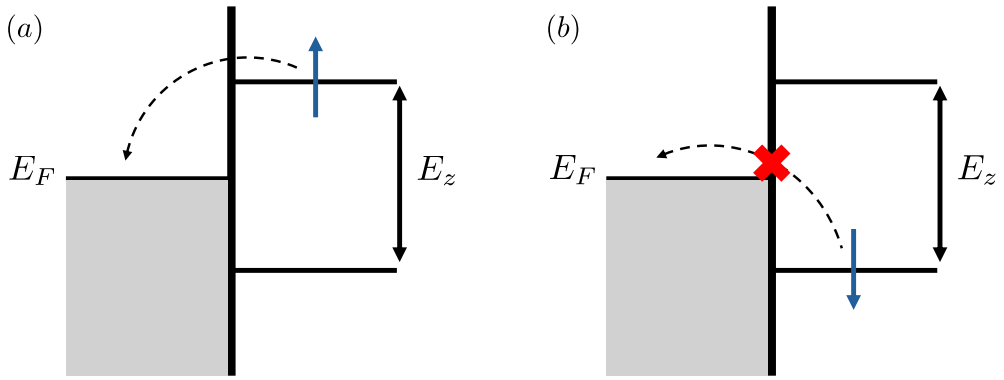

Figure 1.1: Schematic illustration of Elzerman readout where a reservoir (on the left) is coupled to a quantum dot (on the right). (a) An electron in the $|\uparrow\rangle$ state can tunnel out of the quantum dot to the reservoir due to the placement of the Fermi level. (b) An electron in the $|\downarrow\rangle$ state cannot leave the quantum dot because its energy is insufficient. By tracking the charge displacement, we can measure the spin state of the trapped electron.

[73]. The aim of this subsection is to provide a basic understanding of measurement techniques, for a comprehensive review refer to [19].

Elzerman readout

The simplest method to perform spin readout in silicon-based quantum computers is Elzerman readout [74]. This involves coupling the quantum dot containing the electron to a reservoir as shown in Fig. 1.1. By adjusting the gate voltage on the quantum dot, the electron's energy levels can be positioned so that the reservoir's Fermi level lies between the qubit levels. In this configuration, the electron can tunnel out of the quantum dot only if it is in the $|\uparrow\rangle$ state, while an electron in the $|\downarrow\rangle$ state will remain in the quantum dot. A nearby charge sensor can detect such tunnelling event, and hence readout the spin associated to the electron. Note that when the electron is in the $|\downarrow\rangle$ state, no electron from the reservoir can tunnel in the dot because of Coulomb blockade.

However, Elzerman readout has several practical drawbacks. First, the Fermi level of the reservoir is temperature-dependent, requiring $E_z \gg k_B T$ to ensure that E_F is between the spin levels, which leads to high magnetic fields. Second, the method relies on waiting for a tunnelling event to happen, which can result in

long measurement times than may exceed the qubit decoherence time. Lastly, the necessity of a nearby reservoir makes it hard for this technique to scale.

Pauli spin blockade readout

To overcome the limitations of Elzerman readout, Pauli spin blockade (PSB) based readout can be used [75]. This method requires two electrons in a DQD. By modifying the detuning of the DQD, it becomes energetically favorable for the electron in one dot to tunnel to the other dot if the spins are different. If the spins are the same, the transition is forbidden by the Pauli exclusion principle. For instance, a reference spin in the right dot can be used to measure the state of the qubit in the left dot by detecting the charge transition provoked by the change in detuning.

PSB readout addresses the issues associated with Elzerman type readout. By using a DQD, it suppresses the need for an electron reservoir, enhancing scalability and lifting the requirement for high magnetic fields. Moreover, this readout mechanism is faster because the tunnelling event is induced by the modification of the detuning.

In summary, I have explained how silicon spin qubits meet each of the five DiVincenzo criteria, establishing them as a strong candidate for the implementation of large-scale quantum computers.

2

Ab initio modelling of quantum dot qubits

This chapter presents research conducted in collaboration with Simon Benjamin, based on our preprint [76]. While the work presented is my own, Simon independently developed code to validate the evolution of the exchange coupling and discovered the robust pulse sequence discussed here. I then performed the corresponding numerical simulations. Some text I originally wrote is taken verbatim from [76].

Contents

2.1 Introduction 17
2.2 Model 19
2.2.1 Hamiltonian 19
2.2.2 Grid-based simulation 20
2.3 Exchange coupling 24
2.3.1 State preparation 24
2.3.2 Numerical results 26
2.4 Exchange gate dynamics 27
2.4.1 Qubit states 27
2.4.2 Numerical results 29
2.5 Impact of charge noise on exchange gate 31
2.5.1 Static charge 32
2.5.2 Fluctuating charge 33
2.6 Robust pulse sequences 36
2.6.1 Types of errors 36
2.6.2 Robust pulse sequence 38
2.6.3 Numerical results 40
2.7 Future directions and conclusion 42

2.1 Introduction

As mentioned in Chapter 1, silicon spin qubits emerged as a strong candidate for the development of scalable quantum computers due to their small size, long coherence times, fast gate operations, and compatibility with current industry standards [23–26]. However, the vast design space offered by semiconductor foundries makes identifying promising chip architectures challenging. The iterative nature of chip development, coupled with the time and cost of a chip design cycle, highlights the need for accurate predictive modelling tools to speed up the discovery of scalable designs. Given a candidate chip layout defined using standard commercial software (as employed for traditional chips), such tools would extract key metrics and enable one to optimise the design prior to fabrication.

Electron spin qubits are confined by potentials generated by the device's electrodes, and the design space for electrode layouts and corresponding voltage ranges is wide. Of particular interest are double dot potentials, where two electrons can be isolated and experience exchange coupling: the spin-spin interaction resulting from underlying charge state symmetries. The strength of the exchange coupling, denoted as J, and its controllability ultimately limit the speed and quality of any native two-qubit gate [21, 30, 47]. As a result, predicting the exchange coupling for a given double-dot potential by either analytic or numerical methods has been a long-standing effort [77–80]. Due to the complexity of the task, simplifications to the device layout and surrounding environment are often needed to make such predictions feasible.

A critical factor in optimising the fidelity of two-qubit gates is the sensitivity of the exchange coupling to charge noise [81–83]. Recently, several papers have addressed this challenge [84–86] and computed the exchange coupling for realistic scenarios. These works employed techniques to make the computationally expensive numerics more tractable, such as tailoring the basis to obtain a compact form for the Coulomb interaction, coarse-gaining the model in some direction(s), or by using a path integral formulation. Ref. [85] focused on optimising devices geometries to reduce the sensitivity of J to electrical fluctuations, while [84, 86] explicitly added

charge noise to their model using a two-level fluctuator model. The validity of this approach was recently confirmed by experiments [87].

The approach taken in this chapter is *ab initio* in the sense that it makes very few assumptions about the anticipated physics, allowing key properties to emerge once suitable parameters are found. I use a uniform real-space grid method to find the low-lying two-electron states for *any* reasonably smooth potential landscape. As a result, the numerical modelling is fairly demanding but is capable of addressing general scenarios. Here, I focus on a double well potentials derived via a Poisson solver from a realistic device's electrode geometry. Demonstrating the feasibility of this approach is a key step toward the goal of rapidly iterating device layouts virtually, prior to committing to an optimised design to fabrication (see Fig. 2.1).

The chapter is organised as follows. In Section 2.2, I briefly explain how the double quantum dot and electrons' wavefunctions are modelled using real-space grids. Then, in Section 2.3, I show how this tool can be used to compute the exchange coupling in an idealised environment. In Section 2.4, I directly model the dynamics of a $\sqrt{\text{SWAP}}$ gate by continuously deforming the potential, propagating the full two-electron state forward with a split-operator (SO) method. At gate completion, I reassess the validity of the qubit representation, i.e. whether there is an increase in the $(2,0)$ & $(0,2)$ charge configurations, which correspond to qubit loss. This is reported separately from the fidelity of the gate operation within the $(1,1)$ subspace. I observe the effect of non-adiabatic voltage ramp on this loss probability, and verify that a carefully designed one leads to negligible loss. I then assess the impact of unwanted trapped charges in the environment in Section 2.5, either as static charges or as fluctuators that switch position during the gate operation. By reducing the behavior of the full model into a reduced basis, I rapidly explore options for robust pulse sequences in Section 2.6– i.e. a multi-stage gate that corrects for intrinsic axis-error as well as charges in the environment. Adapting methods from the NMR literature [88–91], we exhibit a 9-pulse sequence with remarkable robustness against both effects, albeit at the cost of speed. Finally, I conclude in Section 2.7

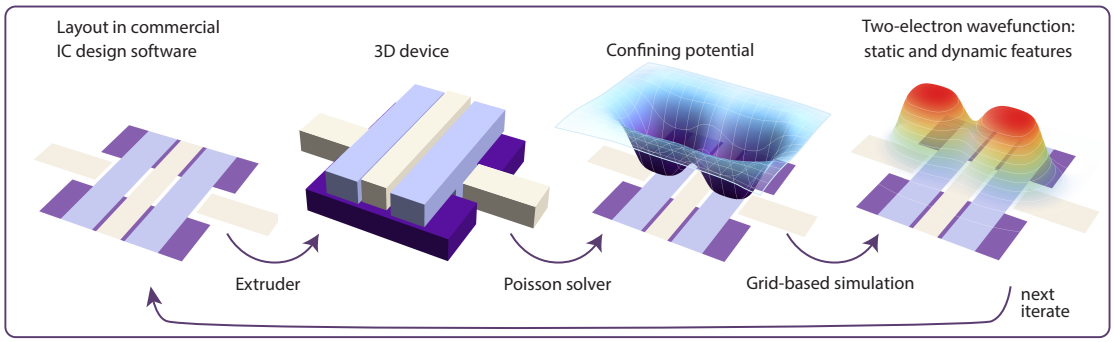

Figure 2.1: Schematic of the iterative design process of a quantum dot device. Starting with an idea of an integrated circuit layout, we would like to know the performance of the chip without sending it through the lengthy fabrication process. Using an extruder, a 3D version of the chip is generated, with specified materials. This then allows me to generate realistic quantum dot potentials. Finally, the grid-based modelling is used to extract the device's key performance metrics. For instance, I can discard designs that show a small exchange coupling between neighbouring dots. Apart from static properties, I can also study the qubits' dynamics and e.g. compute two-qubit gate fidelities, how they are impacted by the presence of charge noise and how we can mitigate this effect. This process informs new layouts, and through repeated iterations, leads to more promising designs that can eventually be sent for fabrication. This figure is extracted from [76] and was made by Tyson Jones from a sketch I designed.

2.2 Model

2.2.1 Hamiltonian

In this chapter, I focus on the study of two electrons trapped in a double quantum dot (DQD) potential (see Fig. 2.2). Under the effective mass approximation, the Hamiltonian of the system is given by,

$$H = \frac{\hbar^2}{2m^*}\mathbf{p_1^2} + \frac{\hbar^2}{2m^*}\mathbf{p_2^2} + V(\mathbf{r_1}) + V(\mathbf{r_2}) + \frac{e^2}{4\pi\epsilon_0\epsilon_r|\mathbf{r_1}-\mathbf{r_2}|} + g(\mathbf{r_1})\mu_B\mathbf{B}\cdot\mathbf{S_1} + g(\mathbf{r_2})\mu_B\mathbf{B}\cdot\mathbf{S_2}, \tag{2.1}$$

with $m^* = 0.19m_e$ the effective mass of an electron in silicon, $\epsilon_r = 11.68$ silicon's relative permittivity, and $g(\mathbf{r_i})$ the position-dependent electron g-factor. The potential V is generated using a Poisson solver on realistic devices. In this initial study, I neglect certain additional effects, while noting that their inclusion would be entirely possible within the approach and represents a potential direction for further work. Specifically, I neglect the impact of the magnetic field on the orbital

motion of the electrons (through its associated vector potential), as my focus is on flat or nearly-flat 2D dots subject to an in-plane field. Moreover, I neglect the valley degree of freedom, given that the valley splitting is usually large in CMOS devices. Lastly, I also neglect spin-orbit coupling, noting that it is small in silicon.

2.2.2 Grid-based simulation

Introduction to grid-based and split-operator methods

A common approach to modelling single or multiple quantum particles, is to represent their state by storing a complex amplitude at each point of a regular grid. This grid can be in real-space, or k-space. This method is broadly a discrete variable representation (DVR) [92], and the approach to driving dynamics used in this paper is a split-operator Fourier transform (SO-FT) method, which dates back to the work of Feit and Fleck in 1976 [93]. Widely employed across various fields, the method involves repeatedly Fourier transforming between the real-space and the dual k-space representations, as described below.

It is also worth noting that this method has recently received attention from researchers seeking efficient algorithms for modelling many-body systems on future quantum computers. The use of the Fourier transform is a natural fit here since this transformation is well-understood and efficiently implementable on quantum machines. The specific variant of the SO method that we use here is one that has been well-explored in that context (see e.g. Refs. [94, 95]).

State representation

For simplicity of presentation, I will start by working in k-space and focus on the case of a single particle in 1D. The state of this system is represented as,

$$\Psi_{\mathrm{KS}}(x) = \frac{1}{\sqrt{L}} \sum_{k=-m}^{m-1} c_k \exp\left(\tfrac{2i\pi kx}{L}\right),$$

where the subscript KS denotes k-space. As it can be rescaled as needed in calculations, I choose a box size $L = 1$, resulting in:

$$\Psi_{\mathrm{KS}}(x) = \sum_{k=-m}^{m-1} c_k \ \psi_k(x) \quad \text{with} \quad \psi_k(x) = \exp\left(2i\pi kx\right). \tag{2.2}$$

One can see that the state is periodic and will repeat itself outside of the box defined by $-\frac{1}{2} < x < \frac{1}{2}$. I am only interested in this central region, and as the confining potential ensures that the wavefunction amplitudes of relevant states vanish well before reaching the borders, I do not have to worry about what is happening outside of the central box region. The scale is such that the wavefunctions have very little amplitude outside of the middle part of this box ($|x|$, $|y|$ less than 1/4). This is important because of the form of the Coulomb interaction, as explained by my co-author in our paper [76].

The index k runs from $-m$ to $m-1$ for a total number of basis states of $M = 2m = 2^q$ for some integer q. Some prior papers make this choice (e.g. Ref. [95]) while others have used an odd total number of k states (e.g. Ref. [94]) which slightly alters the form of the real-space states. These real-space states are found by applying the Fourier transform to the k-space representation:

$$\Psi_{\mathrm{RS}}(x) = \sum_{n=-m}^{m-1} b_n \phi_n(x),$$

where

$$\phi_n(x) = \frac{e^{i\pi(x-p_n)}}{\sqrt{M}} D_m(2\pi(x-p_n)) \quad \text{and} \quad p_n = \frac{n}{M}. \tag{2.3}$$

Here D_m is a Dirichlet kernel,

$$D_m(u) \equiv \frac{\sin\left((m+\frac{1}{2})u\right)}{\sin\left(\frac{u}{2}\right)}.$$

Each $\phi_n(x)$ is peaked at p_n, which are the real-space grid points. With a simulation box running from $x = -\frac{1}{2}$ to $x = +\frac{1}{2}$, p_n runs from $-\frac{1}{2}$ through zero to $\frac{1}{2} - \frac{1}{M}$. Due to the periodicity, $-\frac{1}{2}$ and $\frac{1}{2}$ are identical. The basis in Eq. (2.3) is also called the dual basis, real-space basis, or sometimes the Dirichlet basis.

In my simulations, I initially represent the state in k-space by storing the amplitudes c_k. The transformation to the real-space picture thus simply requires the computation of the terms b_n as follows:

$$b_n = \frac{1}{\sqrt{M}} \sum_{k=-m}^{m-1} \exp\left(\tfrac{2i\pi n\, k}{M}\right) c_k.$$

Importantly, $\phi_n(p_m) = 0$ for all $m \neq n$, meaning that at the peak of any given basis function, all others are equal to zero. This leads to a simple decomposition of any given target wavefunction $f(x)$ into this form. I only need to evaluate the target function at the grid points, yielding:

$$f(x) \approx \Psi_{\mathrm{RS}}(x) = \sum_{n=-m}^{m-1} f(p_n)\phi_n(x). \tag{2.4}$$

This approach can naturally be extended to higher dimensions. In two dimensions, for example, we get:

$$f(\mathbf{r}) \approx \Psi(\mathbf{r}) = \sum_{n_x,n_y} f(p_{n_x}, p_{n_y})\, \phi_{n_x}(x)\, \phi_{n_y}(y), \tag{2.5}$$

with $\mathbf{r} = (x, y)$. Here, the basis size M and the box width of unity are used in both dimensions, though rescaling one dimension or using different basis sizes is straightforward.

In this chapter, I am only focusing on strictly 2D dots using the representation in Eq. (2.5). However, when computing J coupling strengths in Section 2.3, I will also present results from approximate 3D simulations conducted by my co-author, as detailed in our paper [76].

While it is possible to extend Eq. (2.5) to include the z-direction, this would require more computational resources. Since the z-dimension is small compared to the other two, one can choose to 'freeze out' the particle in its ground state in this dimension. This is equivalent to assuming that the gap between the ground and first excited z-state is too large to be relevant in our calculations [19, 96]. This approach modifies the Coulomb term in the Hamiltonian to account for the z-dimension's effects.

In practice, a 2D model can still be used by using an average Coulomb term, obtained by sampling the Coulomb interaction at different separations (in the z direction) between the particles according to the ground state's probability distribution. For more details on this generalisation derived by my co-author Simon Benjamin, see [76]. He has demonstrated that this approximation is valid as it yields similar exchange couplings to those obtained from full 3D simulations.

Finally, note that the spin degree of freedom can be easily added by having one entire charge state of the form Eq. (2.5) *for each* spin basis state. Similarly, the valley degree of freedom could be added in future studies. For multiple particles, the respective one-particle states are combined via the tensor product, although the rapid growth in RAM and processing resources limits the model to two electrons in practice.

The formalism I use closely follows Ref. [95], with some changes in notation and revisions. Interested readers may refer to that work for a more comprehensive survey of grid based methods.

Dynamics

To propagate a state in time, I use the split-operator (SO) method. In this approach, the total evolution time t is divided in N shorter time intervals $\delta t = t/N$ such that:

$$e^{-iHt} = \underbrace{e^{-iH\delta t} \cdots e^{-iH\delta t}}_{N \text{ times}}. \tag{2.6}$$

Since the Hamiltonian in Eq. (2.1) can be separated into a kinetic part and a potential part, each short time propagator $e^{-iH\delta t}$ can be approximated using Trotterisation,

$$\begin{aligned} e^{-iH\delta t} &= e^{-i(H_{kin}+H_{pot})\delta t} \\ &\approx e^{-iH_{pot}\delta t/2} e^{-iH_{kin}\delta t} e^{-iH_{pot}\delta t/2}. \end{aligned} \tag{2.7}$$

The error between the exact and the second-order Trotter approximation operators, computed using the spectral norm, is on the order of $O(\delta t^3)$. We find that it is sufficient for our purposes.

To obtain the time-evolved state $|\psi(t)\rangle$, I apply the operator $U_{SO}(\delta t) := e^{-iH\delta t}$ to the initial state $|\psi(0)\rangle$ N times. In the grid-based formalism, this can be done efficiently by alternating between the *real-space* and *k-space* representations. Here, due to my encoding choice, H_{kin} is diagonal in *k-space* while H_{pot} can be approximated as diagonal in *real-space* (see Appendix F of [76]).

2.3 Exchange coupling

In this section, I address the task of estimating the exchange coupling J of a DQD structure. Since J is directly related to the speed of two-qubit gates, any device that cannot generate a substantial exchange coupling should be disregarded.

The J coupling causes the spin singlet state to acquire phase with respect to all triplet states, effectively inducing a continuous SWAP operation between the two qubits. This effect does not come from any direct interactions (e.g. dipolar interaction) between spins but rather it is a consequence of fermionic particle exchange antisymmetry: the spin singlet and triplet states must be associated with a symmetric and antisymmetric spatial wavefunctions respectively. The energy gap between these states is highly sensitive to the size and shape of the confining double dot potential.

Specifically, for two electrons in a DQD, the exchange coupling is defined as the energy difference between the first two eigenstates of the spinless two-electron system, which have opposite exchange symmetry. These eigenstates can be obtained by diagonalising the Hamiltonian within our real-space basis. However, in the case of two electrons in 2D with M grid points per particle per direction (a total of M^4 basis states), diagonalisation becomes rapidly intractable. Nevertheless, by exploiting the distinct symmetries of the eigenstates with respect to electron swapping and employing a state-preparation technique known as imaginary-time evolution (see e.g. [97]), the grid-based method provides an efficient way of preparing them.

2.3.1 State preparation

Intuitively, for two electrons in a double quantum dot, the ground and first excited states should be close to a symmetric and antisymmetric combinations of finding one electron on the left and the other on the right respectively.

Starting from an initial guess state which has some overlap with the true ground state – such as a symmetric combination of two Gaussians centered in each dot – propagating the state using an imaginary time guarantees convergence to the true ground state $|s\rangle$ [97]. In practice, this is achieved by applying $U_{SO}(i\delta t)$ to our

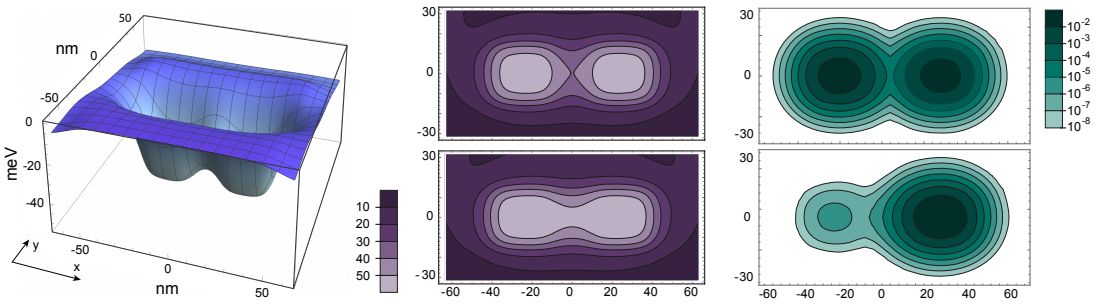

Figure 2.2: Potential well constituting the double dot (left and centre) and states of the confined electrons (right). The 3D visualisation and the upper contour plot show the 'barrier up' case, while the lower contour plot shows the 'barrier down' potential. In the right-most plots we show the probability densities on a logarithmic scale. The upper-right plot represents the ground state probability density for the 'barrier up' scenario, where the distribution is identical for either particle due to symmetry. Then we explore the subspace corresponding to "a given electron is definitely on the left of the structure" and plot the probability distribution for *the other electron*. This shows that with a small, but non-zero probability, the second particle can also be found on the left. In this case, we are in the $(2,0)$ charge configuration which does not constitute a legitimate qubit state. The total probability of finding the electrons in the $(2,0)$ and $(0,2)$ is equal to 0.04% here. This plot was made by co-author Simon Benjamin.

initial state until convergence is achieved. The same approach can also be used for preparing the first excited state, given that the ground and first excited states have distinct symmetries, and imaginary-time evolution preserves the symmetry of the starting state. Therefore, starting with an antisymmetric combination of Gaussians will lead to the antisymmetric state with the lowest energy, namely the first excited state $|a\rangle$. By computing the expectation values of $|s\rangle$ and $|a\rangle$ with respect to the Hamiltonian, I can estimate their respective energies and thus determine the exchange coupling by computing their difference. This method is completely general and can be applied to any DQD structure.

For realistic potentials, $|s\rangle$ and $|a\rangle$ will include a finite probability of finding both electrons in the same dot. However, when the barrier is large, we expect to be able to approximate them by,

$$|s\rangle \approx \frac{1}{\sqrt{2}}\left(|\mathrm{LR}\rangle + |\mathrm{RL}\rangle\right),$$
$$|a\rangle \approx \frac{1}{\sqrt{2}}\left(|\mathrm{LR}\rangle - |\mathrm{RL}\rangle\right), \tag{2.8}$$

with the probability of finding the electrons in the states $|\mathrm{LL}\rangle$ and $|\mathrm{RR}\rangle$ decreasing exponentially with the barrier height. Although this probability is expected to be small, it is important to characterise it since it represents states for which qubits are not defined – finding the system in such a state would correspond to irreversible qubit loss.

Throughout the paper, I will use the state $|\mathrm{L}\rangle$ and $|\mathrm{R}\rangle$ to denote the state of an electron that is definitely on the left or right dot, respectively. This notation will be particularly useful when defining the qubit states in Eq. (2.9). However, it is important to note that when approximating the symmetric and antisymmetric state as in Eq. (2.8), I am referring to specific charge wavefunctions (the eigenstates) localised on each dot.

2.3.2 Numerical results

I swept through a range of potentials from 'barrier up' to 'barrier down' (see Fig. 2.2) and considered various basis sizes. These potentials were obtained from a commercial Poisson-solver tool, which was applied to a certain device layout involving a central electrode (see Fig. 2.1 and Appendix A.1).

The results are presented in Fig. 2.3. In all cases, the exchange coupling J exhibits the general exponential dependence on the barrier gate voltage, as expected from analytic treatments (see e.g. [19]). Panel (a) shows the results for a strictly 2D dot. While all resolutions generally follow the expected behaviour, there is a deviation at low V_g (the 'barrier up' case) when using 16 basis states per particle per dimension. However, using 32^4 basis states adequately recovers the expected exponential variation over the full range of V_g considered.

The lower panel (b) compares the J versus V_g curves for 2D and 3D dots. The grey curves correspond to the three curves from panel (a); along with an additional 64^4 basis states simulation for the 2D dot. One observes that all these curves are relatively similar, with a J coupling at the 'barrier down' end of the range of approximately 150 MHz. For 3D dots, where the electrons are separated on average by 5 nm in the z direction, the coupling is stronger, reaching around 200 MHz. The

choice of the dots' depth appears to be the most important factor in this model, outweighing the differences in basis size or the specifics of the Coulomb matrix.

Intuitively, it is not surprising that the introduction of finite width significantly modifies J. The Coulomb interaction actually *reduces* the exchange coupling by forcing electrons further apart, reducing their amplitude in the central region, where the state's symmetry is most significant. By allowing the electrons to move in a third dimension, the Coulomb interaction is diminished, leading to an increase in exchange coupling [19].

Since the assignment of a mean-z depth makes the model specific to dots of that geometry, I will use the pure 2D Coulomb interaction (which was my work in this project) for the remainder of this chapter. While this is perhaps the least 'realistic' option, the phenomena I am studying remains unchanged even in the presence of a small finite z-depth. Using a 2D representation has no qualitative impact on the physics observed during the dynamics. The only meaningful change will be the speed of the gates.

2.4 Exchange gate dynamics

I now turn to the dynamics of the two-electron system, focusing on determining two-qubit gate fidelities and loss probabilities for various voltage profiles. But first, I need to define the qubits' states.

2.4.1 Qubit states

For two electrons in a DQD in the $(1,1)$ charge configuration, the qubit states are defined as follows,

$$\begin{aligned}
|00\rangle &\coloneqq \frac{1}{\sqrt{2}}\left(|\mathrm{LR}\rangle - |\mathrm{RL}\rangle\right)|{\downarrow\downarrow}\rangle\,,\\
|01\rangle &\coloneqq \frac{1}{\sqrt{2}}\left(|\mathrm{LR}\rangle\,|{\downarrow\uparrow}\rangle - |\mathrm{RL}\rangle\,|{\uparrow\downarrow}\rangle\right),\\
|10\rangle &\coloneqq \frac{1}{\sqrt{2}}\left(|\mathrm{LR}\rangle\,|{\uparrow\downarrow}\rangle - |\mathrm{RL}\rangle\,|{\downarrow\uparrow}\rangle\right),\\
|11\rangle &\coloneqq \frac{1}{\sqrt{2}}\left(|\mathrm{LR}\rangle - |\mathrm{RL}\rangle\right)|{\uparrow\uparrow}\rangle\,,
\end{aligned} \tag{2.9}$$

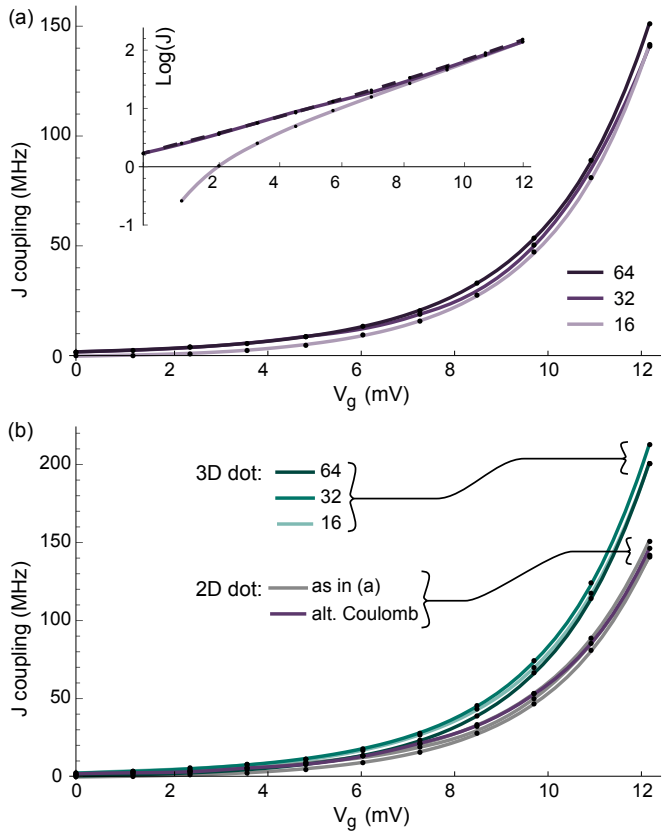

Figure 2.3: Variation of the exchange coupling J with the barrier gate voltage V_g. (a) A strictly 2D dot modelled with 16, 32 and 64 basis states per particle, per dimension. In the inset, curves for 32 and 64 are near-identical. (b) 2D versus 3D models: The curves from the upper figure, as well as a curve for a 2D dot using an alternative Coulomb matrix, are clearly distinct from models with a mean z-thickness of 5 nm. Figure made by my co-author Simon Benjamin. Both our codebases gave similar results.

where the arrows represent the spin states of the electrons as explained in Section 1.2.2. Note that the singlet $S(1,1)$ and triplet $T_0(1,1)$ states can be retrieved as follows,

$$S(1,1) = \frac{1}{\sqrt{2}}\left(|01\rangle - |10\rangle\right),$$
$$T_0(1,1) = \frac{1}{\sqrt{2}}\left(|01\rangle + |10\rangle\right). \tag{2.10}$$

In order to study two-qubit gates within the model, I need to be able to initialise the electrons in the qubit states defined in Eq. (2.9). However, it is important to keep in mind that the valid qubit states defined in Eq. (2.9) will not be sufficient to completely describe the state of a real double-dot structure. This is because the true eigenstates of the system, $|s\rangle$ and $|a\rangle$, are only imperfectly described using

the $|\mathrm{LR}\rangle$, $|\mathrm{RL}\rangle$ states (with the imperfect association as given in Eq. (2.8)). In reality, there is always some finite probability associated with the $(2,0)$ and $(0,2)$ charge configurations, i.e. $|\mathrm{LL}\rangle$, $|\mathrm{RR}\rangle$, which lie outside the qubit subspace and thus constitute qubit loss outcomes. With that said, the approximate association is

$$\begin{aligned}
|00\rangle &\approx |a\rangle\,|{\downarrow\downarrow}\rangle\,,\\
|01\rangle &\approx \frac{1}{2}\left((|s\rangle + |a\rangle)\,|{\downarrow\uparrow}\rangle - (|s\rangle - |a\rangle)\,|{\uparrow\downarrow}\rangle\right),\\
|10\rangle &\approx \frac{1}{2}\left((|s\rangle + |a\rangle)\,|{\uparrow\downarrow}\rangle - (|s\rangle - |a\rangle)\,|{\downarrow\uparrow}\rangle\right),\\
|11\rangle &\approx |a\rangle\,|{\uparrow\uparrow}\rangle\,,
\end{aligned} \tag{2.11}$$

with the approximation becoming better as the barrier gets larger. This encoding naturally enforces a coherent oscillation (with frequency J) between the states $|01\rangle$ and $|10\rangle$.

As noted earlier, adding the spin degree of freedom is done by taking the tensor product of the charge states defined in Eq. (2.5) with an array of length two encoding the spin state. The generalisation to two particles is straightforward. It is worth noting that as we are using a first-quantised representation, the antisymettry of the wavefunction is enforced in the states and is conserved through the evolution in time.

2.4.2 Numerical results

While multiple types of two-qubit gates arise from the exchange interaction [47], I focus here on the entangling $\sqrt{\mathrm{SWAP}}$ gate, which naturally appears when considering a homogeneous magnetic field. To simplify, I will then initially suppose that the g-factor in each dot is the same. In this scenario, the gate is implemented by pulsing the barrier gate voltage such that $\int_0^{T_{\sqrt{\mathrm{SWAP}}}} J(t)dt = \frac{\pi}{2}$. In Section 2.6, I will explore the impact of having different g-factors. For now, I focus on simple pulse sequences in which I smoothly ramp the potential from one with a large barrier (exchange off) to one with a small barrier (exchange on). It is worth noting that our tool can handle any evolution of the potential with time.

Starting from the state $|01\rangle$, as defined in Eq. (2.11), I perform a $\sqrt{\mathrm{SWAP}}$ operation by dynamically changing the barrier gate voltage according to the schedule

shown in Fig. 2.4. The rise and fall times are chosen to prevent non-adiabatic effects (as discussed later) while minimising gate time – ramping over the course of approximately one ns is consistent with current technological capabilities.

After simulating the dynamics, I evaluate the gate's performance. Finding that the state at the end of the gate satisfies $\text{Prob}(|01\rangle) = \text{Prob}(|10\rangle) \approx 1/2$ is of course insufficient to conclude that I performed a $\sqrt{\text{SWAP}}$ gate. A more nuanced assessment is thus required to measure the gate's quality. Suppose we have a general quantum state $|\psi\rangle$ encoded in real-space as follows:

$$|\psi\rangle = \sum_i |\Omega_i\rangle \left(a_i^{\downarrow\downarrow} |\downarrow\downarrow\rangle + a_i^{\downarrow\uparrow} |\downarrow\uparrow\rangle + a_i^{\uparrow\downarrow} |\uparrow\downarrow\rangle + a_i^{\uparrow\uparrow} |\uparrow\uparrow\rangle\right)$$

where the constants can have any values provided that overall antisymmetry is respected. The set of states $\{|\Omega_i\rangle\}$ represents a basis of two-particle charge states (e.g. the grid-based Dirichlet basis).

I now want to measure how close this state is to a given qubit state $|Q\rangle$ which would typically be written in the form,

$$\begin{aligned} |Q\rangle =& \, a_{00} |00\rangle + a_{01} |01\rangle + a_{10} |10\rangle + a_{11} |11\rangle \\ =& \frac{1}{\sqrt{2}} \Big(|\text{LR}\rangle \left(a_{00} |\downarrow\downarrow\rangle + a_{01} |\downarrow\uparrow\rangle + a_{10} |\uparrow\downarrow\rangle + a_{11} |\uparrow\uparrow\rangle\right) \\ & - |\text{RL}\rangle \left(a_{00} |\downarrow\downarrow\rangle + a_{01} |\uparrow\downarrow\rangle + a_{10} |\downarrow\uparrow\rangle + a_{11} |\uparrow\uparrow\rangle\right) \Big). \end{aligned} \tag{2.12}$$

In this target state, the $|\text{LR}\rangle$ notation indicates that electron 1 is definitely in the left side of the structure, while electron 2 is definitely in the right (and conversely for $|\text{RL}\rangle$). States such as $|\text{LL}\rangle$ or $|\text{RR}\rangle$, with both electrons on the same side, are not legitimate qubit states as explained before. To assess the match between $|\psi\rangle$ and $|Q\rangle$, we can start by projecting out any $|\Omega_i\rangle$ which correspond to the $(2,0)$ and $(0,2)$ configurations, recording the probability to be in this non-meaningful subspace as p_{loss}.

Now, I need to find the fidelity within the $(1,1)$ subspace. The exact shape of the charge distribution is not important. For example, if the a_i coefficients in

Eq. (2.12) are correct, then the fidelity is equal to one regardless of the specifics of the charge states $|\mathrm{LR}\rangle$ and $|\mathrm{RL}\rangle$. Therefore, we can write

$$F(|Q\rangle, |\psi\rangle) \coloneqq \sum_{\text{all } |\mathrm{LR}\rangle} |\langle Q|\psi\rangle|^2,$$

where I sum over all variants of the charge part of the target $|Q\rangle$, defined by Eq. (2.12) using a complete basis of states $|\mathrm{LR}\rangle$ (i.e. all the states with a $(1,1)$ charge distribution). Fortunately the calculation of this fidelity in the grid-based representation is straightforward.

Here, I chose the pulse such that $|Q\rangle = \frac{|01\rangle - i|10\rangle}{\sqrt{2}}$. Beginning with a state in the two lowest eigenstates, there is initially a small amplitude associated with the states having a $(2,0)$ or $(0,2)$ charge distribution. For the optimised gate, the final p_{loss} corresponds only to this initial imperfection, meaning the gate itself introduces *no additional loss*. Furthermore, the infidelity within the qubit subspace process is very small, with $1 - F = 1.32 \times 10^{-5}$. Thus, with the chosen voltage profile, a near-perfect $\sqrt{\mathrm{SWAP}}$ gate is achieved starting from $|01\rangle$ (and by symmetry, also when starting from $|10\rangle$).

As a comparison, if I abruptly switch the barrier (see Appendix A.2), p_{loss} increases dramatically, sometimes exceeding 1%. Even after accounting for this loss, the fidelity is limited to $F = 97.9\%$. This highlights the importance of a proper voltage switching profile to achieve near-adiabatic dynamics.

Whereas the ideal system can be adjusted to very high (near-perfect) gate operation, this does not account for imperfect environments. In the next section, I will discuss the impact of environmental factors such as charge noise.

2.5 Impact of charge noise on exchange gate

To adequately model the noise hindering the performance of silicon spin qubits, I first need to understand its nature. Several experimental groups have performed noise spectroscopy, revealing that it follows a $1/f$-like behaviour [42, 98–100]. Such type of noise is commonly associated with charge noise, and it is believed that it can be well-modelled using two-level fluctuators [87]. While various works have

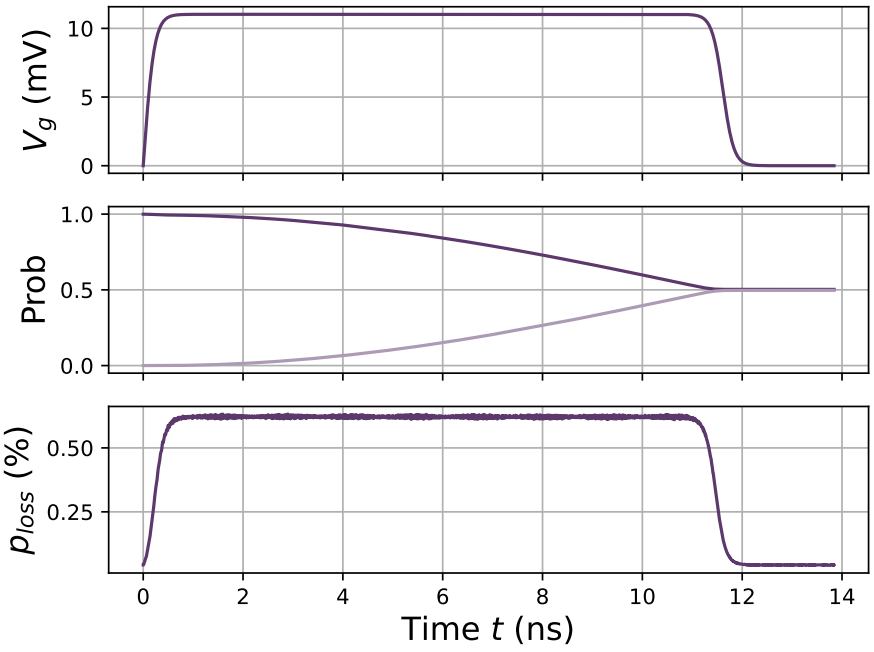

Figure 2.4: Smooth $\sqrt{\text{SWAP}}$ pulse. (top) Barrier gate voltage profile with V_g being the difference in potential between the bottom of the wells and the barrier at the midpoint ($x = y = 0$). I smoothly switch on the exchange interaction to minimise loss. (middle) Probability to find the electrons in the $|01\rangle$ (dark) and $|10\rangle$ (light) states. I observe a near perfect transfer to the desired superposition at the end of the process with an infidelity $1 - F = 1.32 \times 10^{-5}$. (bottom) Probability p_{loss} to find the electrons in the charge $(2, 0)$ or $(0, 2)$ configurations and hence outside of the qubit subspace. It is worth noting that for the initial state, I have a small probability $p_{loss} = 0.04\%$ to be in the $(2, 0)$ or $(0, 2)$ configurations. At the end of the process, I recover a practically identical value for p_{loss}, and conclude that our adiabatic pulse doesn't increase the loss.

considered baths of fluctuators [87, 101, 102], I complement those works by exploring the interesting case of a single fluctuator strongly coupled to the DQD [84, 86]. As shown in [86], considering the time it takes to perform a single $\sqrt{\text{SWAP}}$ gate, one expects that a charge defect will remain fixed. Hence, I will first focus on static charges before addressing the rarer, though interesting, case of fluctuating charges.

2.5.1 Static charge

The trapped charge is modelled classically by adding an extra coulomb interaction term to Eq. (2.1) with respect to each electron in the DQD,

$$H_{sc,i}(\mathbf{r}) = \frac{e^2}{4\pi\epsilon_0\epsilon_r|\mathbf{r} - \mathbf{r_i}|^2}, \tag{2.13}$$

with $\mathbf{r} = (x, y, z)$ and $\mathbf{r_i}$ the position of the charge and the position of the ith electron respectively. The charge is randomly placed at the silicon oxide interface $z \approx 2$ nm above the DQD and in the symmetry axis of the double dot (the x axis, i.e. $y = 0$). By placing the charge along this axis, I explore situations in which the charge has the most impact. In Appendix A.3, I explore the more general setting of off-axis charge defects.

When a single charge is located near the DQD, it displaces the electrons, altering the overlap between their wavefunctions and thereby modifying the exchange coupling. This effect increases as the charge gets closer to the DQD (see Table 2.1). However, this trend breaks down when the charge is between the two electrons. I assume that the gate has been calibrated in the absence of a charge defect, meaning that I am unaware of it and employ the same pulse as in the noiseless case to perform the gate. Since the exchange coupling is modified by the presence of the charge, the pulse no longer yields a high-quality gate.

In Fig. 2.5, I report the evolution of the qubits' state and p_{loss} for a charge positioned at $x = 50$ nm (purple) and $x = 100$ nm (green) with respect to the center of the double dot structure ($x = y = 0$). Assuming the system has previously relaxed into the eigenstates of the modified potentials, the initial state $|01\rangle$ now differs from that computed earlier, and now has a larger initial value of p_{loss}. Although the control pulse still allows for adiabatic evolution, I find that the fidelity decreases drastically as the charge moves closer to the DQD, because of the timing error. In Table 2.1, I report the evolution of F and p_{loss} for various charge positions.

2.5.2 Fluctuating charge

Although most charge defects are likely static during a $\sqrt{\text{SWAP}}$ operation, it is important to study the impact of a charge that moves during a gate operation. While a vast range of interesting scenarios can be studied by the methods I use here, such as a charge moving randomly in 3D, I will focus on the case where a charge fluctuating instantaneously between two traps.

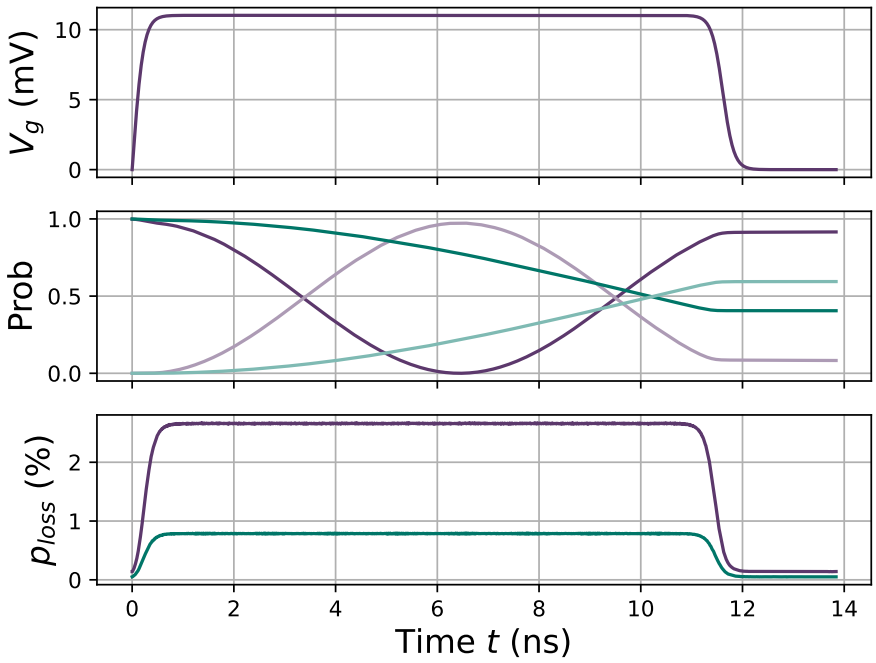

Figure 2.5: Smooth $\sqrt{\text{SWAP}}$ pulse in the presence of charge noise. (top) As I assume the operator does not know that the defect charge is present, I use the evolution of the barrier gate that works optimally in the defect-free case (i.e. upper the panel of this figure is the same as in Fig. 2.4). (middle) Probability to find the electrons in the $|01\rangle$ (dark) and $|10\rangle$ (light) for a charge positioned at $x = 50$ (purple) and $x = 100$ (green) nm respectively. The presence of charge noise generates an error in the $\sqrt{\text{SWAP}}$ gate as it modifies J. There is also an increased value for p_{loss} during the gate operation, as compared to the noiseless gate (Fig. 2.4), but at the end of the gate it returns to the initial value and thus the operation does not *increase* the loss – it is still effectively adiabatic.

Charge position (nm)	50	75	100	150	200
J (MHz)	507	202	168	155	153
p_{loss} (%)	0.14	0.06	0.05	0.05	0.04
F (%)	22.41	92.11	99.10	99.95	99.99

Table 2.1: Exchange coupling, loss probability and state-to-state fidelity as a function of the position of the charge. The charge position (in the dot line) is given as a distance from the center of the DQD which is positioned at (0, 0, 0). As the exchange coupling is a static property, it has been computed with the 32^4 resolution. However, as the loss and fidelities are obtained after simulating the dynamics, I only used the 16^4 resolution which I do not expect to change much (see Fig. 2.3)

.

I approach this in two ways. First, I consider a charge oscillating between its initial position, $\mathbf{r}_{init} = (100, 0, 2)$ nm, and a second location, $\mathbf{r} = (x, 0, 2)$ nm, for various values of x. I expect that the impact will be greater as the two positions are far from one another. I assume that the experiment has been calibrated so that the barrier gate voltage profile is optimised for a $\sqrt{\text{SWAP}}$ gate in the presence of a charge at $\mathbf{r}_{init}$. In the second approach, discussed in the context of robust pulses, I consider fluctuators with a fixed separation of 10 nm but various mean locations.

In Fig. 2.6, I explore the effects of a charge oscillating between $\mathbf{r}_{init}$ and $\mathbf{r} = (75, 0, 2)$ nm. I randomly choose the times (indicated by dashed green lines) at which the charge moves, limiting the number of fluctuations to a maximum of three (see Appendix A.4 for other examples). I set the limit to three as these scenarios become less and less probable as the number of fluctuations increases.

It is interesting to note that instantaneously changing the position of the charge during the gate produces a similar effect to an overly rapid switching of the barrier voltage (see Appendix A.2). The sudden (non-adiabatic) nature of the change is detrimental. However, in this example, the charge does not move too close to the DQD, resulting in only a small change in potential and thus a relatively minor impact from the fluctuation (see the last row of Appendix A.4 for more drastic effects). Regarding p_{loss}, note that even after the end of the gate, we still have some oscillations. By averaging over the time following the gate operation, I find that $p_{loss} = 0.07\%$ and $F = 99.71\%$. Unlike previous cases, this introduces an increase in p_{loss}, which is the worst type of error as it cannot be corrected. A fluctuating charge is thus more damaging than a static one when it comes to gate quality.

Having identified how charge noise affects gate performance, I now turn to the vast literature on robust pulse sequences to explore strategies for mitigating these errors.

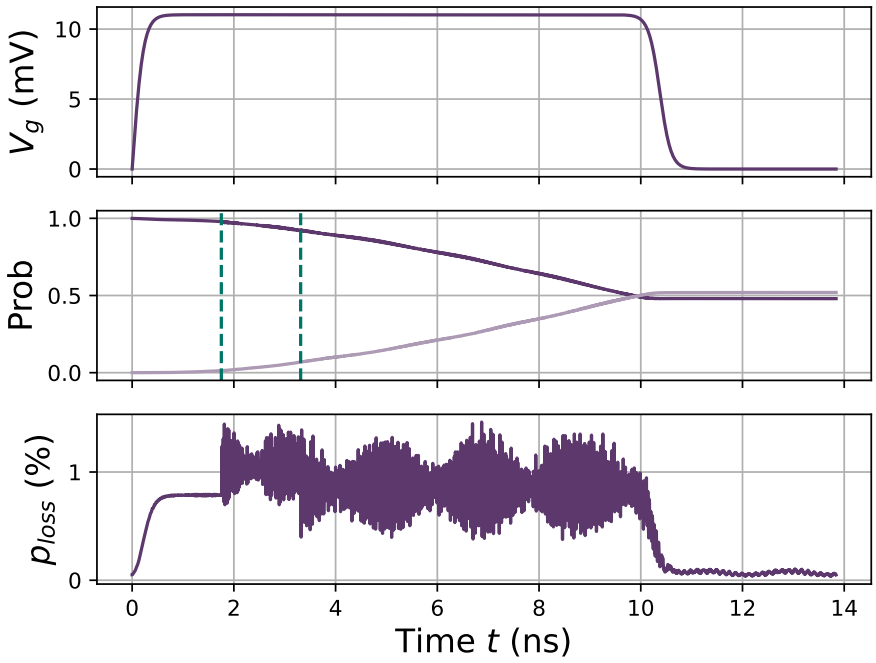

Figure 2.6: Smooth $\sqrt{\text{SWAP}}$ pulse in the presence of a fluctuating charge. The charge oscillates between its initial position $\mathbf{r}_{init} = (100, 0, 2)$ nm and $\mathbf{r} = (75, 0, 2)$ nm at each green dashed lines.

2.6 Robust pulse sequences

2.6.1 Types of errors

When performing a gate, three types of imperfection can arise. Two of these have already been discussed in the previous section. The first is the loss caused by driving electrons into the $(0, 2)$ and $(2, 0)$ configurations, taking them outside of the qubit subspace when modifying the potential. The second arises from charge noise, which alters the exchange coupling J between the electrons, leading to uncertainty in the gate duration and reducing the fidelity of the operation. The third imperfection, which I have not yet tackled, involves magnetic detuning between dots.

If the g-factors between the dots are different, the qubits within the double dot will not experience a pure SWAP-type operation. To understand this, imagine a Bloch sphere for the qubit subspace spanned by $|01\rangle$ and $|10\rangle$ (as defined in Eq. (2.9)). In the presence of different g-factors, the qubit state will rotate around a tilted axis rather than performing an on-axis rotation. For simplicity, I assume that

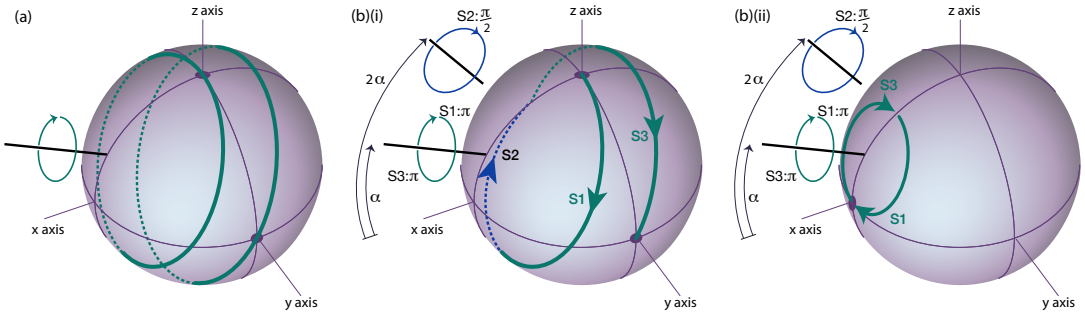

Figure 2.7: (a) Illustration of the off-axis nature of the 'natural' gate when the two dots have distinct g-factors: a pure x-rotation is not achieved, meaning that the two states marked with purple dots do not evolve directly into each other. (b) Illustration of the basic 'ABA' pulse sequence which corrects the off-axis rotation error. The first and last of the three steps (marked green) correspond to the same rotation, by π about an axis making angle α with respect to the x-axis. The middle step (marked blue) is a $\pi/2$ rotation about an axis which makes an angle 2α with respect to the x-axis. In (b)(i), one sees that the state $|01\rangle$ at the north pole undergoes a net rotation to $(|01\rangle + i\,|10\rangle)/\sqrt{2}$ as desired (the two filled purple dots mark these states). In (b)(ii), the same sequence maps the state marked by the purple dot onto itself, only acquiring phase, since the first pulse leaves it an eigenstate of the second pulse (i.e. on that pulse's axis). Together, b(i) and b(ii) show that the sequence synthesises the desired ideal rotation about the x-axis. Note that the angle α is exaggerated in this schematic compared to its actual value in the pulse sequence. This figure was made by my co-author Simon Benjamin.

the g-factor is uniform across each half of the double dot structure, with an abrupt change at the center. This is a simplification, but it is important to note that the numerical methods described here can handle any spatially varying g-factor.

In the case of a well-characterised environment, the first type of error (qubit loss) can be reduced by using an adiabatic pulse. However, as I observed previously, this error cannot be corrected in the presence of unpredictable charge fluctuations. Nonetheless, as I will explore below, the other types of errors can be mitigated even in the presence of unknown charge environments.

It is also worth observing that even though a two-qubit gate is off-axis in the sense described above, it can still be used as a computationally universal gate, given suitable single-qubit rotations. One could thus argue that this natural rotation is acceptable once characterised, and does not require correction at the pulse level. However, in a device containing thousands or millions of such structures, the ability to homogenise their action is likely to be valuable.

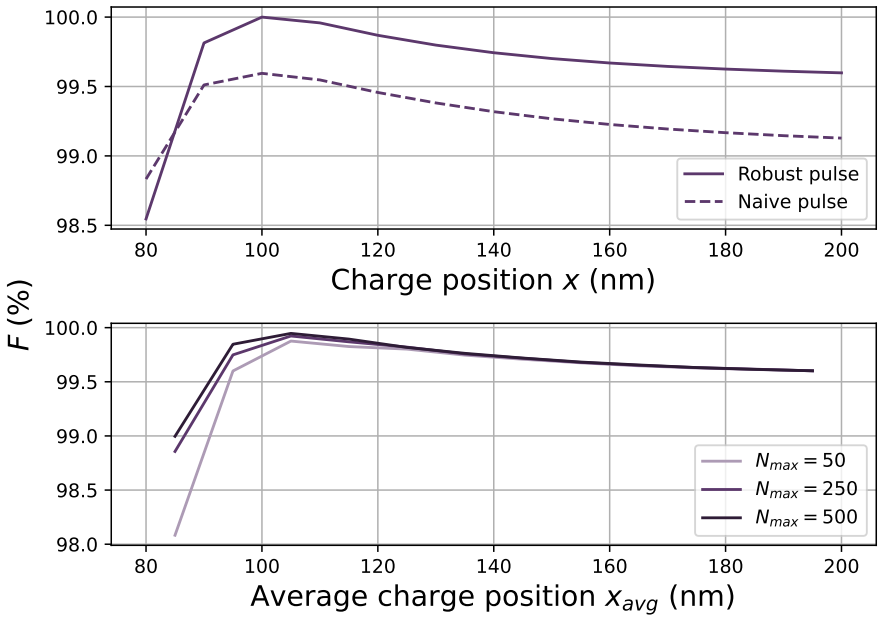

Figure 2.8: Evolution of the fidelity when using the robust pulse sequence for a static (top) and fluctuating charge (bottom). For the fluctuating case, the charge oscillates between two traps located at a fixed 10 nm distance; the horizontal axis in the graph is the mean charge position. The number of fluctuations during the pulse is drawn uniformly at random between $N = 1$ and $N = N_{max}$. Additionally, the switching times are randomly chosen between $t = 0$ and $t = T_{\sqrt{\text{SWAP}}}$. Each curve is obtained by averaging 10^4 instances of such random environment. For the naive case, I used a maximal J coupling equal to the strongest one appearing in the robust pulse sequence.

2.6.2 Robust pulse sequence

Robust or composite pulse sequences have been developed and implemented over several decades by the NMR community to mitigate the effect of experimental imperfections [88]. These techniques can also be applied in quantum computation to reduce the last two types of errors mentioned above (see e.g. Refs. [89, 90]). Composite pulses fall into two categories: those that are efficient for a subset of initial states and those which are error-tolerant for any initial state (also known as fully compensating pulses). In this work, I focus on a fully compensating pulse designed to implement a $\sqrt{\text{SWAP}}$ gate.

This type of pulse (or pulse sequence) typically takes longer to apply than a naive one. As a result, grid-based methods can become impractical for searching for robust pulse sequences, since the large number of time steps required to propagate the full wavefunction must be repeated across various charge defect scenarios. Fortunately,

I can restrict the use of full grid-based model to characterising p_{loss} and extracting key parameters, which are then used to construct a simpler model within a smaller subspace. In this case, the relevant subspace is spanned by the states $|01\rangle$ and $|10\rangle$. The time-dependent Hamiltonian in this subspace is given by,

$$H(t) = J(V_g(t))\sigma_x + \Delta E_z \sigma_z, \tag{2.14}$$

with the dependence of J on V_g having been extracted from the grid-based method. Here ΔE_z is the difference in Zeeman energy between the two dots due to their difference in g-factor and σ_z, σ_x are the Pauli Z and Pauli X operators respectively. I set $\Delta E_z = J_{on}/20$ (with J_{on} the maximal exchange value obtained when switching the interaction on) which is a typical value for the magnetic detuning. Since ΔE_z is assumed to be fixed and non-zero, a pure x-rotation cannot be directly achieved for any value of the controllable parameter J. However, different values of J result in rotations around different axes within the relevant quadrant of the x-z plane (see Fig. 2.7), providing the flexibility to construct interesting sequences.

The ideal composite pulse would implement a high fidelity on-axis rotation (around the x-axis), despite both the off-axis error and environmental charge noise. The goal is to develop a fully-compensating sequence that does so for *any* input state, i.e. a good match to the ideal unitary.

We explored several potential sequences and report the performance of one successful option in Fig. 2.8, a relatively complex sequence with 9 steps. When creating this sequence, my co-author's approach was to begin with a more simple sequence of the form ABA known to correct off-axis errors [91], which was confirmed to perform well in the absence of charge noise (see Fig. 2.7). Then, he divided this sequence into two equal parts as $(A\frac{B}{2})(\frac{B}{2}A)$, and looked for additional pulses that could be inserted in the middle to compensate for charge noise effects. To address this, he adapted another solution employed in the NMR literature [91] by introducing a π-rotation about an axis approximately perpendicular to the x-z plane (requiring two pulses in our case), followed by a full 2π rotation about an

axis in the x-z plane, and then reversing the first π rotation. The effect of this sequence is to perform a rotation of opposite sign to that which occurs in the original ABA sequence. The charge-induced errors which occur 'naturally' are thus largely compensated by those which we 'deliberately' accumulate.

This sequence works for different charge scenarios. Moreover, it can be further tuned to optimise the performance for impurities fixed at a given distance. This comes at the cost of not being able to implement a perfect gate in the absence of charge defects. This is the approach taken for the sequence whose performance is shown in Fig. (2.8).

While this pulse sequence improves gate fidelity, it significantly increases the total gate time. Compared to the naive pulse which implements a $\sqrt{\text{SWAP}}$ gate in $T_{\sqrt{\text{SWAP}}} \approx 12$ ns in Fig. 2.4, the robust pulse sequence results in a gate time $T_{\sqrt{\text{SWAP}}} \approx 1.9$ µs. We emphasise that we did not make a systematic search for more compact sequences, and this is an interesting challenge for future work.

2.6.3 Numerical results

As I am interested in fully-compensating pulses, the measure of fidelity should focus on comparing the implemented unitary to the target one, rather than specific states. In this section, the fidelity F is defined as [103],

$$F(U, U_{target}) \coloneqq \frac{d + \text{Tr}\left(U^\dagger U_{target}\right)^2}{d(d+1)}, \tag{2.15}$$

with $U_{target} = U_{\sqrt{\text{SWAP}}}$ restricted to the subspace spanned by the states $\{|01\rangle, |10\rangle\}$ and $d = 2$.

At the top of Fig. 2.8, I plot the evolution of F for different positions of a static charge impurity using both the naive and robust pulses. In this section, I use a maximal J coupling equal to the strongest one appearing in the robust pulse sequence for the naive pulse. By tuning the robust pulse sequence such that the maximal fidelity is obtained when $x = 100$ nm, the fidelity is increased over a large range of positions. However, the pulse cannot generate a perfect $\sqrt{\text{SWAP}}$ gate in the limit of an infinitely remote impurity (or no impurity). As stated earlier, as the charge

moves closer to the DQD, the fidelity eventually drops significantly. The worst case corresponds to having a charge directly above the DQD, which would render the dot structure temporarily non-functional. A finer impurity model would be necessary to study these scenarios meaningfully, likely treating the impurity as a full particle.

Given that the pulse sequence is two orders of magnitude slower than the naive one, we might expect to have significant fluctuations during the process. It is important to note that by switching from the grid-based method to this lower-dimension model, I lost the ability to track the loss probability. However, for the positions of the charge considered here, p_{loss} remains small and is not drastically affected by charge fluctuations, as reported in Fig. A.3.

At the bottom of Fig. 2.8, I plot the evolution of F for different average charge positions x_{avg} along the dot line, which is the worst-case scenario. An average position x_{avg} corresponds to a charge oscillating between two traps positioned at a fixed distance of 10 nm around x_{avg}. For instance, for $x_{avg} = 95$ nm, the charge oscillates between $x = 90$ and $x = 100$ nm. Since the number of fluctuations N and their timings $\{t_i\}_{1\leq i\leq n}$ are unknown, both are randomly generated such that N is uniformly selected between $N = 1$ and $N = N_{max}$ and the t_i are drawn randomly between $t = 0$ and $t = T_{\sqrt{\text{SWAP}}}$. Each curve is the result of averaging over 10^4 instances of such random environments.

This plot reveals two distinct regimes. When the charge's average position is far from the DQD, the number of fluctuations has negligible impact. This aligns with the trend observed in Table 2.1 – the farther the charge is from the dot, the smaller the change in the exchange coupling. Hence, fluctuating between two positions around x_{avg} that yield similar J couplings leads to a fidelity similar to that of a fixed charge at x_{avg}, regardless of the number of fluctuations. This also explains the similar behavior to the static case (solid purple line on the top of Fig. 2.8). On the other hand, when the charge oscillates closer to the dot, the impact on J between the two traps is larger, lowering the fidelity. However, as the number of fluctuations increases, the fidelity improves and tends to converge toward that of the static case.

In both static and fluctuating charge scenarios, the pulse sequence demonstrates improved fidelity over a wide range of positions, thus accommodating for various charge environments. This shows that the modelling approach used here can effectively search for and validate robust pulses, potentially allowing for the discovery of sequences which yield high-fidelity gates.

2.7 Future directions and conclusion

In this chapter, I present a versatile modelling approach for simulating electrons and their spins in double quantum dots. Using realistic potentials, the tools employed can identify promising devices by computing the exchange coupling, an important metric in the realisation of two-qubit gates. Moreover, these tools can fully simulate qubit dynamics, allowing for the study of different voltage control pulses to enhance gate fidelities while reducing the risk of losing qubits.

Using a two-level fluctuator model, I explored how a single trapped charge near the DQD affects gate quality. For a static charge, the fidelity can be heavily impacted, though the loss remains constant when a suitable adiabatic control pulse is used. However, for a randomly oscillating charge, both fidelity and loss are affected. While the impact on the fidelity can be mitigated using robust pulse sequences (as shown using a reduced model), the effect on the loss cannot be removed for unknown environments.

The grid-based methods used here offer many possibilities for future work. One could think first about a complete treatment of the third dimension. However, as our method is computationally demanding, optimising the code is necessary. My co-author, Simon Benjamin is currently working on that. Additionally, expanding the model to include spin-orbit interaction, valley degrees of freedom, as well as valley-orbit coupling terms could be valuable. Finally, the introduction of interactions between two-level fluctuators or treating them as distinct quantum particles opens further research possibilities.

In conclusion, the approach and tools developed in this work set the stage for a software pipeline that takes a chip layout (defined in software used to design

traditional chips) and outputs two-qubit gate performances under realistic noisy environments. While this work does not explore single-qubit gates or other types of two-qubit gates, extending the methods to study these would be straightforward.

3

Harnessing electron transport for global spin qubit control

This chapter presents early stage research on global control for spin qubits. The project originated from discussions between Adam Siegel, Dr. Fernando Gonzalez-Zalba and myself. While I came up with the idea of homogenising frequencies using the exchange coupling or shuttling presented in Section 3.3 and in Section 3.4 respectively, the final protocol implemented in Section 3.4 was Adam's idea. We developed the code base collaboratively, with Adam focusing on implementing the time-dynamics simulations and me focusing on the clustering heuristics. After submitting this thesis, my co-authors and I finalised our research, and the complete results are now available in a preprint [104].

Contents

3.1 Introduction 45
3.2 Preliminaries 46
3.2.1 Theoretical model 46
3.2.2 Problem statement 47
3.2.3 Clustering 48
3.2.4 Workflow 50
3.2.5 Metrics 51
3.3 Effective frequencies using SWAP gates 52
3.3.1 Two-qubit case 53
3.3.2 n-qubit case 55
3.3.3 Performance on a 2×N array 57
3.4 Effective frequencies using shuttling 59
3.4.1 Single-qubit case 59
3.4.2 n-qubit case 61
3.4.3 Performance on a 2×N array 62
3.5 Discussion 65

3.1 Introduction

In the previous chapter, I explored the basic physics of spin qubits and developed a modelling platform allowing for the design of devices that could lead to high-fidelity two-qubit gates. However, similar to current experiments [44–46, 105], I have only shown the reliable implementation of this basic operation for a limited number of qubits. For spin qubits to represent a viable option for building scalable quantum computers, methods for controlling a large number of qubits – potentially millions to implement error correction [8, 10] – need to be developed.

A single qubit gate is performed by controlling the spin of an electron using a pulse in resonance with the qubit's frequency, as explained in Chapter 1. Single-spin control can be achieved through EDSR, though this would require engineering micromagnets on top of all the dots. Alternatively, a nearby current can be used to create an oscillating magnetic field to perform ESR, but localising this magnetic field to a single electron is challenging. Additionally, as the number of qubits increases, the number of wires needed to generate these currents can become problematic.

One solution to the wiring problem is to use global (i.e. targeting all qubits) magnetic field for ESR. In this case, specific qubits can be targeted by modifying the field's frequency as explained in Chapter 1. However, each spin qubit's frequency is determined by its environment, particularly by the heterostructure's interface roughness above the dots [106]. As the number of qubit grows, frequency crowding – where two or more distinct qubits have similar frequencies – can lead to crosstalk. This means that a non-target qubit can be unintendedly modified while manipulating target qubits. One way to circumvent this issue is by using the The Stark shift, which alters an electron's g-factor (and thus its frequency) via a change of the corresponding gate voltage [48, 63]. Nevertheless, the extent to which the g-factor can be modified is often insufficient to fully suppress crosstalk.

Previous works have tried to tackle this issue. Ref. [107] demonstrated that the difference in frequency between spins can be reduced by carefully selecting the angle of the static field $\vec{B}_0$ defined in Eq. (1.14). Stark shift can also be used to place qubits into frequency bins to limit the impact of frequency crowding. As such,

the minimal spacing between two frequencies does not evolve with the system size [108]. In this approach, a drive with multiple tones corresponding to each frequency bin is used for control. This work was then extended in [109] to be able to target a single qubit rather than all of them. However, none of these works allow for the reliable implementation of single-qubit gates on an arbitrary subset of qubits.

In this chapter, I present two protocols based on electron transport that allows one to do exactly that. Both methods require a global magnetic field and local electrical control only, and respectively rely on the ability to modify the exchange coupling between dots or to shuttle electrons. Note that we developed a third protocol based on the modification of the detuning between dots, but I will not present it here as I did not directly contribute to it.

The chapter is divided as follows. First, in Section 3.2, I introduce the model we used, as well as several technicalities needed to easily understand our schemes. Then, in Section 3.3, I go into the details of how one can create effective qubits using the exchange coupling between dots, finding that this method is limited and will not scale. In Section 3.4, I show how electron shuttling can be leveraged to perform high-fidelity single-qubit operations on any given set of qubits. Finally, I conclude on the use of transport for performing global single-qubit gate and discuss potential extensions of our work in Section 3.5.

3.2 Preliminaries

3.2.1 Theoretical model

Supposing thereafter that $\hbar = \mu_B = 1$ the Hamiltonian of a system composed of n electrons in $N = n$ quantum dots reads,

$$H = \sum_{i=1}^{N} g_i \vec{\sigma}_i \cdot \vec{B} + \sum_{i=1}^{N-1} J_i \vec{\sigma}_i \cdot \vec{\sigma}_{i+1}, \tag{3.1}$$

with J_i the exchange coupling between neighbouring quantum dots, $\vec{\sigma}_i = (\sigma_i^x, \sigma_i^y, \sigma_i^z)^T$ single-qubit Pauli matrices, g_i each electron's g-factor and $\vec{B} = (B_x, 0, B_z)$ a global magnetic field. The magnetic field has both a DC component $B_z = B_0$ (used for splitting the energies of the spin states $|\uparrow\rangle$, $|\downarrow\rangle$) and an AC component

$B_x = B_1 \cos(\omega t)$ (used for controlling the spins). In practice, B_0 ranges from 0.1 to 1 T and $\Omega_i \approx 1$ MHz. Using values from the lowest end of this range is beneficial for increased qubit uniformity (thus reducing off-resonant errors, as predicted by the Rabi model). However, it makes the addressing of qubits more challenging as their frequency difference increases with B_0. In this chapter, where I require a specific B_0 for a calculation, I will suppose that $B_0 = 0.1$ T, showing that the techniques we developed are powerful even in this stringent regime. Exchange coupling J of tens of MHz have been reported in experiments [44–46] and a recent experiment showed a J on the order of 100 MHz [110]. Theoretical studies showed that even larger J could be realistic as well [76, 84, 86].

For electrons in silicon heterostructures, the g-factors depend on the interface roughness. Exactly modelling this dependence remains a challenge as it would require simulations of interface physics. Usually, this dependence is simply modelled by randomly assigning to each qubit a g-factor following a normal distribution $\mathcal{N}(g_0, \Delta g)$, where $g_0 = 2$ and Δg ranges from 10^{-3} to 10^{-2} [111]. Note that these random g-factors can further be adjusted using the Stark shift by an amount $\delta g \approx 0.1\Delta g$ [48, 112, 113]. Here, we will choose $\Delta g = 10^{-3}$ and $\delta g = 10^{-4}$ to be in the worst case.

With $B_0 = 0.1$ T, I report in Table 3.1 the values of the different parameters of interest. I also added the time resolution of the control electronics and the shuttling speed $v = 10$ m/s (leading to a shuttling frequency $v/d = 100$ MHz for an interdot spacing $d = 100$ nm) in order to understand the different frequency scales used in this chapter.

3.2.2 Problem statement

Now that I have specified the model, I can clearly state the problem we tried to solve. Given n_q qubits, how one can perform an X gate on n_t target qubits while applying an I gate to the $n_q - n_t$ remaining qubits with the highest fidelity.

An X rotation is performed on the ith qubit by choosing a global field B_x such that the drive frequency ω_i matches that of the qubit $\omega_{q,i} = g_i B_0$ as explained in

Parameter	Symbol	Order of magnitude
Time resolution	τ	10 GHz
Qubit frequency	ω_q	1 GHz
Shuttling speed	v	100 MHz
Exchange coupling	J	$10 - 100$ MHz
Average Rabi frequency	Ω	1 MHz
Average qubit splitting	$\Delta\omega_q$	1 MHz
Stark-shift tunability	$\delta\omega_q$	0.1 MHz

Table 3.1: Relevant parameters and corresponding frequency values.

Chapter 1. The ith qubit will thus rotate at the Rabi frequency $\Omega_i = g_i B_1$. This means that each target will have slightly different gate times $T_i = \pi/\Omega_i$ which might be problematic. Suppose that instead, we send the global pulse for a duration $T = \pi/\Omega$ with $\Omega = g_0 B_1$ obtained using the average g-factor. Each qubit will thus be under- or over- rotated by an angle $(\Omega_i - \Omega)T = \pi\Delta g/g_0 = O(10^{-3})$. The fidelity of the operation is linked to the Rabi formula presented in Eq. (1.22). This formula allows us to conclude that the infidelity due to this timing mismatch will be on the order of $(\pi\Delta g/g_0)^2 = O(10^{-6})$ which is negligible. In conclusion, the most important thing to do is to increase the separation between qubit frequencies to reduce crosstalk which is the goal of the two methods presented in this chapter.

3.2.3 Clustering

The techniques presented in this chapter rely on fast qubit transport to homogenise the g-factors of electrons. However, one might not want to homogenise the frequencies of every qubit in the device. For instance, it would be desirable to have different frequencies for the target and non-target qubits. What would be interesting is to homogenise the frequencies of a given group or *cluster* of electrons. This idea of clustering will be key in both methods.

Suppose you have a cluster of qubits, each of which has a different g-factor. Using either the exchange interaction (see Section 3.3) or shuttling (see Section 3.4), we can arrange that every qubit of the cluster will explore every dot such that their

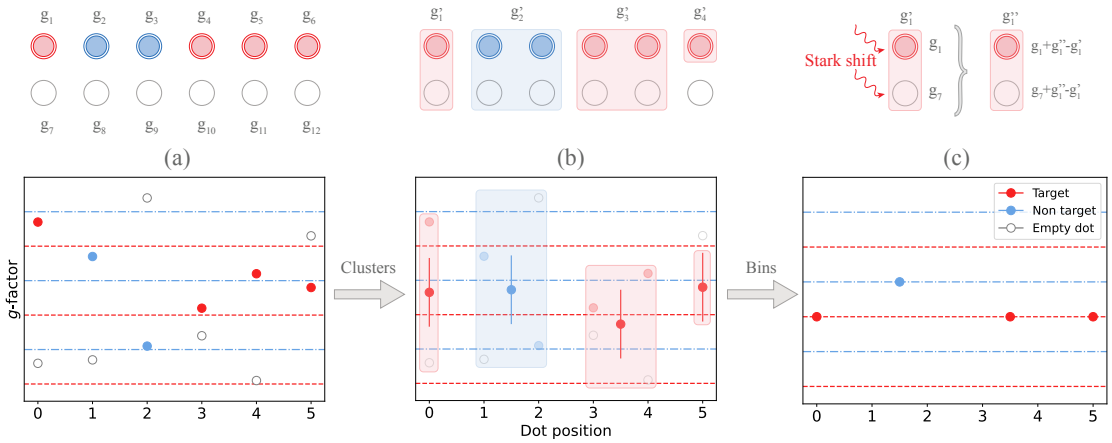

Figure 3.1: Schematic of the clustering and binning workflow. In the upper panels, I show the 2×N architecture I used to test our methods. Note that empty dots (grey circles) are used here to allow for shuttling in loops as detailed in Section 3.4. In the lower panels, I show the corresponding g-factor distributions. Red and blue disks correspond to the target and non-target qubits respectively. Red and blue dashed lines correspond to target and non-target frequency bins. (a) Initial configuration. (b) Clustering step. Every qubit within a similar cluster, represented as a box, obtains the same g-factor $g_i^{'}$. (c) Binning. Using Stark shifts for all the physical qubits, I can shift the cluster's frequency from $g_i^{'}$ to $g_i^{''}$ and put them in frequency bins. In this example, we went from a spectrum containing six distinct frequencies to only two. Using only binning would have led to four bins being populated. This figure was made by my co-author Adam Siegel.

g-factor gets homogenised to the mean g-factor of the dots composing the cluster. Later in the chapter, I verify this simple intuition through numerical simulations.

Protocols based on homogenisation have two consequences. First, they allow one to reduce frequency crowding as the number of different frequencies will be reduced to the number of clusters. Second, they allow for an increased g-factor tunability. For instance, suppose you have two qubits with respective g-factors equal to $g_0 - \Delta g$ and $g_0 + \Delta g$. Using our protocols makes it possible to modify their g-factor to g_0, which represents a tunability of Δg which is an order of magnitude larger that the one offered by the Stark shift which is equal to δg.

Clustering represents thus a key aspect of our protocols, and is highly dependent on the device structure. Finding the best clustering is computationally demanding as it requires comparing all possible subsets of a given set. To address this, we developed tailored heuristics for both protocols in order to find suitable clusterings. Fig. 3.1(b) illustrates the effect of clustering.

3.2.4 Workflow

Each section in this chapter will present a new global control technique and will have a similar breakdown. First, I introduce the scheme and test it on a small number of qubits. Then, I test it on a realistic device. Here, I focus on a 2×N quantum dot architecture. The choice of this architecture is motivated by the fact that it could be built in the near term and can allow for the implementation of error correction [114]. To test our techniques, I perform $N_{runs} = 10^4$ Monte-Carlo simulations of the following experiment:

1. Set the experimental parameters: B_0, B_1, Δg, δg
2. Set the total number of qubits n_q and select n_t target qubits uniformly at random
3. Generate the qubits' g-factors according to a normal distribution $\mathcal{N}(g_0, \delta g)$ (Fig. 3.1(a))
4. Use one of our schemes presented in Section 3.3 and Section 3.4 for the relevant qubit clusters to reduce frequency crowding (Fig. 3.1(b)). Here, every qubit within a cluster will acquire a similar effective g-factor $g^{'}$ equal to the average of the ones composing the cluster.
5. Further reduce frequency crowding by tuning the effective g-factors. This can be done using Stark shift and putting the effective g-factors into bins (Fig. 3.1(c)). See Appendix B.1 for details on binning. The output of this step is a new set of achievable g-factors $g^{''}$ from which we can compute the drive frequencies ω_i. To make things clear, to get from the initial g-factors g_i to the final ones $g_i^{''}$ we perform two operations: (i) create the clusters (by turning on the exchange couplings or shuttling the electrons) and (ii) modify the initial g-factors g_i by the amount $g_i^{''} - g_i^{'}$.
6. Compute the fidelity of $U_{target} = X^{\otimes n_t} \otimes I^{n_q - n_t}$ by performing the full Hamiltonian simulation for a time $T = \pi/\Omega$. The fidelity is computed using Eq. (3.2) presented in the next section.

3.2.5 Metrics

In this chapter, I will need to characterise the fidelity of different operations. In this subsection, I define different fidelity metrics for clarity purposes.

Global control fidelity

To evaluate the quality of our techniques, I fully simulate the dynamics of the system and compare the implemented unitary U to the target one $U_{target} = X^{\otimes n_t} \otimes I^{n_q - n_t}$ using the following matrix norm:

$$\mathcal{F} = \frac{1}{4} \left| \mathrm{Tr}(U_{\mathrm{target}}^{\dagger} U) \right|^{2/n_q} . \tag{3.2}$$

We took the n_q-th root to obtain an average single-qubit fidelity. As we perform an evolution under a time-dependent Hamiltonian, U is obtained by discretising time with time steps $dt = T/N_{steps}$, with $T = \pi/\Omega$ and $N_{steps} = 10^4$. The number of time steps has been chosen such that convergence is reached.

When the qubits are not interacting, the fidelity can be broken down as a product of single-qubit gate fidelities. We then modify these fidelities to remove the *known, but unwanted* Z-rotations via virtual Z gates (or equivalently applying a change of rotating frame) [40]. It is more complicated to perform such manipulation when the qubits are interacting, as it is the case when using the exchange-based method. However, it should not be a problem in our case, as we study the system in the reference frame rotating at the mean frequency. Any Z rotation performed while one qubit is in a dot will be compensated when it is swapped.

Fidelity proxy

As performing the full time-dynamics is computationally expensive, we need a fast and easy-to-compute fidelity proxy when comparing two clusters in the heuristics (see below). Inspired by the Rabi model Eq. (1.22), we defined the following fidelity proxy $\mathcal{F}_{proxy}$,

$$\mathcal{F}_{proxy} = \prod_{i \in \mathcal{D}} \tilde{\mathcal{F}}_i^{target} \times \tilde{\mathcal{F}}_i^{non-target}, \tag{3.3}$$

where,

$$\tilde{\mathcal{F}}_i^{target} = \prod_{j \in \mathcal{T}_i} \frac{\Omega^2}{\Omega^2 + (\omega_i - \omega_{q,j})^2} \tag{3.4}$$

$$\tilde{\mathcal{F}}_i^{non-target} = \prod_{j \in \mathcal{N}_i} \frac{(\omega_i - \omega_{q,j})^2}{\Omega^2 + (\omega_i - \omega_{q,j})^2}. \tag{3.5}$$

$\mathcal{D}, \mathcal{T}_i, \mathcal{N}_i$ denote the set of drive frequencies, the set of qubits targeted by the drive of frequency ω_i and those which are not respectively.

Process fidelity

The last fidelity I am going to use is the one characterising the quality of the homogenisation of the g-factors within a cluster. For instance, when using the exchange-based method (see Section 3.3), the process fidelity is computed by evaluating the fidelity (as in Eq. (3.2)) of the implemented unitary obtained through Hamiltonian simulations with $n_q = n_t = 2$, with a drive at a frequency corresponding to the average of the two dots, and the exchange coupling on. For shuttling, we perform similar simulations with $n_q = 2$ and $n_t = 1$.

Rather than computing these fidelities for more qubits, we pre-compute them for both the exchange and shuttling-based methods for various values in the parameter space (see Fig. 3.4 and Fig. 3.7) and interpolate between them.

3.3 Effective frequencies using SWAP gates

To reduce frequency crowding my first idea was to leverage the exchange interaction between neighbouring quantum dots. My idea is quite simple, suppose that you have two electrons trapped in neighbouring dots such that their respective wavefunctions are well localised. I can then assign to each of them a g-factor which I will denote g_1 and g_2. By lowering the barrier between the two dots, they will start to swap positions as seen in the previous chapter. If this swapping is fast enough compared to the global drive, one could think that each electron will obtain an effective g-factor $g^{'}$ corresponding to the average of the individual ones $g^{'} = (g_1 + g_2)/2$.

Focusing first on the case $n_q = 2$, I show that by using a global drive with a single tone at a frequency $\omega_q = (\omega_{q,1} + \omega_{q,2})/2$, I can perform an X gate on both qubits.

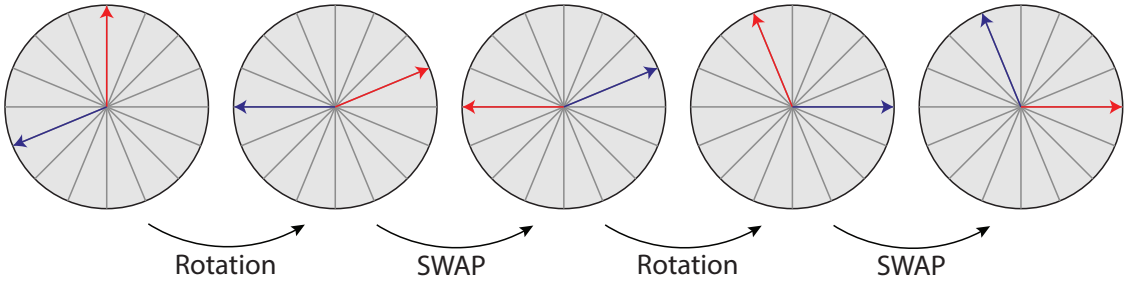

Figure 3.2: Schematic time-evolution of two qubits of distinct frequencies $\omega_{q,1}$ (blue) and $\omega_{q,2}$ (red) when $J \gg \omega_{q,1}$ and $J \gg \omega_{q,2}$ which results in instantaneous swaps. At each time step, the blue arrow rotates by one increment while the red one advances by three increments. The resulting dynamics is a precession of both qubits at the effective frequency $(\omega_{q,1} + \omega_{q,2})/2$ corresponding to two steps. This figure was designed by Adam Siegel and is used here with his agreement.

3.3.1 Two-qubit case

In the absence of an AC field, both electrons precess around the z-axis with a frequency corresponding to their individual g-factor. Now, suppose that I lower the barrier between them to turn on the exchange interaction. To get an idea of the evolution of the system, I discretise time and look at the evolution of the qubits in Fig. 3.2. When lowering the barrier, the exchange coupling should be sufficiently large compared to any other frequency such that we can consider that we have instantaneous SWAPs between the qubits at each time step. In Fig. 3.2, we notice that the two qubits are now effectively rotating at an average frequency ω_q which would lead one to suppose that we could perform independent X gates on each of them by sending a single drive at this frequency. We can either say that we homogenised the qubits' frequencies or consider that we have now created an effective qubit whose frequency is equal to the average of the original ones.

To confirm this assertion, let me now add the global AC magnetic field at the frequency $\omega_q = (\omega_{q,1} + \omega_{q,2})/2$. Starting from a given state, both qubits should experience individual X gates within a reference frame rotating at frequency ω_q. In Fig. 3.3, we set $J = 20(\omega_{q,2} - \omega_{q,1})$ and $\omega_{q,2} - \omega_{q,1} = \Omega$, and we start from the (arbitrary) initial state $|\psi_0\rangle$ defined as,

$$|\psi_0\rangle = |0\rangle \otimes \left(\sqrt{\frac{3}{4}}\,|0\rangle + \sqrt{\frac{1}{4}}\,|1\rangle\right), \tag{3.6}$$

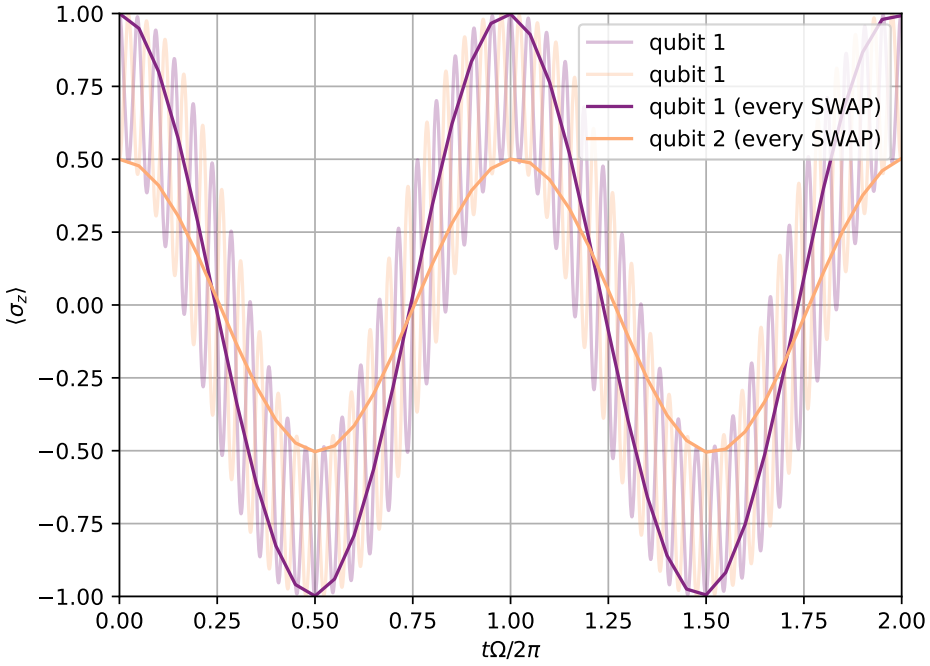

Figure 3.3: Time evolution of the expectation value of the Pauli Z operator of two qubits under a global field composed of a single frequency $\omega_q = (\omega_{q,1} + \omega_{q,2})/2$. By sampling the value of the expectation value at each point corresponding to a full SWAP cycle we notice that we are performing an X operation on both qubits.

and plot the evolution of the expectation value of the Pauli Z operator with respect to the time-evolved state. In light purple and orange my co-author plots the quantities $\langle\psi(t)|\, Z_1 \otimes \mathbb{1}\, |\psi(t)\rangle$ and $\langle\psi(t)|\, \mathbb{1} \otimes Z_2\, |\psi(t)\rangle$. We notice that on top of the slow X-like oscillations (envelope of the light purple and orange signals), we have fast oscillations corresponding to both electrons getting swapped. By sampling the expectation values at times corresponding to full SWAP cycles – every $2\pi/J$ when each electron is back to its original dot – we obtain the dark purple and orange curves. They show that we retrieve X-like oscillations for both qubits. This confirms that when the exchange is turned on, both qubits share the same effective frequency ω_q.

It is important to note that the exchange coupling also generates entanglement between the qubits, which is something I would like to avoid as my goal is to implement gates on individual qubits. To prevent the operation from creating net entanglement, one has to tune the exchange coupling J such that when the target unitary $U_{target} = X \otimes X$ has been performed, both electrons are back to their

original dots. This means that J and Ω are related by,

$$\frac{\pi}{\Omega} = \frac{2p\pi}{J},\ p \in \mathbb{N} \tag{3.7}$$

which leads to an exchange coupling,

$$J = 2p\Omega,\ p \in \mathbb{N}. \tag{3.8}$$

However, achieving an X-like rotation for a particular initial state is not sufficient to confirm that I implemented the target unitary U_{target}. To characterise how close I am to the target unitary, I can use the process fidelity $\mathcal{F}_{process}$ defined in Section 3.2.5. This fidelity is dimensionless and depends on the ratio between the different frequencies of the problem, J/Ω and ω_q^{12}/Ω with $\omega_q^{12} = \omega_{q,2} - \omega_{q,1}$ the difference in frequency between the dots. In Fig. 3.4, I illustrate the evolution of the process infidelity $1 - \mathcal{F}_{process}$ for different ratios, assuming that the condition defined in Eq. (3.8) is satisfied.

As the ratio J/Ω increases, the infidelity decreases monotonically regardless of the ratio ω_q^{12}/Ω. This agrees with my expectations: as J gets larger (with ω_q^{12} constant), we approach the instantaneous SWAP model presented in Fig. 3.2, which ideally yields a process fidelity $\mathcal{F}_{process} = 1$. Conversely, if ω_q^{12} decreases (and J is fixed), the two dots become more similar until they eventually obtain the same frequency, also leading to $\mathcal{F}_{process} = 1$.

Using realistic numbers extracted from Table 3.1, one obtain ratios $J/\Omega = 100$ and $\omega_q^{12}/\Omega \gtrapprox 1$ leading to a process infidelity between 10^{-4} and 10^{-3} on average, which suggests that our scheme could be practical.

3.3.2 n-qubit case

In a similar fashion as in the two-qubit case, one could control n qubits with a single tone at their average frequency by homogenising their g-factors. Contrary to the previous section, having the electrons visit each dot and return to their original positions cannot be performed by turning the exchange interaction between every dot at the same time. However, it can be performed by turning on the exchange

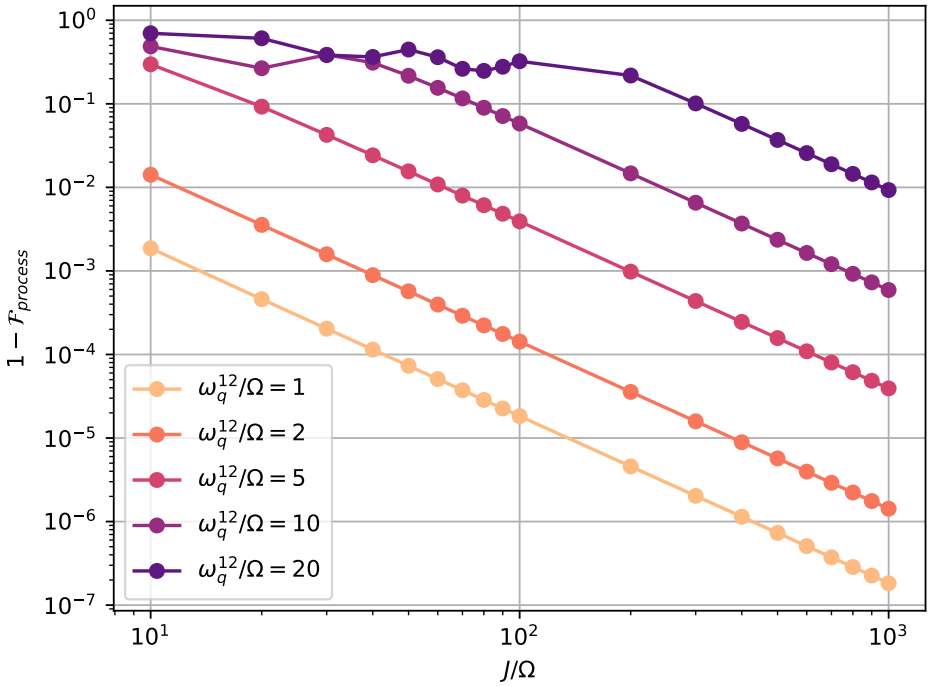

Figure 3.4: Process infidelity $1 - \mathcal{F}_{process}$ of the exchange-based scheme for different ratios J/Ω and ω_q^{12}/Ω. J is chosen such that Eq. (3.8) is verified.

interaction between pairs of dots and alternate SWAP operations, as illustrated in Fig. 3.5. This results in a greater number of SWAP gates to complete a *single* cycle, I would thus require an even larger J than before to obtain a high fidelity. It is worth noting that the additional number of gates depends on the architecture, as presented in Fig. 3.5. In the case of this four-qubit example, using a 2D architecture reduces the required number of SWAP gates by half.

In addition to the requirement for a larger J (which could pose practical challenges), the n-qubit scenario involves numerous uses of the exchange interaction, unlike the two-qubit case where it is activated only once. Considering the susceptibility of the exchange gate to charge noise, as explored in Chapter 2, even a slight infidelity in SWAP operations could render our protocol impractical for $n > 2$ qubits. Moreover, note that charge noise will lead to a violation of Eq. (3.8). I will thus limit myself to using this method on two qubits at a time. As the purpose of this chapter is to tackle the addressability issue, I will not model charge noise.

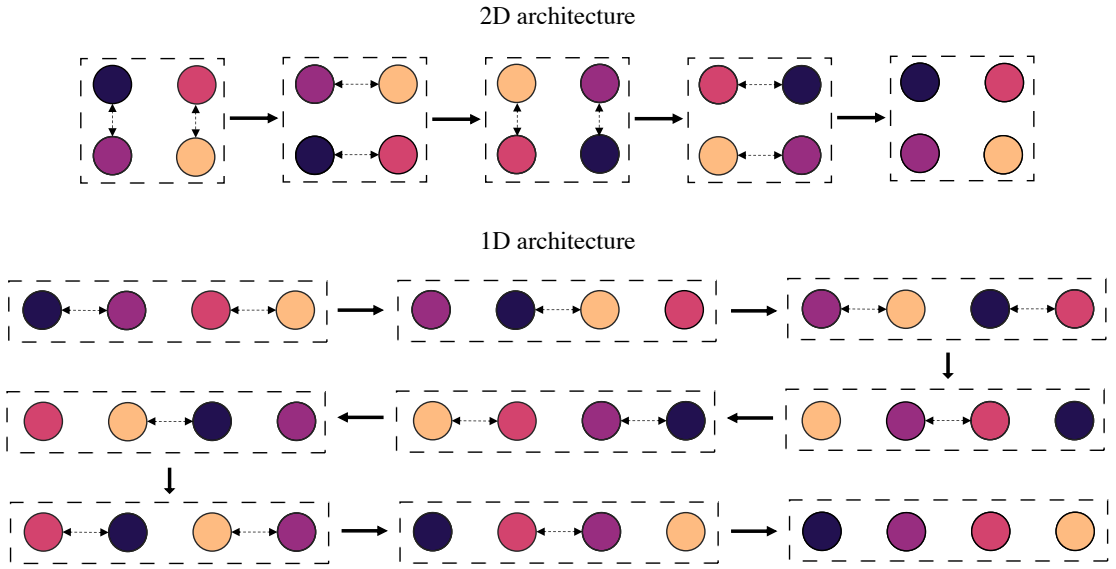

Figure 3.5: Schematic illustration of one cycle in which every electron explores every dot for a 1D and 2D architectures. Note that the number of SWAPs (dashed arrows) is smaller for the 2D architecture, which means that one could get away with a better fidelity compared to the 1D case for a fixed J. Or equivalently, one would need to use a larger J with a 1D architecture to match the performances of the 2D version.

3.3.3 Performance on a 2×N array

To test the performance of the exchange-based scheme, I consider a 2×N array with $n_q = 10$. I now have two different kinds of qubits: target qubits, for which I want to apply an X gate, and non-target qubits which should remain unaffected. Using the previously described technique, I can reduce the number of distinct qubit frequencies by pairing qubits of the same type.

For small 2×N architectures, it might be possible to find the optimal pairing (i.e the set of pairs of qubits) by exploring all possibilities. However, as n_q increases, this task quickly becomes intractable. To find a pairing $\mathcal{P}$ yielding a high fidelity, I developed a simple heuristic. Starting from an empty pairing $\mathcal{P} = \{\}$ perform the following steps:

1. Choose a qubit q.

2. Copy $\mathcal{P}$ into a new pairing $\mathcal{P}'$.

3. Pair q with one of its neighbouring qubit of the same type (target or non-target) q' and add the pair to $\mathcal{P}'$. If q' was already part of a pair (e.g. with qubit q''), replace this pair with the new one. $\mathcal{P}'$ is thus a new pairing in which the pair (q, q'') is replaced with (q, q').

4. Compute the proxy fidelity of $\mathcal{P}'$ as in Eq. (3.3), and the process infidelity as defined in Section 3.2.5.

5. If the the product of the proxy and process fidelities is improved compared to $\mathcal{P}$, set $\mathcal{P} = \mathcal{P}'$.

6. Go back to Step 2 until all neighbours of q have been explored.

7. Go back to Step 1 until every qubit has been explored.

In Fig. 3.6, I plot the evolution of the infidelity $1 - \mathcal{F}$ as defined in Eq. (3.2). Here, J is chosen such that $J/\Omega \approx 100$, which could be achievable in practice according to Table 3.1.

Without using binning, I notice that the improvement in fidelity between the naive (purple) and exchange-based (orange) methods is marginal. This could be due to the fact that for our values of $\Delta\omega_q/\Omega$, it might not be always beneficial to pair qubits together as it could lead to a high process infidelity. I am currently exploring new pairing heuristics as well as other types of proxy fidelities. I only averaged the results over $N_{runs} = 200$ repetitions, as this is enough to get an idea of the protocol's quality. Note that this is the only simulation for which we have not used $N_{runs} = 10^4$, as the results was not promising. When using binning (not shown here), I currently find that there is no net advantage to using our method and we are still trying to understand why. Further investigation is needed to reach a conclusion on the practicality of this scheme.

Even if the SWAP-based scheme allows one to reduce the infidelity, the improvement seems limited. It is therefore not a viable strategy for scalable control techniques, at least in this form. In the following section, I present a shuttling-based solution that results in high-fidelity global control.

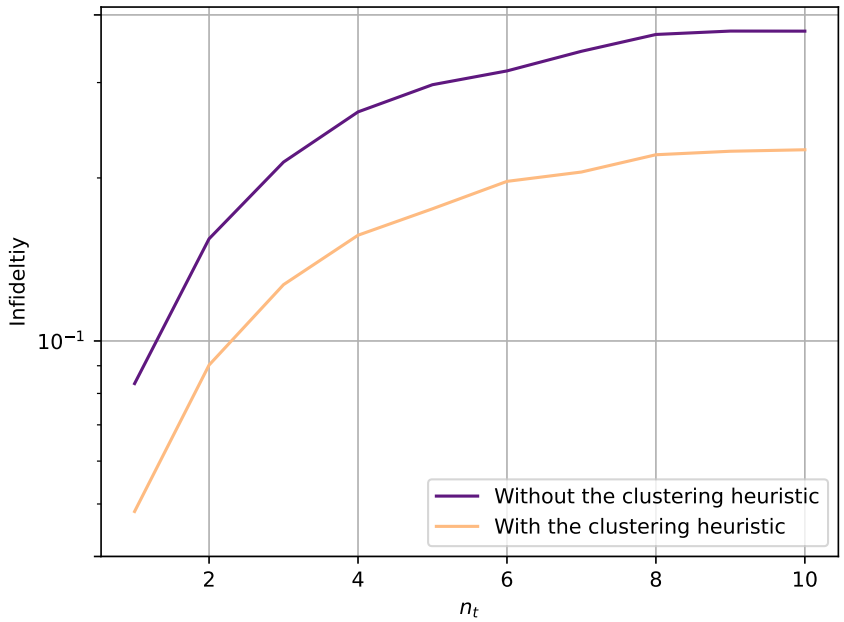

Figure 3.6: Evolution of the infidelity with the number of targets n_t for $n_q = 10$. The difference in fidelity between the naive (purple) and exchange-based (orange) methods is marginal. Note that we have not used the binning technique here.

3.4 Effective frequencies using shuttling

In the previous section, I showed that by making an electron explore a double quantum dot using the exchange interaction, I can effectively modify its g-factor and thus its frequency. This principle can be generalised: any mechanism that allows for an electron to explore multiple dots and come back to its original position rapidly enough will alter its frequency: it will now be equal to the average of the frequencies of the visited dots. In this section, I focus on a more straightforward way to move qubits: electron shuttling [115–117], which consists in physically moving the electrons through electrical control. Diving into the details of shuttling is not essential here (a brief explanation is provided in Chapter 4).

3.4.1 Single-qubit case

For ease of presentation, I will begin by explaining the scheme for a single electron, which can be transferred between two quantum dots using shuttling. This

corresponds to the continuous version of the situation presented in Section 3.3.1 with the qubit frequency evolving as $\omega_q : x \mapsto \omega_q(x)$ along the shuttling path. Using the same intuition as before, by shuttling the electron fast enough, it will obtain an effective frequency equal to the average of that of the individual dots. Note that this frequency narrowing effect has been observed before [118], but was not used for reducing crosstalk errors.

In practice, electron shuttling can be achieved by modifying the energy detuning between the dots. Yoneda et al. found in [115] that the g-factor varies continuously with detuning, following an S curve. Instead of evolving between two discrete frequencies, $\omega_{q,1}$ and $\omega_{q,2}$, we can thus suppose that the electron's frequency evolves as follows,

$$\omega_q(x) = \frac{\omega_q^{12}}{1 + \exp^{-\lambda x}}, \tag{3.9}$$

with λ a variable parameter. Note that I am looking at the problem from a frame rotating at the left qubit's frequency, i.e. $\omega_q(x) = 0$ for $\lambda x \to -\infty$ and $\omega_q(x) = \omega_q^{12}$ for $\lambda x \to \infty$. This function is only chosen as an example (motivated by experiments [115]), but our scheme should still work for other functions $\omega_q : x \mapsto \omega_q(x)$.

Determining the drive's frequency becomes more complex than simply averaging the g-factors of the individual dots. It requires detailed characterisation of the g-factors at each position, which could be performed beforehand. However, I expect that any protocol requiring electrons to be shuttled will benefit from this knowledge.

I suppose that the electron has a speed v and travels back and forth between dots separated by a distance d. Practically, we will have $d = 100$ nm and $v \approx 10$ m/s [118]. For simplicity, I model the evolution of the electron's position $x(t)$ as a triangular pulse between $-x_{\max} = d/2$ and $x_{\max} = d/2$ with period $2d/v$. The electron's frequency should then be equal to the average value of $\omega_q(x(t))$ over time, which is equal to $\Omega/\pi \int_0^{\pi/\Omega} \omega_q(x(t))\, dt = (\omega_{q,1} + \omega_{q,2})/2$. Note that π/Ω needs to be equal to a multiple of d/v for the electron to come back to its original position at the end of the π pulse. I thus obtain a similar condition as in Eq. (3.8),

$$\frac{dp}{v} = \frac{\pi}{\Omega},\ p \in \mathbb{N}, \tag{3.10}$$

leading to,

$$v = \frac{dp\Omega}{\pi},\ p \in \mathbb{N}. \tag{3.11}$$

By applying a single tone with frequency $\omega = (\omega_{q,1} + \omega_{q,2})/2$, I should be able to perform an X gate. In Fig. 3.7, my co-author plotted the process infidelity as defined in Section 3.2.5, as a function of the ratios $v/d\Omega$ and ω_q^{12}/Ω, confirming that we indeed implement a single-qubit X gate. To match the the variation of the frequency ω_q to experimental results presented in [115], we set $\lambda d = 30$.

The evolution of the infidelity matches my expectations: it decreases with increasing shuttling speed, as we approach the instantaneous swap model presented in Fig. 3.2. However, as ω_q^{12} increases at a fixed speed, the infidelity gets larger as more frequencies need to be averaged. If $\omega_q^{1,2}$ decreases, the frequency will not change much along the shuttling path which allows one to perform a high-fidelity gate by driving at $\omega = \omega_{q,1}$.

Using realistic numbers extracted from Table 3.1, one obtain ratios $v/d\Omega = 100$ and $\omega_q^{12}/\Omega \gtrapprox 1$ leading to a process infidelity $\approx 10^{-10}$. Our scheme has thus a high fidelity in a practical regime in the noiseless case (i.e without shuttling noise). Below, I will elaborate more on this point.

3.4.2 n-qubit case

It is import to highlight that unlike the exchange-based approach, I am not restricted to dealing with pairs of electrons. The generalisation to the n-qubit case follows naturally: I use shuttling to simultaneously move multiple electrons in a closed loop of dots without any interaction between them. Contrary to the exchange-based scheme, no entanglement is built between the electrons (at least to a first approximation). The only remaining requirement is that at the end of a drive period, the electrons should return to their original positions (see Eq. (3.11)).

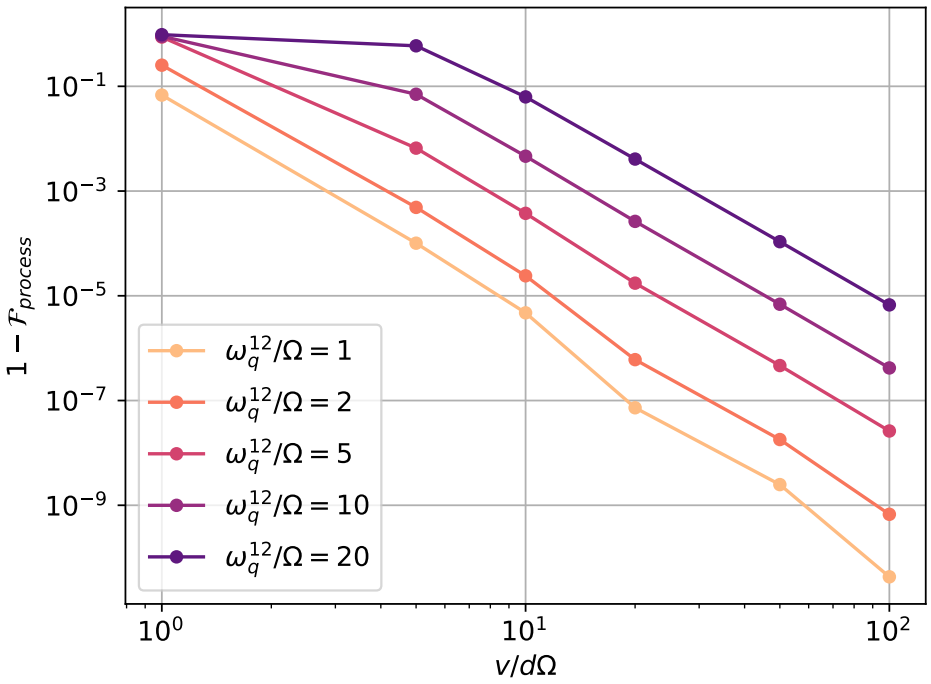

Figure 3.7: Process infidelity $1 - \mathcal{F}_{process}$ of the shuttling-based protocol for different ratios $v/d\Omega$ and ω_q^{12}/Ω.

3.4.3 Performance on a 2×N array

Here, I focus on 2×N arrays of quantum dots with filled dots on the top row and empty dots on the bottom row to allow shuttling in loops. Similarly as in Section 3.3, I need a clustering heuristic.

My first intuition was the following. To minimise the number of frequencies in the spectrum, one would like to have the largest clusters possible. However, large clusters lead to large process infidelities. To balance between the two, one should start from the largest cluster and split them only if it improves the fidelity. After simulations, I found that this method was not performing well within our workflow.

After multiple discussions, my co-author Adam Siegel found another method which was performing better. In this approach, we start with a 2×N array and move the non-target qubits to the bottom row. Then, we shuttle every qubit around the full device. As explained in Fig. 3.8, we end up with two clusters following each other: a red cluster composed of target qubits and a blue one composed of non-target qubits. If we do not perform any other operations, every qubit will just

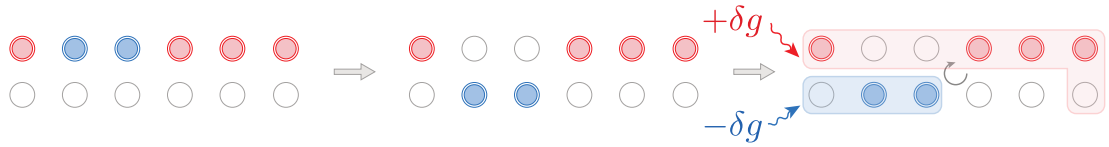

Figure 3.8: Shuttling protocol. Starting from a 2×N array, we move the non-target qubits to the bottom row. We then shuttle all the electrons around the device and dynamically modify the Stark shift, by $+\delta g$ for the target (red) cluster and $-\delta g$ for the non-target (blue) cluster.

acquire a g-factor equal to the average of g-factors over the full device. We will thus not be able to distinguish between the different types of qubits. To solve this issue, we propose dynamically modifying the Stark shift such that every dot containing an element of the red (blue) cluster is Stark-shifted by $+\delta g$ $(-\delta g)$.

By using this protocol, the frequency spectrum now only contains two frequencies bins separated by $2\delta\omega_q = 2\delta g B_0$. We can further increase the fidelity by synchronising the rotations of the non-target qubits such that they perform a 2π rotation (identity gate) while the target qubits perform a π rotation (X gate) as explained in Appendix B.1. Note that, contrary to the binning presented in Fig. 3.1 in which this synchronisation will be approximate (because of the large number of non-target bins), it can be made exact here.

Suppose that we managed to perfectly homogenise the g-factors, the time it takes to perform an X gate is $T = \pi/\Omega$. As showed in Eq. (1.22), the non-target qubits will rotate at a frequency $\sqrt{\Omega^2 + 4\delta\omega_q^2}$ as they are detuned by $2\delta\omega_q$. The synchronisation requirement leads to the following condition,

$$\frac{\pi}{\Omega} = \frac{2p\pi}{\sqrt{\Omega^2 + 4\delta\omega_q^2}}, \; p \in \mathbb{N}. \tag{3.12}$$

By enforcing this condition, we can guarantee that crosstalk vanishes. The only remaining source of infidelity is the process infidelity.

In Fig. 3.9, we plot the evolution of the infidelity $1-\mathcal{F}$ as a function of the number of qubits n_q and number of target qubits n_t. The squares represent the infidelity obtained when directly putting the electrons into bins as explained in Fig. 3.1 without using shuttling, while the circles represent the infidelity when using the protocol

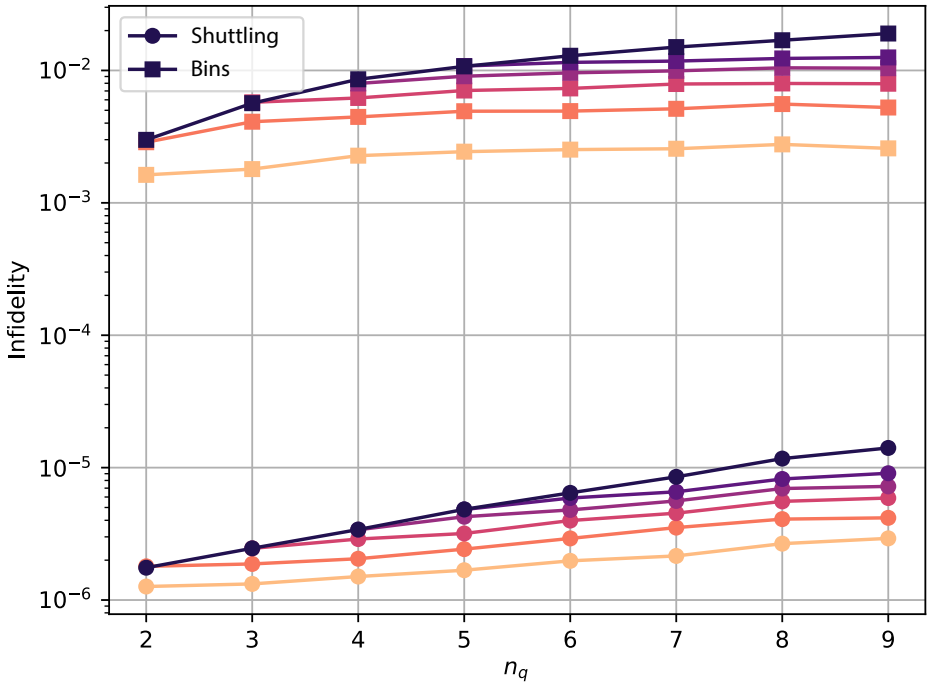

Figure 3.9: Evolution of the infidelity $1-\mathcal{F}$ for different n_q when only using the binning (squares) or using a combination of the shuttling and binning (circles) protocols as in Fig. 3.1. We plot the infidelity for various number of target qubits: $n_t \in \{1, 2, 3, 4, 5, n_q\}$ from lighter to darker colors.

presented in this section. We see that there is a drastic improvement as without shuttling, crosstalk errors dominate, while with shuttling it is the process infidelity.

It is important to note that we have not introduced shuttling noise in our simulations. If we did, the infidelity would saturate at the shuttling noise level as the per-dot shuttling fidelity has been experimentally evaluated to be 10^{-4} [119]. Using values from Table 3.1 (i.e. 10ns to move an electron from one dot to the other), one would need to move between 100 dots during an X gate of gate time ≈ 1 μs. This leads to an overall shuttling infidelity of 10^{-2}. However, as shuttling is currently receiving a lot of attention and is expected to be a key enabler in the development of large-scale quantum computers, one can hope that this infidelity will keep decreasing.

While promising, the scheme presented here requires the dynamical modification of the Stark shift while the qubits are shuttled. As this might be challenging to implement in practice, we suggest an alternative architecture which relaxes this requirement. This alternative requires a 3×N architecture which is a natural

extension of 2×N devices. Note that qubits are placed in the middle row, while the top and bottom ones are empty. Following a similar idea as the protocol presented previously, one could Stark shift all the upper dots by $+\delta g$, all the lower ones by $-\delta g$ and shuttle the electrons according to their type. The target qubits will loop around the first two arrays (up and middle), while the non-target qubits will loop around the last two (middle and bottom). This allows us to obtain similar results as before, without the requirements of dynamically modifying the g-factors. However, it is important to note that using such an architecture requires a lot of empty dots.

By proposing this alternative, I hope to convey the idea that the shuttling-based frequency homogenisation is closely related to the choice of architecture. With more freedom on this choice, one could come with even better protocols.

3.5 Discussion

In this chapter, I explore various techniques to improve the global control of spin qubits, which will be crucial for the development of large-scale spin-based quantum computers.

First, I focus on a method generating effective qubits based on the natural exchange gate between quantum dots. While promising for two-qubit control, this approach proved impractical for more than two qubits due to the small fidelity improvement, at least at the moment. Then, I demonstrated that by using shuttling and grouping qubits in moving clusters, one can achieve global X gates with a greatly improved fidelity. Note however that I focused on near-term 2×N architectures and that our protocol could perform even better on different architectures as hinted at the end of the previous section.

While the current simulations provide valuable insights into the potential of this new control method, several experimental limitations need to be considered for full validation. First, we must account for possible off-resonant driving, which can occur due to the uncertainties in the g-factors landscape and from the fact that the driving frequency will generally not be composed of a single tone but will have a finite-bandwidth. Second, we assume that the exchange coupling and the

drive can be switched on instantaneously, given the that in fact finite time will be required, this could lead to non-adiabatic effects. These can be mitigated by carefully designing the shape of the pulses. Third, factors like charge noise and excited valley states were not included, as we focused on improving addressability.

Finally, the idea of generating effective qubits is more general than what I have presented in this chapter. Any method allowing for electrons within a region of varying g-factors to explore a fixed set of g-factors and come back to their initial position fast enough (compared to the drive's period) will suffice. Additionally, it is worth noting that I focused on protocols in which the electrons spend an equal amount of time in each dot, but it would be interesting to relax this constraint. For instance, I would like to investigate whether I can generate effective frequencies corresponding to a weighted average of the individual g-factors by letting the electrons spend different time in each dot.

4

Multicore quantum computing

This chapter is based on the research published in Physical Review Applied [120] and closely follows that manuscript. The research was conducted by Brennan Undseth, Zhenyu Cai, Bálint Koczor, Simon Benjamin and myself. While Simon Benjamin contributed to the core of the multicore idea, Zhenyu Cai and Bálint Koczor explored its applications which I summarise in Section 4.6. Brennan Undseth and I modelled different methods to link distant spin qubits. I focused on the cavity-based approach which I present in Section 4.2 and Section 4.3 whereas Brennan Undseth explored shuttling as briefly presented in Section 4.4. I use text verbatim from [120] where I am the original author.

Contents

4.1 Introduction 68
4.2 Linking distant spin qubits using a cavity 69
4.3 Generating Bell Pairs using cavities 72
4.3.1 The model 72
4.3.2 Bell pair generation 74
4.3.3 Entanglement purification 78
4.4 Linking distant spins using shuttling 83
4.4.1 Overview of electron shuttling 83
4.4.2 Distributing Bell pairs using shuttling 84
4.5 Comparison of entanglement distribution modes . . . 84
4.6 Applications 86
4.6.1 Error mitigation 86
4.6.2 Other applications 90
4.7 Discussion 91

4.1 Introduction

As seen in Chapter 3, increasing the number of qubits on a single chip leads to qubit control issue. Note that this scaling issue is also faced by the leading qubit platforms, i.e. superconducting- and ion-based qubits. For instance for linear ion traps, it is expected that scaling beyond $50 - 100$ ions will be difficult due to issues like decrystallisation, spectral crowding etc. [121–123]. Conversely, the large footprint of superconducting qubits poses daunting infrastructure challenges if one wishes to build a chip with millions of qubits. Moreover, the scaling of monolithic (a single-chip for a single-core) processors will come with increasing complications in providing power [124, 125], controlling the qubits (cross-talk and frequency crowding) [126, 127] and chip cooling. Additionally, as the number of qubits grows, the probability to find a faulty component somewhere in the device grows as well because of finite yield. While a finite number of defects should be tolerable [128], switching to a modular architecture will entirely alleviate this issue as one could remove the faulty cores rather than the entire chip. IBM, for instance, announced that they will switch to a multicore approach as they are already reaching the limits of single-chip designs [129]. Considering these factors alongside the fact that we can already design moderate-scale devices, we can suppose that a multicore system will be created long before monolithic structures become truly scalable.

Even if one assumes that the ultimate goal is to create a large monolithic chip, the multicore era might be a pivotal stage in the field. First, moderate-size single-core quantum computers are built [4, 7, 8, 10, 130, 131], then multiple cores are used in parallel without quantum interlinks to speed up sampling-based tasks like variational quantum algorithms [132] or error mitigation [120, 133]. Next, we use quantum interlinks to join these cores [134, 135] as we will discuss in this chapter. Finally, as cores grow and interlinks become more dense, we will be able to reach architectures suitable for error correction codes [136–138].

Here, I study a multicore architecture based on silicon spin qubits (shown in Fig. 4.1), while noting that it could be transferred to other qubit modalities. Although the performance of the studied interlinks is encouraging, the fidelities

that we obtain for realistic device characteristics are well below those expected of short-range, intra-core operations. This could be seen as a fundamental constraint drawing from the insights of classical computing in which the quality of the links deteriorates with the distance. Even as quantum computing matures, the quality of long-range links may never match that of short-range ones. Through our detailed modelling, we find that this statement holds for near-future devices in terms of the fidelity of inter- versus intra-core links. Fortunately, it is possible to increase this fidelity using entanglement purification techniques [139–144] which I define later. Note however that entanglement purification will reduce the speed of inter-core operations. By suitably tailoring purification protocols, I find fidelities (with respect to the closest of the four canonical Bell states) of 99.5% can be achieved after a few successive uses of the interlink. Higher values are achievable in principle, but as anticipated in [139] the noise in intra-core operations becomes the limitation.

This chapter is organised as follows. First, I review techniques used to link distant spin qubits using a cavity in Section 4.2 and present our linking protocol based on the generation of Bell pairs in Section 4.3. Then, I summarise my co-author's work exploring shuttling as a Bell pair transfer mechanism in Section 4.4 and compare the quality of our protocols in Section 4.5. Finally, I present my co-authors' work on applications of this multicore architecture for context, focusing on the implementation of a powerful error mitigation technique.

4.2 Linking distant spin qubits using a cavity

In the context of superconducting qubits, the coupling of distant qubits can be achieved by mediating quantum information through a superconducting cavity within the circuit quantum electrodynamics (cQED) framework [145–148]. By using a similar method, one can generate a coherent coupling between the resonator's modes and silicon spin qubits over few millimetres. For simplicity and for consistency with the shuttling alternative that I will summarise later, I will restrict myself to LD qubits while recognising that singlet-triplet qubits may have advantages in cavity coupling [149].

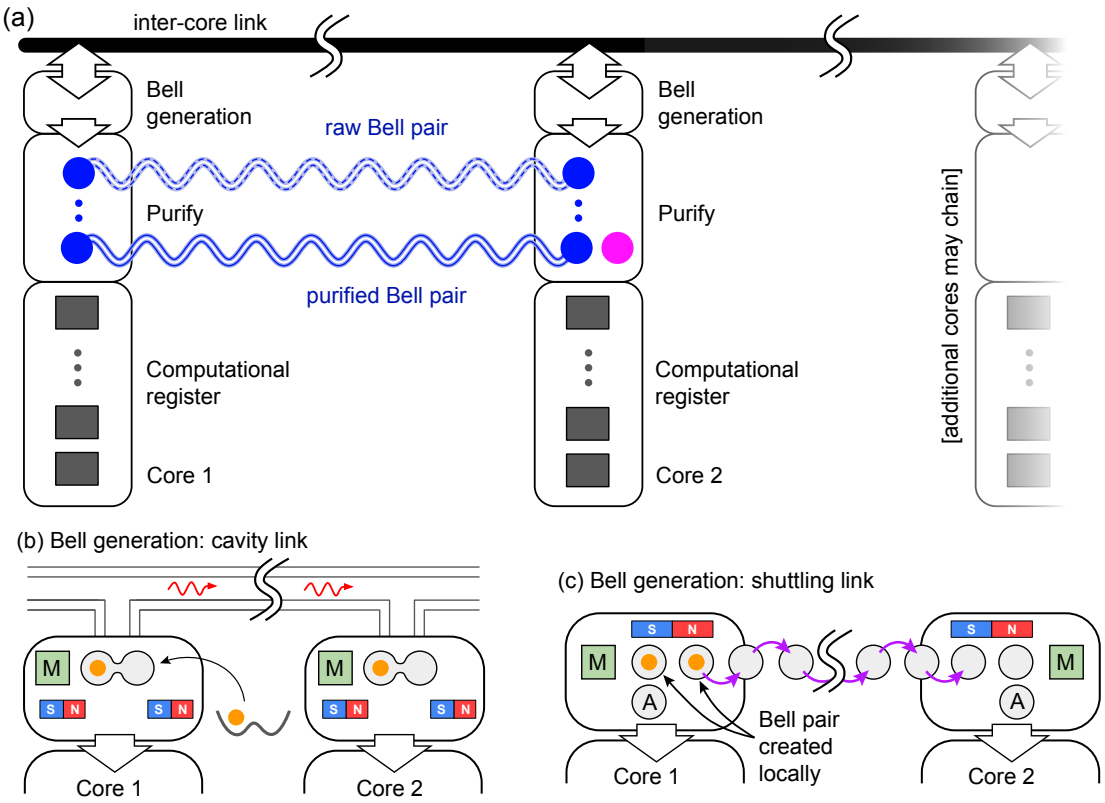

Figure 4.1: (a) Distant quantum cores (core 1, core 2 etc.) linked via purified Bell pairs (wavy lines). In this chapter, I consider distributing Bell pairs via a cavity (b) and by shuttling (c) (not to scale). (b) Schematic of two spin registers connected via a superconducting cavity. Using micromagnets, I can increase the spin-orbit coupling of the electrons and thus couple distant qubits through microwave photons (red squiggles) in a cavity. Charge measurements are used to speed up the process. A successful measurement outcome will lead to a raw Bell pair, whose quality can be improved through purification. (c) Schematic of two spin registers connected via a quantum dot shuttling channel. A Bell pair consisting of two electrons (small orange dots) is created locally. One of the electrons is shuttled to the other core (this chain does not use micromagnets as we do not need to control the spin degree of freedom during shuttling). This figure has been extracted from [120] with the agreement of my co-authors.

However, applying the above procedure to LD qubits is hard. The reason being that electronic spins have a small magnetic dipole, which means that the spin-photon coupling is slower than the spin's decoherence rate. Ref. [150] gives a comprehensive review of cQED with single-electron spin qubits.

In recent years, numerous experimental works have explored the possibility of coherently coupling a semiconductor spin qubit with microwave photons [62, 151, 152]. By placing a cobalt micromagnet above a single electron confined in a double quantum dot (DQD), Refs. [62, 151] have leveraged the following two-step process

to effectively couple electron spins with a cavity. First, the large magnetic field gradient generated by the micromagnet increases the spin-orbit coupling which leads to a hybridisation of the spin and charge degrees of freedom. Then, the dipole-dipole interaction between the cavity and the charge indirectly couples the spin to the cavity resulting in significant coupling rate (on the order of several MHz).

Despite the fact that hybrid spin qubits are more prone to charge noise, Refs. [62, 151] have reached the strong coupling regime, i.e the regime in which the spin-photon coupling rate is greater than the spin decoherence rate and the cavity loss rate. Borjans et al. [153], demonstrated that one can coherently couple two spin qubits with a resonator by extending the architecture in [151] and Harvey-Collard et al. [61] reported the coupling of distant spins through virtual photons. The latter represents a significant step towards the experimental realisation of cavity-mediated two-qubit gates.

The success of these experiments has stimulated further theoretical work towards implementing long-range two-qubit gates. While Ref. [154] showed how an iSWAP gate can be implemented with fidelities above 90% by optimising the spin-charge hybridisation with current hardware, Warren et al. [155] have predicted the possibility of 99.5% fidelity $\sqrt{\text{iSWAP}}$ gates by exploiting the low-energy dynamics of the system. More recently, the authors proposed a photon-mediated cross-resonance gate that is more robust to charge noise [156]. It is important to note that all these works assume a dispersive parameter regime: this regime may reduce cavity loss, but it results in relatively slow gates.

Here, I focus on generating specific Bell states using cavity mediated interactions (see Section 4.3), in contrast to long-range general unitary quantum gates in the dispersive regime which were the focus of the aforementioned prior works [154, 155]. This allows me to work in a resonant parameter regime that generates entanglement on a significantly faster timescale at the cost of introducing small (coherent) errors. I then apply powerful purification techniques to reduce such errors in the raw Bell pairs thereby non-deterministically generating high-fidelity Bell pairs in Section 4.3.3. After purification, these Bell pairs can be used as a resource

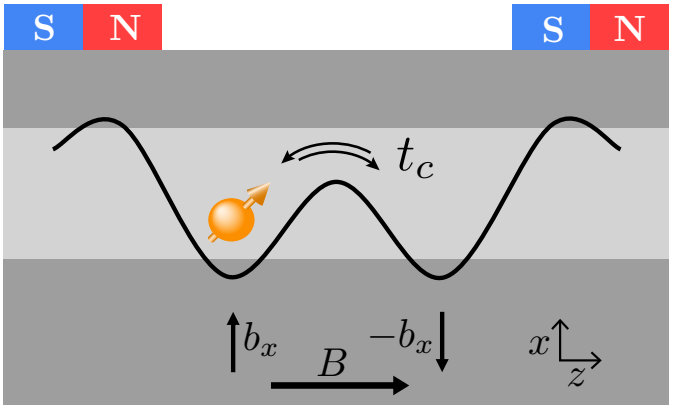

Figure 4.2: Sketch of a single electron confined in a DQD. The DQD is generated by gates on top of a material heterostructure (a Si/SiGe heterostructure for instance). Cobalt micromagnets are placed above the DQD to generate the transverse magnetic field b_x which will hybridise the charge and spin degrees of freedom. By lifting the spin degeneracy, the parallel magnetic field $(0, 0, B)^T$ allows us to define a qubit on the electron's spin.

to perform arbitrary two-qubit gates via gate teleportation. As I explain below in Section 4.6, more generally these Bell pairs can be used as valuable resources for a range of specific practical tasks beyond gate teleportation.

4.3 Generating Bell Pairs using cavities

4.3.1 The model

By confining a single electron in a DQD, I obtain a *spin qubit* encoded in the spin of the electron with the basis states $|\uparrow\rangle, |\downarrow\rangle$ and a *charge qubit* encoded in the position of the electron with the basis states $|\mathrm{L}\rangle, |\mathrm{R}\rangle$ as explained in Chapter 1. The spin and charge qubits are coupled by a magnetic field gradient across the two dots. It is important to note that in my model, I do not consider higher excited states of the charge degree of freedom nor the possible coupling with the valley degree of freedom [157]. To first order these would not change our results in the present section due to other mechanisms of more significant imperfections as explained below (in contrast my co-author did need to take into account the valley coupling for the shuttling-based alternative). As such, the full Hamiltonian of this system can be written as

$$H_{\mathrm{DQD}} = \left(\frac{\epsilon}{2}\tau_z + t_c\tau_x\right) + \frac{1}{2}Bs_z + \frac{1}{2}\tau_z \otimes (\vec{b}\cdot\vec{s}). \tag{4.1}$$

Above and throughout the chapter I set $\hbar = 1$. Here τ_i and s_i are Pauli matrices σ_i with $i \in \{x, y, z\}$ acting on the charge and spin qubits, respectively. The first term above represents the Hamiltonian for a charge qubit in a DQD with ϵ and t_c being the detuning and tunnel coupling between the two dots. The second term represents the Zeeman splitting B of the spin qubit due to a magnetic field that defines the spin direction. The last term represents the coupling between the spin and charge qubits due to the *difference* $\vec{b} := (b_x, b_y, b_z)^T$ in magnetic fields between the two dots. In this section, I assume the main gradient of the magnetic field is along the x-direction ($\vec{b} = (b_x, 0, 0)^T$). This gradient can be generated using micromagnets as illustrated in Fig. 4.2. While using micromagnets in coupling spins to cavities may present technical challenges [61, 153], I expect it is not a fundamental bottleneck for the cavity-based approach given only two sets of micromagnets per pair of connected cores need to be implemented.

Here, I consider a resonator described by the usual quantum-harmonic-oscillator Hamiltonian $H_r = \omega_r a^\dagger a$ where ω_r is the resonator frequency and $\hat{a}^\dagger$, $\hat{a}$ are the corresponding (bosonic) creation and annihilation operators. Furthermore, I model the coupling between the resonator and the charge qubit via the Hamiltonian [158],

$$H_{\tau r} = g_c \tau_z \otimes \left(a + a^\dagger\right), \tag{4.2}$$

where g_c is the charge-photon dipole coupling factor. Hence, via interactions mediated through the charge qubit, we can couple two spin qubits in two distant DQDs via the resonator. I write the full Hamiltonian for such a system as

$$H_{srs} = H_r + \sum_{i=1}^{2} (H_{\mathrm{DQD},i} + H_{\tau r,i}), \tag{4.3}$$

where $i \in \{1, 2\}$ are labels of the two distant DQDs, each containing a spin qubit and a charge qubit. The resonant condition corresponds to the Zeeman splittings of the two DQDs matching the resonator's frequency: $B_1 = B_2 = \omega_r$. Note also that in my numerical simulations, I model up to six photons per resonator mode – this guarantees a very good approximation of the exact dynamics given that the expected occupation number is much lower for higher energy states.

4.3.2 Bell pair generation

In this section, I simulate the above model and numerically search for a set of optimal evolution parameters. I explicitly take into account hardware noise that is comparable to that of current devices. I find that the cavity-based system is capable of generating raw Bell pairs of sufficiently high fidelity that can then be used as input for purification.

In particular, I numerically simulate the evolution of the system's state under the Hamiltonian defined in Eq. (4.3) while also accounting for different noise processes via the general Lindblad master equation as

$$\frac{\partial \rho}{\partial t} = -i\left[H_{srs}, \rho\right] + \sum_i \gamma_i \left(A_i \rho A_i^\dagger - \frac{1}{2}\left\{A_i^\dagger A_i, \rho\right\}\right). \tag{4.4}$$

Here γ_i are positive rates, and $\hat{A}_i$ are Lindblad terms. Thereafter, I denote each noise process using a tuple (γ_i, A_i). In this chapter, I focus on three sources of noise. First, I consider cavity loss (κ, a) as defined by a cavity loss rate κ and the annihilation operator a. Then, I consider spin dephasing $\left(1/(2T^*_{2,s}), s_z\right)$ characterised by the spin decoherence time $T^*_{2,s}$ and the Pauli operator s_z. Finally, I consider charge noise $\left(1/(2T^*_{2,c}), \tau_z\right)$ as defined by the charge decoherence time $T^*_{2,c}$ and the Pauli operator τ_z. Our model assumes a time-independent Hamiltonian and the high-frequency fluctuations of the detunings are accounted for in the charge noise model [159]. While the charge dephasing $T^*_{2,c} \leq 400$ ns is significantly faster than spin dephasing, note that in our protocol quantum information is never stored in the charge state – the charge degree of freedom merely acts as an interaction mediator for a brief period of time ≈ 15 ns.

As described above, my goal is to entangle the two remote spins by generating a specific Bell state – which may be noisy due to imperfections. I stress that my model in Eq. (4.3) implemented in the resonant regime leads to highly non-trivial evolutions and even with no decoherence it cannot perfectly generate a Bell pair as detailed in Appendix C.4. Intuitively, the reason is that I assume I only have the ability to abruptly alter the interaction Hamiltonian – and in the full Hamiltonian I only have control of at most 10 individual parameters, whereas at the end of

the evolution a number of restrictive conditions need to be satisfied. First, the two spins have attained the highest degree of entanglement; second, the spins are separable from the charge degrees of freedom and from the cavity; third, the cavity is nearly in the vacuum state. As such, given the dimensionality of the Hilbert space is $2^4 \times N_{cav}$ (allowing for the fact that only the first $N_{cav} \approx 7$ energy levels are populated in the cavity), the dynamical system is clearly under-parameterised. Therefore, I optimise the system's trajectory to come as close as possible to the desired condition, but do not expect to perfectly meet it.

In order to obtain the best possible fidelity, I will perform projective measurements at a given optimal time on the interaction-mediating charge degrees of freedom, thereby also increasing the entanglement in our raw Bell pairs at the cost of a slight probability of failure as explained in Appendix C.4. My ultimate goal is that given a sufficiently high level of entanglement of our raw Bell pairs, I can improve its fidelity to arbitrarily high levels by using purification techniques.

To find an optimal set of parameters in Eq. (4.3), we performed a gradient-based optimisation that maximises the entanglement between the spins. The optimal parameter values can be found in Appendix C.1. As a figure of merit for the optimisation, I use the concurrence [160] of the joint state of the two spins given a fixed charge measurement outcome. Averaging this concurrence for the charge measurement outcomes that we accept then defines our cost function. I also introduce a set of constraints for the optimisation. Firstly, the resonant condition $B_1 = B_2 = B = \omega_r$, and, secondly $2t_{c,1} = 2t_{c,2} = 2t_c = B$ (achievable by tuning the tunnelling barrier of the DQD [161]) which gives us the maximum coupling between the spin and the charge [158]. Apart from these fixed parameters, we optimise all other parameters that define our model and the initial state of each electron.

In Fig. 4.3, I report the evolution of our system under an optimal set of parameters. In particular, Fig. 4.3(a) shows the probabilities of different charge measurement outcomes, while Fig. 4.3(b) shows the corresponding evolution of the concurrence between the spins. The concurrence peak nicely matches a probability peak given charge measurement outcomes in the odd parity subspace (as defined

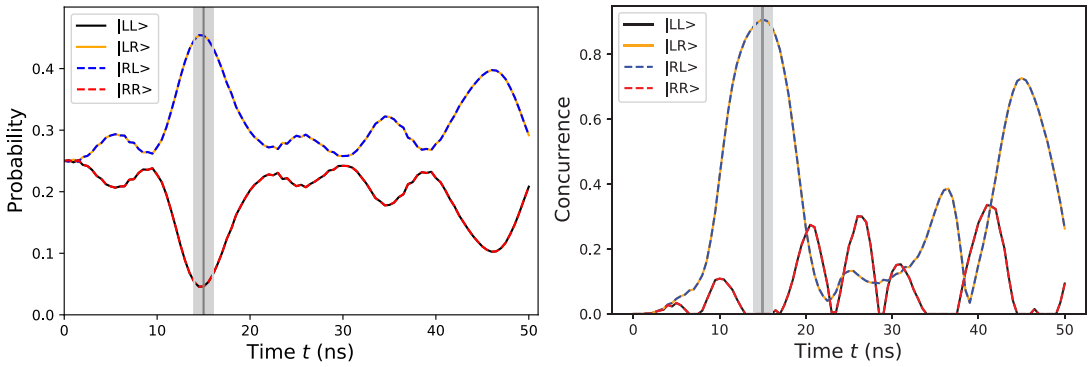

Figure 4.3: (a) Evolution of the probabilities of measuring the charges at different locations. (b) Evolution of the concurrence between the distant spins. Both plots were computed using the optimised evolution parameters. The optimal initial state is given by: $|\psi\rangle \approx \frac{(|\mathrm{L}\rangle - |\mathrm{R}\rangle)|\uparrow\rangle}{\sqrt{2}} \otimes \frac{(|\mathrm{L}\rangle - |\mathrm{R}\rangle)|\uparrow\rangle}{\sqrt{2}} \otimes |0\rangle$, with the first two terms describing the state of the first and second electrons respectively, and the last one the state of the cavity. To account for potential difficulties to stop the interaction exactly at the maximal probability, I averaged the final density matrix over different stopping times (shaded area) weighted by a Gaussian distribution centred at 15ns (see text). Here, $T^*_{2,s} = 120\mu$s and $T^*_{2,c} = 400$ ns.

by the states $|\mathrm{LR}\rangle$ and $|\mathrm{RL}\rangle$). I discuss in Appendix C.4 that the discarded charge measurement outcomes eliminate the dominant coherent error in which the charge state $(|\mathrm{LL}\rangle - |\mathrm{RR}\rangle)/\sqrt{2}$ is coupled to the unentangled spin state $|\downarrow\downarrow\rangle$ due to energy conservation on-resonance – which explains the nearly zero concurrence of the discarded measurements in Fig. 4.3(b)(black and red lines).

As I perform a charge measurement at the maximal peak, i.e., at around $t \approx 15$ ns, I find a high success probability ≈ 0.9 while the resulting state achieves a relatively high concurrence ≈ 0.89. It is expected that in typical experiments the coupling between each spin and the cavity is electrically controllable [151, 158]. This allows me to rapidly switch off their interactions at the optimal evolution time—which corresponds to the peaks in Fig. 4.3—by electrically adjusting either detuning or the tunnelling couplings. However, it might be difficult to stop the interaction exactly at the maximal probability in practice. I take this into account by averaging the final density matrix over different stopping times weighted by a Gaussian distribution centred at 15 ns (with $\sigma = 0.5$ ns).

Given state-of-the-art technology achieves smaller jitter by orders of magnitude, ours is a very conservative estimate to account for the possibility of integrating (likely lower quality) pulse-control electronics within the chip. Furthermore, the

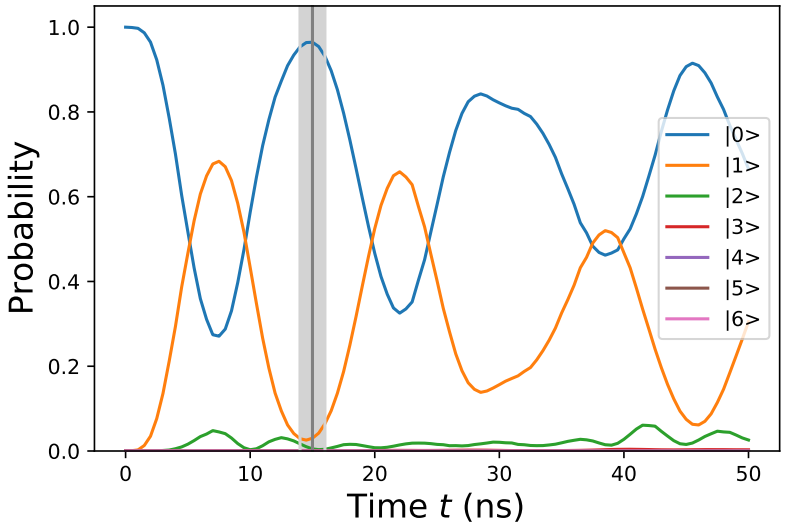

Figure 4.4: Evolution of the photon's state after a successful charge measurement. When I stop the interaction and make the measurements, there is a ≈ 0.96 probability to find the cavity in the vacuum state.

purified fidelities I observe are nearly oblivious to jitter times, in particular I observe fidelities 99.60%, 99.59%, 99.51% with a jitter time of 0 ps, 50 ps, 500 ps, respectively, which speaks for the robustness of the approach. As an artefact, the present dynamics might introduce a deterministic phase to the quantum state, however, if this phase is stable over consecutive generations of Bell pairs it is automatically removed by the purification process.

To be able to reuse the cavity straight after having generated a Bell pair, we need to end up the protocol in the vacuum state. One possibility would be to measure the charges and the cavity, and deem the result as successful if I am measuring both the charges in either the state $|\mathrm{LR}\rangle$ or $|\mathrm{RL}\rangle$ and the cavity in the vacuum state. However, this would lead to a decrease of the success probability compared to the case where I only perform charge measurements. Fortunately, after a successful charge measurement, the probability to find the resonator in the vacuum state is high ≈ 0.96 (see Fig. 4.4). Even though the probability to be in the zero state for the cavity is high, one may need to consider the cavity heating up during consecutively repeated Bell pair generations. Nevertheless, given cavities in current experiments have typical decay rates $\approx$ 1MHz [153], dissipation should occur on a timescale that is smaller than a single round of entanglement purification.

$T^*_{2,c}$ (ns)	400	100	50
Fidelity	94.5%	90.0%	84.4%

Table 4.1: Pre-purification fidelities for different noise parameters. The fidelities are expressed with respect to the Bell pair $|\psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$.

Furthermore, the cavity is only used for a very small fraction of the procedure – allowing for cooling protocols to be periodically performed if necessary.

Note that the quality of the entanglement in the raw Bell pairs (without purification) might not be sufficient to perform high-fidelity long-range operations as I report in Table 4.1. In particular, I expect that charge dephasing noise (which deteriorates our hybridised states) is the dominant source of error, and I report fidelities in Table 4.1 for different values of $T^*_{2,c}$ decoherence times on the order of tens up to hundreds of nanoseconds. For each simulation I assumed a spin decoherence time of $T^*_{2,s} = 120$ µs [162]. It is worth noting that $T^*_{2,s}$ is not a bottleneck, and while it could be reduced without significantly affecting our results, it is expected that $T^*_{2,s}$ will have to be reasonably large given for any practical application one needs to perform at least hundreds of quantum gates.

4.3.3 Entanglement purification

As noted earlier in the introduction, as early as 2003 entanglement purification was explored theoretically as a means to distribute a quantum task over subsystems with non-optimal links [139–141, 163, 164]. Here, we leverage this technique to improve the quality of our raw Bell pairs. But before diving into how I used entanglement purification in the context of the multicore architecture, I start by presenting a simple example explaining the concept.

Simple example: bit-flip error

Suppose that Alice and Bob are sharing multiple copies of a noisy version of the state $|\phi^+\rangle$ described by the state ρ such that,

$$\rho = p\left|\phi^+\right\rangle\!\left\langle\phi^+\right| + (1-p)\left|\psi^+\right\rangle\!\left\langle\psi^+\right|, \tag{4.5}$$

where $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ and $|\psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$ are maximally entangled states and $p > 1/2$. Alice and Bob wish to *purify* ρ in order to increase the fidelity with respect to the state $|\phi^+\rangle$. They only have access to local quantum operations, measurements and a classical communication channel. Entanglement purification consists then in multiple rounds in which two copies of ρ are locally entangled and one copy is measured out. A key characteristic of the state ρ to note here is that $|\phi^+\rangle$ and $|\psi^+\rangle$ have different parities. If the tested Bell pairs have the same parity, then only certain measurement outcomes are possible. If Alice and Bob are not seeing these allowed outcomes, they discard the non-measured copy, otherwise they store it. As a consequence, the remaining Bell pair passed the test, and its fidelity is higher, such that its state after a single round of purification is now given by,

$$\rho_1 = \frac{p^2}{p^2 + (1-p)^2} \left|\phi^+\right\rangle\!\left\langle\phi^+\right| + (1 - \frac{p^2}{p^2 + (1-p)^2}) \left|\psi^+\right\rangle\!\left\langle\psi^+\right|, \tag{4.6}$$

where $p^2/(p^2 + (1-p)^2) > p$ when $p > 1/2$.

Alice and Bob now have a Bell-pair with improved quality and they can repeat the process to further increase the fidelity. Note however, that entanglement purification is a stochastic process and as such it becomes less and less probable to succeed as one increases the number of rounds.

General case

More generally, when performing a round of purification we only suppress the error for one noise channel (e.g. bit-flip) from order p to p^2 while the other (phase-flip) increases from p to $2p$. One should thus create a second validated Bell pair and subsequently combine the two in a higher round of purification mitigating the effect of the remaining noise channel. Hence, using four 'raw' Bell pairs with infidelity of order p we can distil a single Bell pair with infidelity of order p^2 (with a modest prefactor). Additional rounds of purification can be performed to further reduce the noise to order p^3 or higher, albeit at an exponentially increasing cost in Bell pairs.

Even though it is relevant to consider noise in the local operations, it will be mitigated by the purification process too. The noise in the final round is thus the

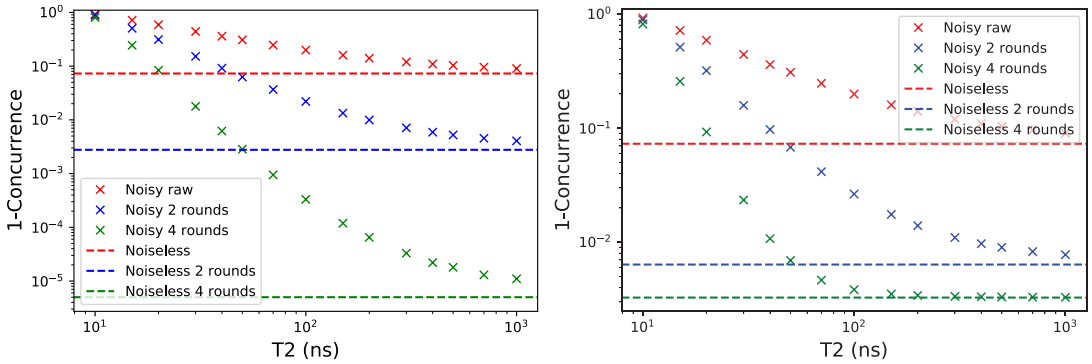

Figure 4.5: Concurrence (measure of entanglement) as a function of charge decoherence for different levels of purification. Left (right) without (with) errors in the local quantum gates in the purification process. The dashed Red line represents a perfect unitary evolution (no decoherence) and perfect charge measurements and the only source of imperfections is the fact that our fast, resonant dynamics in Eq. (4.3) cannot exactly generate a Bell pair.

most significant. Note also that if the raw Bell Pairs have a significant structure, it can be leveraged to tailor the purification circuit. As such, full rank noise (i.e. white noise) is the least desirable.

Bell pairs generated using a cavity

I use the controlled interaction between the spin qubits and the cavity described in the previous section as a process to generate noisy Bell pairs. In Fig. 4.5, I plot the concurrence (entanglement) in the Bell pairs after various rounds of purification as a function of the dominant noise source, the charge decoherence time $T^*_{2,c}$. I also model spin decoherence and imperfections in the local single and two-qubit gates via depolarizing noise as well as measurement errors. Here, I use the variant of the standard purification circuit in Fig. 4.6 that yields the best concurrence after the purification process for the particular structure of noise in my raw Bell pairs (note that this particular purification protocol transforms the initial $|\psi^-\rangle$ into $|\phi^+\rangle$). In Appendix C.5, I report that by adding further single-qubit gates one can obtain better concurrence in the ideal purification case. For simplicity, in the following I will mostly focus on the $T^*_{2,s} = 120$ µs and $T^*_{2,c} = 400$ ns case. Such a charge decoherence rate can be reached with current hardware [165].

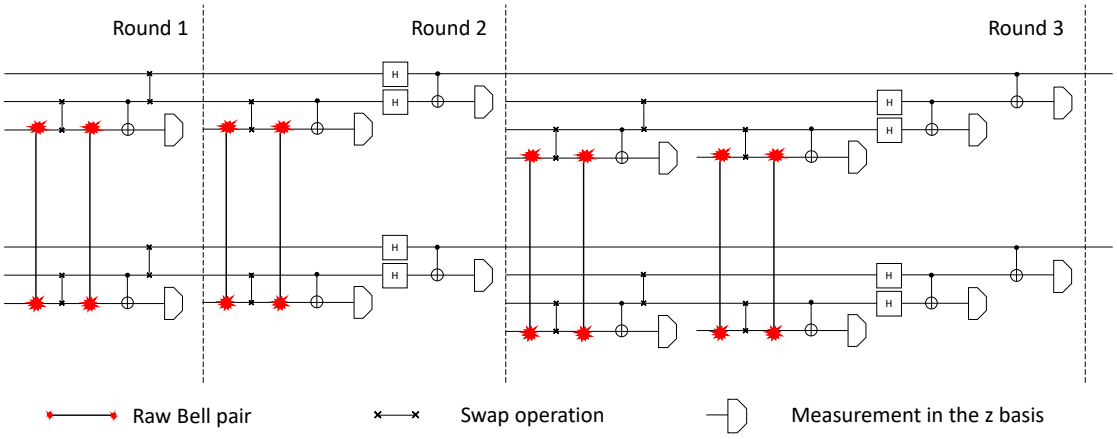

Figure 4.6: Entanglement purification protocol circuit. Here three rounds are displayed but depending on the noise model two rounds might be sufficient. The first round addresses bit flip errors in the input noisy, raw Bell pairs while the second round reduces phase flip errors. Then, in the third round I repeat the same procedure as in the first round but with two level-2 Bell pairs as input. The fourth round (not shown here), reuses the circuit in round 2 but with two level 3 Bell pairs as input. If both bit flip and phase flip errors are present, the protocol should terminate at either the second or the fourth round depending on the noise rate. Errors in the purification process are accounted for by adding depolarizing noise after each local single- and two-qubit gate, and measurement errors (99.9% fidelity) at each round (not displayed here). While the figure shows the canonical choice in entanglement purification, I tailored the protocol towards the specific error model of cavities by inserting further single-qubit rotations.

Since in entanglement purification I need to discard certain measurement patterns, it is a non-deterministic process. In my case, the purification protocol (in both rounds) succeeds if I measure both qubits in the same state. Fig. 4.7 shows the main steps of the purification process. On average, one needs to generate $N_{avg} \approx 5.2$ noisy Bell pairs to obtain a 99.5% fidelity Bell pair. Assuming typical gate/measurement times this results in an average generation rate $g_{Bell} \approx 0.14$ MHz (see Appendix C.6), even though in the limit of very fast gates and measurements entanglement generation is dominated by the cavity interaction time with a generation rate of $g_{Bell} \approx 12.9$ MHz.

Fortunately this rate of 0.14 MHZ can still suffice for powerful applications including the derangement-based error mitigation described in Section 4.6. Prior experiments demonstrated that coherence can be maintained with respect to phase noise for as long as 28 ms using a Carr-Purcell-Meiboom-Gil sequence [162], and indeed even devices fabricated in 300 mm commercial wafers have achieved a T_2

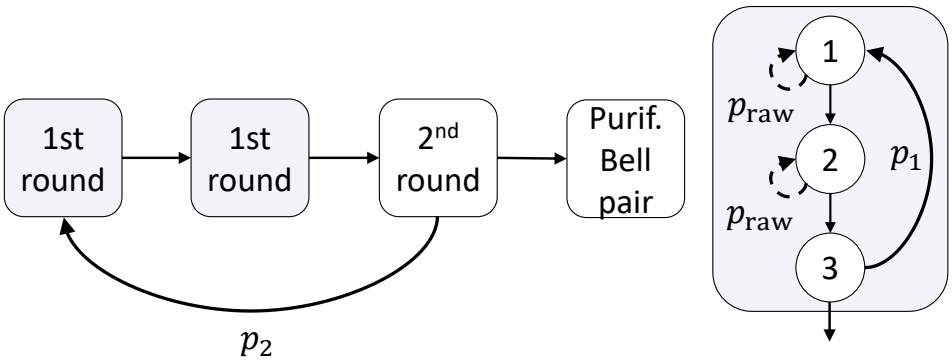

Figure 4.7: Purification process flow chart for two rounds of purification. The blue-shaded box represents the steps for one round of purification. It requires the generation of two raw Bell pairs (steps 1 and 2), a process that fails with probability $p_{\text{raw}} \approx 0.10$. The purification process (step 3) fails with probability $p_1 \approx 0.08$. After having successfully generated two level-1 Bell pairs, I perform a second round of purification to get the final purified Bell pair, which fails with probability $p_2 \approx 0.06$. The probabilities are given for the $T^*_{2,s} = 120$ µs and $T^*_{2,c} = 400$ ns case.

$T^*_{2,c}$ (ns)	400	100	50
2 rounds	99.5% (99.7%)	98.7% (98.9%)	96.6% (96.8%)
4 rounds	99.8% (>99.9%)	99.8% (>99.9%)	99.7% (99.9%)

Table 4.2: Post-purification fidelities for different noise parameters and rounds of purification. In parentheses: without noise in the local operations in the purification process. The fidelities are expressed with respect to $|\phi^+\rangle$.

time of 3.7 ms [24]. I therefore expect our approach allows the generation of tens or even hundreds of purified Bell pairs to be used for 'comparing qubits' between cores using the derangement method in Section 4.6 before dephasing would negate the advantage of the process.

For applications where this rate is insufficient, note that the cavity is in use for only a very small fraction of the procedure (≈ 15 ns), and I can thus significantly improve generation rates by a 'parallelisation' of the purification step: The extended architecture would include n 'purification stations' on both sides of the cavity, and while the cavity is exclusively used to generate a single Bell pair at a time, these raw Bell pairs are fed in an alternating way into the different purification stations. By choosing a modest parallelisation $n \leq 5$ one would expect that the generation rate is effectively increased by the same factor n.

For charge noise worse than $T^*_{2,c} = 400$ ns, four rounds of purification are necessary to obtain an adequate fidelity. In Table 4.2, I collect the fidelities for different charge noise parameters for two and four rounds of purification. The fidelities reported in parentheses are obtained without errors in the local gates and during measurements in the purification process. Let us define the obtained Bell pair after the ith successful purification round as a Bell pair of level i. Four rounds of purification is rather demanding, since in order to go from the ith round to the $(i+1)^{th}$ one needs two level-i Bell pairs. As purification is a stochastic process, performing more rounds leads to a decrease in the success probability and hence to a smaller generation rate. See Appendix C.6 for its computation for $T^*_{2,c} = 100$ ns and $T^*_{2,c} = 50$ ns.

With better hardware, the number of purification rounds will decrease while the generation rate and the fidelity will increase. Indeed, with negligible decoherence, one obtains a fidelity $\approx 99.9\%$ with $g_{Bell} \approx 0.15$ MHz after two purification rounds ($g_{Bell} \approx 13.9$ MHz via fast gates and measurements). However, we will see in Section 4.6 that for many applications one does not need to generate perfect Bell pairs and already a fidelity of 99.5% is enabling for powerful NISQ applications.

4.4 Linking distant spins using shuttling

4.4.1 Overview of electron shuttling

For quantum computers based on neutral atoms, ions or electrons' spins, a natural way of moving quantum information is performed by directly moving (or shuttling) the qubits themselves [8, 115–118, 166–168]. Shuttling has emerged as a key feature for these platforms in the quest for fault-tolerant quantum computers [8, 114].

For spin qubits trapped in gate-defined quantum dots, there are two main modes of transport. The so-called *bucket-brigade* [115, 166, 167] and *conveyor-belt* modes [116–118, 168]. In the former, the electrons move from one dot to another by detuning the dots, and in the latter, a particular set of time-dependent voltages is applied to the different electrodes to create a moving dot. While conveyor-belt seems to be the way forward for shuttling as it is less demanding in terms of

control voltages, experimental realisations were still lacking at the time we wrote the paper. That is why my co-author focused on the bucket-brigade mode which was already implemented [115].

4.4.2 Distributing Bell pairs using shuttling

Previous theoretical papers have been concerned with shuttling spin qubits in silicon [166, 167, 169, 170] and how they can conserve the state over the process. However, my co-author looked at the less demanding task of distributing Bell pairs. When distributing a Bell pair, the coherent errors happening during the shuttling process can be mitigated using entanglement purification. As shuttling can be performed on the same timescale as physical gates, we could continuously generate entanglement and use it as a resource to perform long-distance gates between distant cores (see Fig. 4.1).

Using a restricted subspace to study shuttling of an electron across a double quantum dot, while taking into account spin-orbit and valley-orbit couplings, my co-author found that one could reliably distribute Bell pairs with concurrences $\approx 99.5\%$ after a single round of purification. More details can be found in [120].

4.5 Comparison of entanglement distribution modes

In Fig. 4.8, I compare the imperfect Bell pairs obtained from the cavity and shuttling-based approaches for distributing Bell pairs by plotting their respective density matrices before and after entanglement purification.

The figure demonstrates that both methods are able to produce Bell pairs of sufficient quality, with subsequent entanglement purification resulting in final fidelities of at least 99.5%. It is interesting to note that the low-rank nature of the noise in the case of shuttling (indicated by the fact that there are only two brown bars in the diagonal of the density-matrix in Fig. 4.8 rather than three) allows for a simpler purification process, requiring only a single purification round. Conversely, the full-rank noise predicted for the cavity-mediated link (three brown bars in the

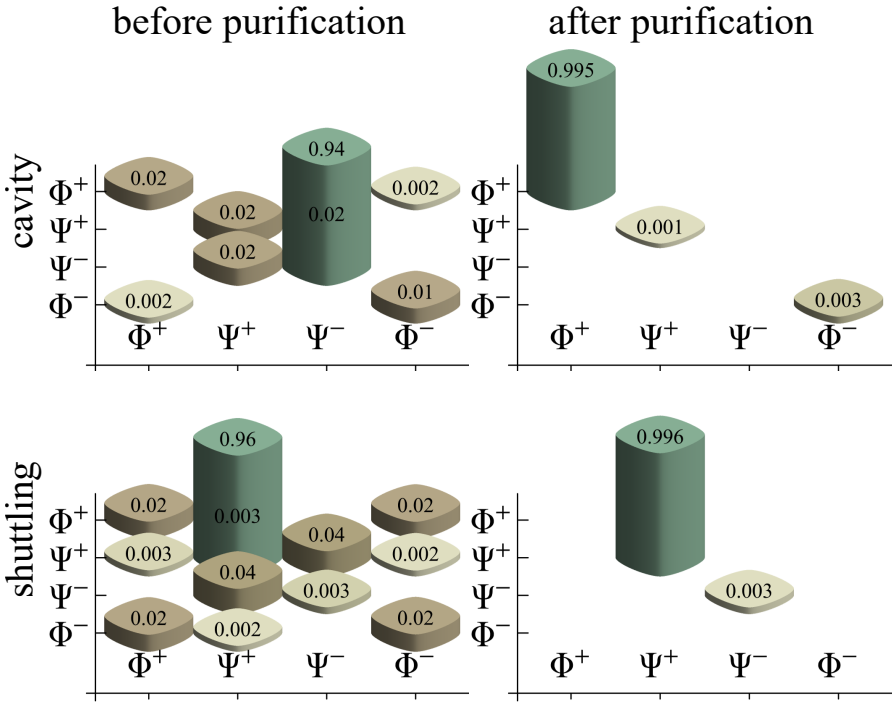

Figure 4.8: Comparing output density matrices in the case of cavity-based and shuttling-based approaches. Absolute values of matrix elements in the Bell basis assuming imperfection parameters $T^*_{2,c} = 400$ ns (top) and $SD_\phi = \pi/4$ (bottom). SD_ϕ represents the standard deviation of the distribution of valley phase between quantum dots . The matrix associated to shuttling before purification corresponds to an average over multiple instances of 25-dot chains and is given up to local rotations. We need 2 rounds for the cavity-based approach, and only 1 for the shuttling-based one. The largest fidelity with respect to the closest of the four canonical Bell states is increased from 94.5% to 99.5% for the cavity-based method, and from 95.7% to 99.6% for shuttling. This figure has been extracted from [120] with the agreement of my co-authors.

diagonal in Fig. 4.8) necessitates multiple rounds of purification to reach high-quality final Bell states. Additionally, it is worth noting that a shuttling channel could simultaneously transport a number of spins in a 'pipeline' mode. In contrast, achieving multiple use of the cavity in parallel might be hard to perform in practice.

While this comparison is insightful, it may be unlikely that a chip architect would have to select between these two different mechanisms in practice. Instead, the choice would likely be constrained by the desired length of the link. Shuttling is expected to be relevant in the $1 - 10\,\mu$m range, while cavity-based links are more appropriate for longer distances up to several mm (or even between chips).

4.6 Applications

4.6.1 Error mitigation

Introduction

Although we are seeing rapid development in error correction [8, 10], we are still far from the fault-tolerant era. While error correction is concerned with recovering a noiseless state from a noisy quantum computer, one could ask for a less demanding task e.g. computing ideal expectation values from noisy devices. This is precisely what error mitigation is concerned with. Error mitigation [133] represents the set of techniques used to reduce the error in computing expectation values without encoding information into logical qubits and thus require fewer resources, making it well-suited for the NISQ and early fault-tolerant eras. I give more details on error mitigation in Section 5.1.

In this section, I will focus on Error Suppression by Derangements (ESD) [171] and Virtual Distillaiton (VD) [172], two error mitigation techniques allowing for an exponential reduction of hardware noise. Both methods rely on the same idea of using multiple copies of a state but differ in their implementation. While my co-author gave a general derivation for N_c copies (with the need of an ancilla) in [171], [172] focused on the case $N_c = 2$ (without an ancilla). Thereafter, we will focus on the general N_c copies and 1 ancilla scenario as illustrated in Fig. 4.9.

Intuition behind Error suppression by derangement

Suppose that we want to compute the expected value of an observable O with respect to an ideal N-qubit state $|\psi\rangle_{id}$. When preparing this state on near-term devices, errors can occur such that the obtained state is given by a statistical mixture, represented by a density operator ρ. Decomposing ρ in its eigenvector basis we get,

$$\rho = \lambda \, |\psi\rangle\langle\psi| + (1-\lambda) \sum_{i=2}^{2^N} p_i \, |\psi_i\rangle\langle\psi_i| \, , \tag{4.7}$$

with $\lambda < 1$, $\sum_{i=2}^{2^N} p_i = 1$ and $\forall i, \ \lambda > (1-\lambda)p_i$. $|\psi\rangle$ is the "dominant" eigenvector, and is equal to $|\psi\rangle_{id}$ up to a coherent error. This error can be discarded as Koczor

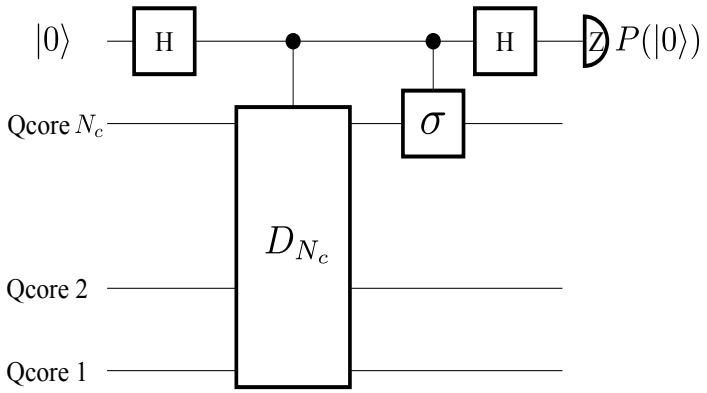

Figure 4.9: Circuit representation of the error mitigation technique. First, N_c quantum cores prepare the same state in parallel. Then we apply the derangement operator D_{N_c} which leads to the copies validating each other. Errors in estimating the observable σ are exponentially suppressed with the increasing the number N_c of cores. Each core is composed of N qubits, the derangement circuit can then be implemented efficiently by distributing $N(N_c-1)$ Bell-pairs between the cores and performing $N(N_c-1)$ controlled-SWAP operations locally. This figure has been extracted from [120] with the agreement of my co-authors.

showed in [173] that it is exponentially smaller than incoherent errors. The aim is then to recover the expected value $\langle O\rangle = \mathrm{Tr}[O\,|\psi\rangle\langle\psi|] = \langle\psi|\,O\,|\psi\rangle$ from O and ρ.

ESD starts with the preparation of N_c independent copies of the same quantum state (see Fig. 4.9) ρ. The most probable state, of these N_c copies taken altogether is $|\psi,\ldots,\psi\rangle$ with a probability λ^{N_c}. In this case, the expectation value of the first register is equal to $\langle\psi,\ldots,\psi|\,O\,|\psi,\ldots,\psi\rangle = \langle\psi|\,O\,|\psi\rangle$, which is the expectation value we are looking for. However, with probability $\lambda^{N_c-1}(1-\lambda)p_i$, we get the second most probable state for which one error occurred in the register i. Without loss of generality, suppose that the error occurs in the first register, meaning that the state prepared is one of the orthogonal eigenvectors $|\psi_i\rangle$. The expectation value is now $\langle\psi_i,\ldots,\psi|\,O\,|\psi_i,\ldots,\psi\rangle = \langle\psi_i|\,O\,|\psi_i\rangle$, which is wrong. Now, if instead we measure the expectation value of the observable $O\mathrm{SWAP}_{1,N_c}$ (where SWAP_{1,N_c} swaps the first and N_c^{th} qubits), the expectation value is equal to $\langle\psi_i,\ldots,\psi|\,O\,|\psi,\ldots,\psi_i\rangle = \langle\psi_i|\,O\,|\psi\rangle\,\langle\psi|\psi_i\rangle = 0$ as $|\psi\rangle$ and $|\psi_i\rangle$ are orthogonal. If an error occurs in one register, swapping it with one that did not experience an error leads to a zero contribution to the expectation value. We can generalise this by swapping every register by using a derangement operator (a derangement being a permutation that has no fixed point).

Koczor showed in [171], that the probability to measure the ancilla qubit in the state $|0\rangle$ in Fig. 4.9 is related to the expectation value of O such that,

$$\begin{aligned} 2\,P(|0\rangle) - 1 &= \mathrm{Tr}\Big(O\rho^{N_c}\Big) \\ &= \lambda^{N_c} \langle\psi| O |\psi\rangle + (1-\lambda)^{N_c} \sum_{i=2}^{2^N} p_i^{N_c} |\psi_i\rangle \langle\psi_i| , \end{aligned} \tag{4.8}$$

where we see that we are exponentially suppressing the errors with N_c. Dividing by λ^{N_c}, we get,

$$\frac{2\,P(|0\rangle)}{\lambda^{N_c}} = \langle\psi| O |\psi\rangle + \mathcal{E}, \tag{4.9}$$

where $\mathcal{E} < CQ^{N_c}$, C is a constant and $Q < 1$ is a suppression factor. If λ is not known, Koczor proved that we can estimate it by running the same procedure without the observable O in Fig. 4.9.

Implementation in the multicore architecture

The multicore architecture is well-suited for implementation of ESD for the following reasons. First, it is straightforward to prepare N_c independent copies of a state using N_c separate quantum cores. Then, the only quantum links between them appear when implementing D_{N_c} right before the ancilla measurement. Finally, Ref. [171] provides a decomposition of D_{N_c} into $N(N_c - 1)$ controlled-SWAP gates between pairs of qubits. The derangement circuit requires less gates than typical circuits used to prepare the quantum states. Using Bell pairs generated via a cavity-based or shuttling-based approach, we can then easily implement D_{N_c}.

While we could use Bell pairs to perform long-range two-qubit gates, we adopt another approach based on teleportation to implement the controlled-SWAP gates (see Fig. 4.1). For ease of presentation, I will explain the protocol for $N_c = 2$. First, I swap the state of a computational qubit from core 2 into a buffer qubit in core 1 using teleportation as illustrated in Fig. 4.10(a). This requires a single Bell pair as well as local operations and classical communication (LOCC). Then, I implement the controlled-SWAP gate locally in core 1 as shown in Fig. 4.10(b). I repeat this procedure N times for each qubits to implement the derangement

operation. For a more detailed explanation of the protocol and how it generalises to an arbitrary number of copies see [120].

Using the Bell pair error rates I derived from my cavity and entanglement purification analysis, my co-author has numerically simulated this protocol.

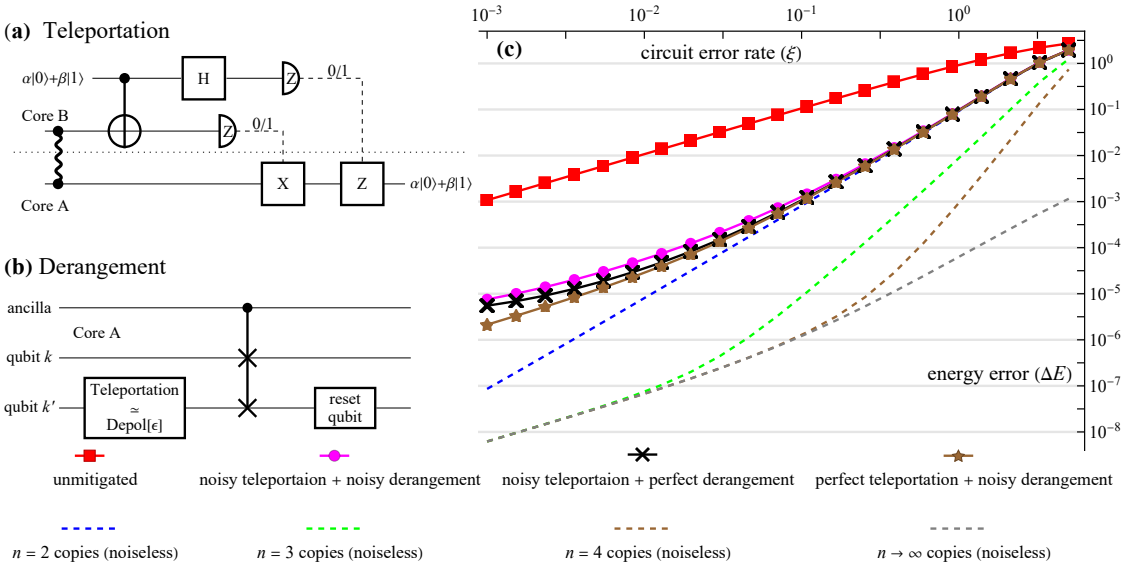

Figure 4.10: (a) Circuit for teleporting a single qubit state from core B to core A using a Bell pair (wavy line) (b) after teleportation we apply apply the derangement circuit equivalent to a controlled-swap between the two qubits in core A. We then reset the lower qubit and repeat this procedure for all computational qubits k which implements the entire derangement circuit. The effect of the imperfections in the Bell pairs is equivalent to adding a depolarising channel during the computation, the derangement circuit then mitigates these imperfections too. (c) Error in estimating the ground state energy as a function of the expected number of errors in the state-preparation circuit (ξ). The unmitigated errors (red squares) are highly suppressed by ESD in the noiseless case using N_c copies of the computational state (dashed lines). In the practically relevant regime when $0.1 \leq \xi \leq 5$, ESD has a similar performance to the ideal one even when taking into account imperfections in the derangement circuit (brown stars), in the Bell pair distribution (black crosses) or in both (magenta circles). This figure has been extracted from [120] with the agreement of my co-authors.

Numerical simulations

Consider a spin-ring Hamiltonian,

$$\mathcal{H} = \sum_{k \in \text{ring}(N)} \omega_k Z_k + J\, \vec{\sigma}_k \cdot \vec{\sigma}_{k+1}, \tag{4.10}$$

with a constant coupling $J = 0.1$ and uniformly randomly generated on-site interaction strengths $\omega_k \in [-1, 1]$. This Hamiltonian appears in the context of many

body localisation [174–176]. We focus on the case $N = 6$ and consider $N_c = 2$ copies of the ground state prepared using a variational Hamiltonian ansatz [177–181]. More details can be found in the appendix of [120].

Using realistic error rates for single and two-qubit gates [120] and considering noisy Bell pairs of fidelity $F = 99.5\%$ prepared either from the cavity-based or shuttling-based methods, teleporting single qubit states is modeled by applying single-qubit depolarising noise of probability 0.5% after every qubit in one of the copies of the computational quantum states [120].

In Fig. 4.10, my co-author plotted the error in estimating the energy of the ground state as a function of the circuit error rate ξ. ξ represents the expected number of errors made when preparing the state. As the measurement cost usually grows exponentially with ξ, let us focus on the practically interesting regime $0.1 \leq \xi \leq 5$. There, the unmitigated error (red squares) is drastically reduced when using noisy teleportation (black crosses), noisy derangements (brow stars) or both (magenta dots). Note that in this regime the curves are superposed. While this shows that ESD is robust against imperfections in the inter-core links, it is important to stress that it incurs an increased measurement cost growing as F^{-N_c} with F being the fidelity of the Bell pairs. See [120] for mor details.

To summarise, ESD/VD techniques are compatible with the multicore paradigm and imperfections in the long-range links are impressively mitigated by the derangement circuit.

4.6.2 Other applications

Using a modular architecture can also be useful in other scenarios. For instance, implementing the SWAP-test is natural for our architecture. It consists in preparing two independent states $|\psi\rangle_1$ and $|\psi\rangle_2$ and computing their overlap $|\langle\psi_1|\psi_2\rangle|^2$ using controlled-SWAP gates. Note that this is reminiscent of ESD/VD in the $N_c = 2$ case. It is an important building block for multiple quantum algorithms [182–188].

The multicore architecture is also naturally efficient for simulating problems with a Hamiltonian that is modular, meaning that it can be broken down in subsystems.

Numerous physical problems of interest from quantum chemistry [189–191] to many-body physics [192, 193] have this property. While the compilation onto hardware with specific topology might be challenging [194–196], previous studies have explored whether such a modular architecture can allow the simulation of large and complex quantum systems [197–199].

Finally, our multicore architecture could be used to implement quantum error correction codes [141, 164, 200]. For instance, one could create 'logical modules' by letting each core implement an error correcting code and then use the long-range links to implement another code on top.

4.7 Discussion

In this chapter, I explore a multicore architecture for silicon-based quantum computers and its applications. With this type of qubit, one could create multiple independent quantum cores (with a small footprint) at a low cost by leveraging the current semiconductor industry. As the inter-core links might be intrinsically more noisy than the intra-core processes, we abandoned the idea of linking remote qubits through the implementation of direct two-qubit unitaries. Instead my co-authors and I decided to distribute Bell pairs between the processors using 'weak' and noisy entangling links. I found that the quality of these raw Bell Pair is not sufficient to be useful, but could be upgraded using powerful entanglement purification protocols.

After briefly reviewing experimental works on 'quantum links', we identified shuttling and cavity-mediated interactions as the two most promising options for our architecture. Through detailed numerical simulations my co-authors and I have explored the advantages and drawbacks of both techniques. Using realistic parameters for the different noise processes, our numerical simulations show that both shuttling and cavities (with the help of entanglement purification) allow for the generation of Bell pairs between quantum cores with fidelities about 99.5%. Our simulations also confirmed that we could generate near-perfect Bell pairs using an increased number of entanglement purification rounds and approach error levels of local, intra-core operations. However, this comes at a smaller average generation

rate. While I have used the best reported experimental values for relaxation times, gate and measurement times, there is no single experiment that would achieve all those values at once. Experiments are constantly evolving and we expect these parameters to improve in the future.

Amongst the multiple applications that are well-suited for a multicore architecture, my co-authors studied in detail ESD/VD [171, 172] techniques which are particularly relevant as they require N_c separate cores performing the same quantum computation independently. The different cores are then entangled through the derangement operation, which can be efficiently implemented using our long-range links. This allows to reduce the errors exponentially with N_c.

By numerically simulating a VQE solving the ground state of a spin-ring Hamiltonian, we confirm that even with Bell pairs of fidelity 99.5% we can drastically suppress the errors in estimating the ground state energy due to noisy quantum circuits. Note that a major advantage of implementing ESD/VD with our architecture is that the long-range links are formally part of the main computation and as such their imperfections are also mitigated.

To conclude, the ideas presented here can easily be transposed to other types of quantum computers. Distributing and purifying Bell pairs between distant cores would enable powerful error mitigation schemes as well as other applications, extending the reach of near-term devices.

5

Introduction to quantum error mitigation and classical shadows

This chapter introduces some concepts necessary for the understanding of Chapter 6. The presentation follows some parts of my paper [201].

Contents

5.1 Quantum error mitigation 94

5.1.1 Probabilistic error cancellation 94

5.1.2 Error extrapolation 97

5.1.3 Symmetry verification 99

5.2 Classical shadows 100

5.2.1 Classical shadows with idealised measurements 101

5.2.2 Classical shadows with noisy measurements 102

5.1 Quantum error mitigation

5.1.1 Probabilistic error cancellation

Introduced by [202], Probabilistic Error Cancellation (PEC) is a powerful error mitigation technique enabling the recovery of noise-free expectation values from a noisy quantum device. PEC has been widely studied [133, 203, 204] and has already been implemented experimentally [133, 205–207]. The key idea behind PEC is that one can decompose an ideal gate $\mathcal{U}$ as a linear combination of noisy ones $\mathcal{G}_k$ (that we can implement physically) such that $\mathcal{U} = \sum_k \gamma_k \mathcal{G}_k$ (note that we use a channel representation for the gates). This is a quasiprobability decomposition rather than a simple probabilistic decomposition, as some γ_k can be negative. This originates from the fact that this decomposition corresponds to inverting a noise channel, which is not a physical operation here. Being non-physical, we can only apply this decomposition at the classical post-processing step when computing the expectation values.

Intuition

For ease of presentation, I will explain PEC using a simple example involving a circuit comprising a single noisy gate before generalising to more complex circuits. Considering a noisy quantum state $\rho = \mathcal{U}_q |0\rangle\langle 0|$ and an observable O, the aim is to retrieve the ideal expectation value $\mathrm{Tr}[O\rho_{id}] = \mathrm{Tr}[O\mathcal{U} |0\rangle\langle 0|]$.

In practice, we have access to a noisy version $\mathcal{U}_q$ of the ideal operation $\mathcal{U}$ which can generally be decomposed as,

$$\mathcal{U}_q = (1-q)\mathcal{U} + q\mathcal{N}, \tag{5.1}$$

with q the gate error rate and $\mathcal{N}$ representing the noise channel. It is thus straightforward to write a decomposition of $\mathcal{U}$ by inverting Eq. (5.1),

$$\mathcal{U} = \frac{1}{1-q}\mathcal{U}_q - \frac{q}{1-q}\mathcal{N}. \tag{5.2}$$

Notice the negative coefficient in the decomposition. We now need to decompose $\mathcal{N}$ in terms of gates that we can implement physically. Generally, the noise can be

modeled as Pauli noise and can thus be decomposed in terms of Pauli gates as,

$$\mathcal{N} = \sum_{k>1} p_k \mathcal{P}_k, \tag{5.3}$$

with $\sum_{k>1} p_k = 1$. Providing that we can physically implement high-quality Pauli gates, we obtain the quasiprobability decomposition of $\mathcal{U}$ in terms of physical operations $\mathcal{G}_k$,

$$\gamma_1 = \frac{1}{1-q},\ \mathcal{G}_1 = \mathcal{U}_p \quad \text{and} \quad \forall k > 1,\ \gamma_k = -\frac{q}{1-q} p_k,\ \mathcal{G}_k = \mathcal{P}_k. \tag{5.4}$$

Note that $\sum_k |\gamma_k| = 1$.

Using the quasiprobability decomposition and the linearity of the trace we get,

$$\begin{aligned} \mathrm{Tr}[O\rho_{id}] &= \mathrm{Tr}[O\mathcal{U}\,|0\rangle\langle 0|] \\ &= \mathrm{Tr}[O\left(\sum_k \gamma_k \mathcal{G}_k\right)|0\rangle\langle 0|] \\ &= \sum_k \gamma_k \mathrm{Tr}[O\mathcal{G}_k\,|0\rangle\langle 0|] \\ &= \sum_k \mathrm{sign}(\gamma_k)|\gamma_k| \mathrm{Tr}[O\mathcal{G}_k\,|0\rangle\langle 0|]. \end{aligned} \tag{5.5}$$

Eq. (5.5) has a clear operational meaning: with probability $|\gamma_k|$ prepare the state $\mathcal{G}_k\,|0\rangle\langle 0|$ and compute $\mathrm{sign}(\gamma_k)\mathrm{Tr}[O\mathcal{G}_k\,|0\rangle\langle 0|]$. The ideal expectation value is then obtained by taking the sum of these expectation values. In practice, this is performed using Monte-Carlo sampling and we give the measurement overhead introduced by PEC [133, 202, 208] below.

One might wonder how to determine each gate's decomposition into physical gates $\mathcal{G}_k$ in general. This problem has been tackled in previous works [133, 209–212] and we will suppose thereafter that such a decomposition is known with high precision. The cost of learning the decomposition (which is only performed once) is negligible compared to the cost of error mitigation and is thus disregarded.

Generalisation to more complex circuits

Consider the case of a circuit $\mathcal{U}_{circ}$ composed of ν gates, such that $\mathcal{U}_{circ} = \mathcal{U}_\nu \circ \cdots \circ \mathcal{U}_2 \circ \mathcal{U}_1$. Decomposing each ideal gate $\mathcal{U}_i = \sum_{k_i} \gamma_{k_i}^{(i)} \mathcal{G}_{k_i}^{(i)}$ and introducing the

vector $\underline{k} = (k_1, k_2, \dots k_\nu)$ we have,

$$\mathcal{U}_{circ} = \sum_{\underline{k}} \gamma_{\underline{k}} \mathcal{G}_{\underline{k}}, \tag{5.6}$$

with $\gamma_{\underline{k}} = \gamma^{(1)}_{k_1}\gamma^{(2)}_{k_2}\cdots\gamma^{(\nu)}_{k_\nu}$, $\mathcal{G}_{\underline{k}} = \mathcal{G}^{(\nu)}_{k_\nu} \circ \cdots \circ \mathcal{G}^{(2)}_{k_2} \circ \mathcal{G}^{(1)}_{k_1}$, the superscript representing the gate in the circuit and the vector $\underline{k}$ spanning all possible gate sequences. Let us now define the quasiprobability decomposition of a circuit.

Definition 1. *We define the quasiprobability decomposition of an ideal circuit $\mathcal{U}_{circ}$ as the set $\Gamma \coloneqq \{(\gamma_{\underline{k}}, \mathcal{G}_{\underline{k}})\}_{\underline{k}}$. To every decomposition Γ we associate the probability distribution $p_{\underline{k}} \coloneqq |\gamma_{\underline{k}}|/\|\gamma\|_1$. Note that the norm factorises as $\|\gamma\|_1 = \prod_{i=1}^{\nu} \left\|\gamma^{(i)}\right\|_1$.*

Note also that for any probabilistic noise model, if when applying the ith gate an error occurs with probability p_i then $\left\|\gamma^{(i)}\right\|_1 = (1+p_i)/(1-p_i)$ [133] (as in Eq. (5.2)). We thus have $\|\gamma\|_1 = \prod_{i=1}^{\nu} \left\|\gamma^{(i)}\right\|_1 = \prod_{i=1}^{\nu}(1+p_i)/(1-p_i) = \mathcal{O}(e^{2\xi})$ with $\xi = \sum_i p_i$ being the expected number of errors in the circuit. Similarly as in the previous subsection, the ideal expectation value of an observable O is obtained by sampling $\text{sign}(\gamma_{\underline{k}})\text{Tr}[O\mathcal{G}_{\underline{k}}\,|0\rangle\langle 0|]$ according to $p_{\underline{k}}$ and linearly-combining them in post-processing [133, 203, 204].

Cost of probabilistic error cancellation

To estimate the cost (or sampling overhead) C_{PEC} of PEC as defined in Appendix D.1, we first reformulate it in terms of statistical estimators.

Lemma 1. *Given a quasiprobability decomposition Γ of the ideal circuit $\mathcal{U}_{circ}$, an unbiased estimator of the ideal expectation value $o_{id} = \text{Tr}[O\mathcal{U}_{circ}\,|0\rangle\langle 0|]$ is given by,*

$$\hat{o}_{id} \coloneqq \hat{o}_{\underline{k}} = \|\gamma\|_1 \text{sign}(\gamma_{\underline{k}})\text{Tr}[O\mathcal{G}_{\underline{k}}\,|0\rangle\langle 0|] \tag{5.7}$$

in that $\mathop{\mathbb{E}}_{\underline{k}}[\hat{o}_{id}] = o_{id}$.

The proof of Lemma 1 follows straightforwardly from the one of Lemma 2 given in Appendix D.2. The cost C_{PEC} then reads,

$$C_{\text{PEC}} = \frac{\text{Var}[\hat{o}_{id}]}{\text{Var}[\hat{o}]} = \mathcal{O}(\|\gamma\|_1^2) = \mathcal{O}(e^{4\xi}) \tag{5.8}$$

with $\hat{o}$ being the estimator obtained by directly measuring the expectation value on ρ. PEC is thus impractical when $\xi \gg 1$ [133, 202, 208].

Generalisation to quantum states

In reality, PEC can do more than just providing an unbiased estimator for the ideal expectation value. It also yields an unbiased estimator of the ideal state itself.

Lemma 2. *Given a quasiprobability decomposition* Γ *of the ideal circuit* $\mathcal{U}_{circ}$*, an unbiased estimator of the ideal state* ρ_{id} *is given by,*

$$\hat{\rho}_{id} \coloneqq \hat{\rho}_{\underline{k}} = \|\gamma\|_1 \text{sign}(\gamma_{\underline{k}}) \mathcal{G}_{\underline{k}} |0\rangle\langle 0| \tag{5.9}$$

in that $\mathbb{E}[\hat{\rho}_{id}] = \mathbb{E}_{\underline{k}}[\hat{\rho}_{\underline{k}}] = \rho_{id}$.

The proof is given in Appendix D.2. Unfortunately, as the quasiprobabilities $\gamma_{\underline{k}}$ are negative, averaging over the obtained density matrices can only be performed classically in post-processing. This would require the complete reconstruction of each density matrices, which is impractical. This limitation disappears when leveraging a compact representation of quantum states such as classical shadows as discussed in Section 6.2.

5.1.2 Error extrapolation

One major disadvantage of PEC is that it requires a lot of resources. Disregarding the cost of learning the quasiprobabilities, one has to run the quantum computation with different circuits which might be demanding in practice. When PEC cannot be performed due to hardware restrictions, a more intuitive and less demanding error mitigation technique called zero-noise extrapolation (ZNE) can be used [133, 202, 208, 213]. It is important to note that contrary to PEC, ZNE yields a biased estimator. ZNE is based on the ability to *boost* the noise in the quantum computation and extrapolate the expectation values back to zero noise.

Intuition

Consider the quantum state ρ prepared using a noisy quantum computer. This state can be viewed as a function of the circuit error rate ξ as defined in Section 5.1.1 and we make this dependence clear by writing $\rho = \rho_\xi$. When ξ is small, which is

the regime in which quantum error mitigation is relevant, we can Taylor-expand $\mathrm{Tr}[O\rho_\xi]$ as follows [202],

$$\mathrm{Tr}[O\rho_\xi] = \mathrm{Tr}[O\rho_0] + \sum_{j=1}^{n-1} c_j \frac{\xi^j}{j!} + \mathcal{O}(\xi^n). \tag{5.10}$$

Let ξ_1 be the circuit error rate of the quantum computer used to prepare ρ. By directly estimating $\mathrm{Tr}[O\rho_{\xi_1}]$, we make an error on the order of $\mathcal{O}(\xi)$. In order to mitigate this error, Ref. [202] used Richardson extrapolation which consists in estimating $\mathrm{Tr}[O\rho_\xi]$ for boosted error rates $\{\xi_1, \dots \xi_n\}$ with $\forall i < j,\ \xi_i < \xi_j$ and linearly combining them as $\sum_{l=1}^{n} \mathrm{Tr}[O\rho_{\xi_l}] \prod_{m\neq l} \xi_m/(\xi_m - \xi_l)$ such that,

$$\mathrm{Tr}[O\rho_0] = \sum_{l=1}^{n} \mathrm{Tr}[O\rho_{\xi_l}] \prod_{m\neq l} \frac{\xi_m}{\xi_m - \xi_l} + \mathcal{O}(\xi^n). \tag{5.11}$$

The error mitigated estimator $\hat{o}_{\mathrm{ZNE}}$ constructed from Eq. (5.11) by directly measuring $\mathrm{Tr}[O\rho_{\xi_l}]$ from the quantum computer for every noise level ξ_l and combining them accordingly yields a biased estimator, with a bias on the order of $\mathcal{O}(\xi^n)$ compared to the naive $\mathcal{O}(\xi)$. Note that as we will always be using the best error rate to run a computation, we can only *boost* errors.

Cost of zero-noise extrapolation

The cost C_{ZNE} of ZNE is given by [133],

$$C_{\mathrm{ZNE}} = \frac{\mathrm{Var}[\hat{o}_{\mathrm{ZNE}}]}{\mathrm{Var}[\hat{o}]} \approx \left(\sum_{l=1}^{n} \left| \prod_{m\neq l} \frac{\xi_m}{\xi_m - \xi_l} \right| \right)^2. \tag{5.12}$$

Note that if one of the error rates is too big or if one of the gaps between two error rates is too small C_{ZNE} blows up. To give a more quantitative understanding of how C_{ZNE} scales, suppose that we set a fixed gap between the error rates e.g. $\xi_l = l\xi_1$. We then get $C_{\mathrm{ZNE}} \approx (2^n - 1)^2$ which scales exponentially with the number of data points. Thereafter, we will only focus on the $n = 2$ case which corresponds to a linear extrapolation. It is worth noting that extrapolation techniques beyond Richardson extrapolation have been explored to address the situations where ξ is not small [133].

5.1.3 Symmetry verification

Intuition

Another way to mitigate errors in the preparation of a state ρ on a noisy quantum computer is to leverage its symmetries [133, 214, 215]. Usually, the problem we are trying to solve and its implementation on a quantum computer is such that we know that certain symmetries must be present in ρ_{id}. For instance, when using a second quantisation representation of an N-particle state on a register of qubits [216], the number of particles N corresponds to the Hamming weight of the register and we know that any state of weight other than N cannot occur.

More precisely, suppose that ρ_{id} obeys the symmetries S from the symmetric group $\mathbb{S}$. To put it differently, if we consider the projector onto the symmetric subspace $\Pi_{\mathbb{S}} = 1/|\mathbb{S}| \sum_{S\in\mathbb{S}} S$, then ρ_{id} verifies,

$$\Pi_{\mathbb{S}}\rho_{id}\Pi_{\mathbb{S}} = \rho_{id}. \tag{5.13}$$

Projecting the state ρ in the symmetric subspace $\mathbb{S}$ prior to measuring its expectation value with respect to O will thus remove the errors which lead to a violation of the symmetries. This can be performed using the Hadamard-test and by post-selecting the runs respecting the symmetries [133, 217] which effectively generate the state

$$\rho_{symm} = \frac{\Pi_{\mathbb{S}}\rho\Pi_{\mathbb{S}}}{\mathrm{Tr}[\Pi_{\mathbb{S}}\rho]}. \tag{5.14}$$

This technique is called symmetry verification (SV). The expectation value $\mathrm{Tr}[O\rho_{symm}]$ can then be expanded as follows,

$$\mathrm{Tr}[O\rho_{symm}] = \frac{\mathrm{Tr}[O\Pi_{\mathbb{S}}\rho\Pi_{\mathbb{S}}]}{\mathrm{Tr}[\Pi_{\mathbb{S}}\rho]} = \frac{1}{|\mathbb{S}|}\frac{\sum_{S,S'\in\mathbb{S}}\mathrm{Tr}[SOS'\rho]}{\sum_{S\in\mathbb{S}}\mathrm{Tr}[S\rho]}. \tag{5.15}$$

While this decomposition has been derived using the post-selected state ρ_{symm}, note that it gives us another way to construct $\mathrm{Tr}[O\rho_{symm}]$. By directly estimating $\mathrm{Tr}[SOS'\rho]$ and $\mathrm{Tr}[S\rho]$ for $S, S' \in \mathbb{S}$, and linearly combining them as in Eq. (5.15), we construct $\hat{o}_{\mathrm{SV}}$ in post-processing. A big advantage of this method is that it does not require the implementation of in-circuit symmetry verification (e.g. Hadamard-test).

It is important to note that SV does not correct for errors that commute with the $S \in \mathbb{S}$ and will generally yield a biased estimator [133].

Cost of symmetry verification

The cost of SV when using post-selection and post-processing is different. When using post-selection, the cost C_{SV} is given by [133]

$$C_{\mathrm{SV}} = \frac{\mathrm{Var}[\hat{o}_{SV}]}{\mathrm{Var}[\hat{o}]} \approx \frac{1}{\mathrm{Tr}[\Pi_{\mathbb{S}}\rho]}, \tag{5.16}$$

where $\mathrm{Tr}[\Pi_{\mathbb{S}}\rho]$ can be seen as the success probability of the in-circuit symmetry verification. When using post-processing, the sampling overhead is larger and is given by,

$$C_{\mathrm{SV}} = \frac{\mathrm{Var}[\hat{o}_{SV}]}{\mathrm{Var}[\hat{o}]} \approx \frac{1}{\mathrm{Tr}[\Pi_{\mathbb{S}}\rho]^2}. \tag{5.17}$$

5.2 Classical shadows

In 2020, inspired by Aaronson's work [218], Huang et al. introduced classical shadows which allows one to classically store and extract properties of a quantum state ρ without going through the computationally expensive task of fully reconstructing the state [219]. Given an N-qubit quantum state ρ, it consists in applying a random unitary Q_j sampled from an ensemble $\mathcal{Q}$ and performing a measurement in the computational basis. In practice, $\mathcal{Q}$ is composed of single-qubit rotations (Pauli measurements) or Clifford unitaries. Let $\underline{b} = \{0,1\}^N$ be the outcome of the measurement in the computational basis, and $l = (j, \underline{b})$ an index storing both the random unitary and the result of the measurement.

Definition 2. *We define the classical shadow of quantum state ρ for a given ensemble of unitaries $\mathcal{Q}$ as the collection of indices $l = (j, \underline{b})$ obtained by running multiple independent executions of shadow tomography on ρ.*

Classical shadows changes the paradigm with which one obtains properties of quantum states. With classical shadow we adopt a 'randomly measure and extract properties later' attitude, rather than performing bespoke quantum measurements for each expectation value we would like to know. However, depending on the observable, the number of measurements needed to estimate the expectation value to a given accuracy can become prohibitive as we will see later.

5.2.1 Classical shadows with idealised measurements

Here, I show how shadows can be used to construct an estimator of ρ. I will adopt the formalism used by [220] rather than the original one developed by [219] as it allows us to easily mitigate errors in the measurement step of the protocol as I will discuss in Section 5.2.2.

Suppose that after running the protocol once, we obtain the index $l = (j, \underline{b})$. This measurement outcome can be represented using the positive operator $E_l = p_j Q_j^\dagger |\underline{b}\rangle\langle\underline{b}| Q_j$, with p_j the probability of drawing Q_j from $\mathcal{Q}$. This outcome occurs with probability $q_l = \mathrm{Tr}[E_l \rho]$. Outcomes of the shadow protocol are thus completely described by the collection of positive operators E such that,

$$E = \{E_l = p_j Q_j^\dagger |\underline{b}\rangle\langle\underline{b}| Q_j \mid Q_j \in \mathcal{Q},\ \underline{b} \in \{0,1\}^N\}, \tag{5.18}$$

with $N_E = |E| = |\mathcal{Q}| 2^N$. The astute reader would have recognised that providing that $\sum_{l=1}^{N_E} E_l = \mathbb{1}$, E represents a generalised measurement or positive operator-valued measure (POVM) [1]. The E_l are called the effects of E. Running the classical shadow protocol on a state ρ can thus be seen as applying to ρ the channel C_E such that,

$$C_E(\rho) = \sum_{l=1}^{N_E} \mathrm{Tr}[E_l \rho] E_l. \tag{5.19}$$

This allows us to construct an unbiased estimator of the state ρ.

Theorem 1 (Classical shadows)**.** *A classical shadow protocol with the POVM measurement E as in Eq.* (5.18) *yields an unbiased estimator of ρ,*

$$\hat{\rho}_l \coloneqq C_E^{-1}(E_l), \tag{5.20}$$

in that $\mathop{\mathbb{E}}_l[\hat{\rho}_l] = \rho$, providing that E spans the whole space of observables.

The proof of Theorem 1 is given in Appendix D.3. In practice, one has to generate a collection of snapshots $S(\rho_{id}, N_S) = \{\hat{\rho}_1, \hat{\rho}_2, \ldots \hat{\rho}_{N_s}\}$ and take their average to estimate ρ. Thereafter we will refer interchangeably to the collections $\{l_1, l_2, \ldots l_{N_s}\}$ and $S(\rho_{id}, N_S)$ as the classical shadows of ρ of size N_s.

It is important to note that when the POVM E consists in individual measurements on individual qubits (i.e. $Q_j = \otimes_{i=1}^{N} Q_j^{(i)}$), $\hat{\rho}_l$ can be decomposed as a tensor product of single-qubit density matrices. This is a big advantage offered by shadows as, rather than storing the full $2^N \times 2^N$ density matrix, one can get away with storing N 2×2 matrices. In the practical case of single-qubit random Pauli measurements the decomposition of $\hat{\rho}_l$ reads [219],

$$\hat{\rho}_l = \bigotimes_{i=1}^{N} \left[3(Q_j^{(i)})^\dagger \left| b^{(i)} \right\rangle\!\left\langle b^{(i)} \right| Q_j^{(i)} - \mathbb{1} \right], \tag{5.21}$$

with $Q_j^{(i)}$ the ith single-qubit unitary in the decomposition of Q_j, and $b^{(i)}$ the ith bit of $\underline{b}$.

While our formalism for quantum error mitigated shadows is general and can be implemented for different unitary ensembles $\mathcal{Q}$ (such as matchgates [221], Clifford circuits [219], and more [222]), I will focus on the practical case of single-qubit Pauli measurements.

5.2.2 Classical shadows with noisy measurements

As mentioned in the previous section, using the POVM formalism in classical shadows allows one to directly mitigate measurement errors [220]. Here, I focus on a simple readout error model in which the outcome of a measurement $\underline{b}$ is flipped with some probability, while noting that for correlated readout errors or errors in the random basis transformations techniques developed in Refs. [133, 223, 224] can be incorporated to our scheme.

Suppose that when obtaining the measurement outcome $\underline{b}$, the ith bit $b^{(i)}$ is flipped with probability α_i^+ (resp. α_i^-) if $b^{(i)} = 0$ (resp. $b^{(i)} = 1$). The noiseless measurement operators for the ith qubit $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$ are thus replaced by $(1 - \alpha_i^+) |0\rangle\langle 0| + \alpha_i^- |1\rangle\langle 1|$ and $(1-\alpha_i^-) |1\rangle\langle 1| + \alpha_i^+ |0\rangle\langle 0|$ respectively. This transformation only modifies the effects E_l of the POVM and as the new C_E remains invertible, the measurement errors are completely removed by inverting C_E to generate the snapshots $\hat{\rho}_l$.

In our case of single-qubit random Pauli rotations, $\hat{\rho}_l$ is now given by,

$$\hat{\rho}_l = \bigotimes_{i=1}^{N} \left[\frac{3}{1-2\alpha_i} (Q_j^{(i)})^\dagger \left| b^{(i)} \right\rangle\!\!\left\langle b^{(i)} \right| Q_j^{(i)} - \frac{1+\alpha_i}{1-2\alpha_i} \mathbb{1} \right], \tag{5.22}$$

supposing that $\alpha_i^+ = \alpha_i^- = \alpha_i$. In the more realistic case when $\alpha_i^+ \neq \alpha_i^- = \alpha_i$, the scheme is still valid but the snapshots have to be computed by numerically inverting (rather than analytically) the channel defined in Eq. (5.19).

6

Quantum error mitigated classical shadows

This chapter presents results from [201] published in PRX Quantum. In Section 6.2 and Section 6.3, I present the contribution of my co-authors for context, before presenting my contribution in Section 6.4. This chapter closely follows [120] and I used text verbatim where I am an original author. The proofs of the different lemmas and theorems done by my co-authors Jonathan Steinberg and Bálint Koczor are given in Appendix E for completeness.

Contents

6.1 Introduction 105
6.2 Probabilistic error cancelled shadows 106
6.2.1 Protocol 107
6.2.2 Rigorous performance guarantees 108
6.2.3 Post-processing algorithms computing expectation values 111
6.3 Further error mitigation techniques 113
6.3.1 Error boosted shadows 113
6.3.2 Symmetry verified shadows 114
6.4 Applications 115
6.4.1 Ground-state preparation 117
6.4.2 Error mitigated estimation of entanglement entropies . . 122
6.4.3 Further applications 123
6.5 Discussion 124

6.1 Introduction

As we saw in the previous chapters, quantum computers are extremely sensitive to noise. While quantum error correction would solve this issue, it remains a formidable engineering challenge [225]. However, it is believed by many in the community that even with noisy devices, one could reach early practical quantum advantage [133, 205–207]. Quantum error mitigation (which we briefly introduced in the previous chapter) will be a key enabler towards this goal [131, 133]. The multicore architecture we designed allows for efficient implementation of powerful error mitigation techniques (ESD/VD), thereby extending the reach of near-term devices. However, a major hurdle remains in achieving early quantum advantage: the large number of circuit runs needed to suppress shot noise [226–229]. To make things worse, error mitigation further increases this issue as the cost to reduce error is an increased number of measurements (see Appendix D.1). Nonetheless, one could imagine using multiple multicore processors, without linking them, to reduce this problem but this will require improved hardware.

A recently introduced technique called classical shadows [219] promises to alleviate this problem as explained in Chapter 5. The purpose of this chapter is to combine quantum error mitigation and classical shadows techniques in order to obtain noiseless properties from noisy quantum devices with few measurements.

Previous works have explored relationships between quantum error mitigation and classical shadows. On the one hand, refs. [230, 231] used classical shadows to perform ESD/VD [171, 172] with access to only a single copy of the state but with a complexity growing exponentially with the number of qubits. On the other hand, refs. [232, 233] explored the mitigation of errors in the randomised measurements. However, these works did not consider errors occurring during state preparation and focused on pre-established quantum states.

Here, I focus on the more practical case where the state ρ comes from a noisy quantum device. The goal then is to mitigate the errors introduced by the quantum gates and extract properties of the ideal quantum state ρ_{id} as generated by a noiseless quantum computer. This is reminiscent of error mitigation in the sense

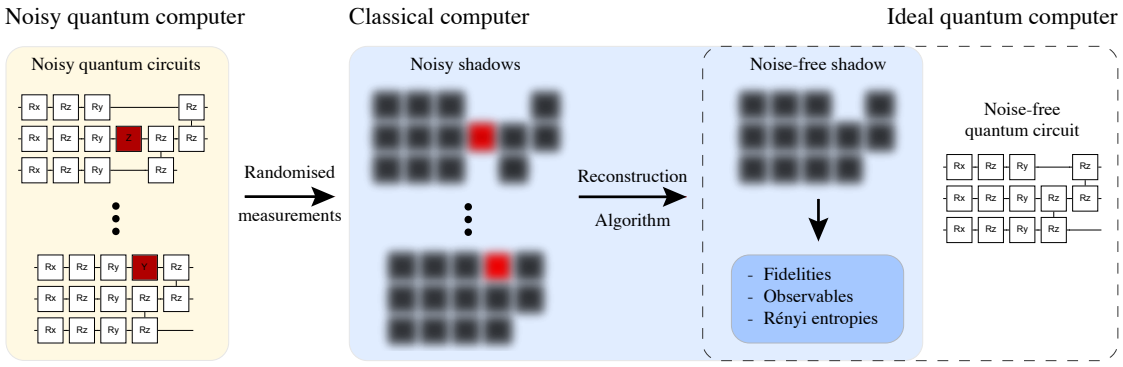

Figure 6.1: Here, we assume we only have access to a noisy quantum computer (left) such that every circuit we run (left, yellow area) gets corrupted by gate noise (unwanted red gate elements). We aim to extract properties of a state that would be prepared by an ideal quantum computer (right) with the use of powerful error mitigation techniques. We provide a rigorous theoretical framework for PEC shadows which effectively allows us to obtain a classical shadow of the ideal quantum state (noise-free shadow) from which we can predict many ideal properties in classical post-processing (middle blue area, classical computer). In our formalism we run a series of distinct quantum circuit variants (left, yellow area) that cast different classical shadows (noisy shadows) due to gate noise and due to our intentional recovery operations (red gate elements and their shadows). Under the assumption that the device's error characteristics have been appropriately learned, we can estimate the noise free shadow (middle) via classical post-processing.

that we extract ideal expectation values $\mathrm{Tr}[O\rho_{id}]$ from noisy ones $\mathrm{Tr}[O\rho]$. However, we are not limited to properties of the state here, we can generate a classical representation of the ideal state ρ_{id} itself from access to a noisy quantum computer as shown in Fig. 6.1. This representation is called a quantum error mitigated classical shadow of ρ or QEM shadow for short.

This chapter is organised as follows. In Section 6.2, I present our main result as Probabilistic Error Cancelled (PEC) shadows. While we find that PEC shadows is the most natural combination, we explore how to connect other error mitigation techniques and classical shadows in Section 6.3. Finally, in Section 6.4, I explore multiple applications of QEM shadows and conclude in Section 6.5.

6.2 Probabilistic error cancelled shadows

In this section, I show that by combining the two random processes at the core of PEC (random generation of the circuits $\mathcal{G}_{\underline{k}}$) and classical shadows (random choice

of the bases transformations and probabilistic measurement outcomes) one can efficiently construct and store a classical representation of the ideal state ρ_{id} using a noisy quantum computer (see Fig. 6.1 for an illustration).

6.2.1 Protocol

In Lemma 2, we saw that PEC can be used to construct an unbiased estimator of ρ_{id} by linearly combining the density matrices obtained by preparing the states $\|\gamma\|_1 \text{sign}(\gamma_{\underline{k}}) \mathcal{G}_{\underline{k}} |0\rangle\langle 0|$ sampled according to $p_{\underline{k}}$. However, that protocol is impractical due to the negative sign of the quasiprobabilities $\gamma_{\underline{k}}$ and the exponential cost associated with the reconstruction of the density matrices. Using classical shadows, we show that this protocol can be modified to obtain an unbiased estimator of ρ_{id} that can be classically stored and manipulated efficiently.

Theorem 2 (PEC Shadows)**.** *Given a quasiprobability decomposition* Γ *of the ideal circuit* $\mathcal{U}_{circ}$ *as defined in Definition 1 and a classical shadow protocol with the POVM measurement* E *as in Eq.* (5.18)*, we define PEC shadows as the set* $H \coloneqq \{(\gamma_{\underline{k}}, \mathcal{G}_{\underline{k}}, E_l)\}_{\underline{k},l}$*. H yields an unbiased estimator of* ρ_{id}*,*

$$\hat{\rho}_{id} \coloneqq \hat{\rho}_{\underline{k},l} = \|\gamma\|_1 \text{sign}(\gamma_{\underline{k}}) C_E^{-1}(E_l), \tag{6.1}$$

in that $\mathbb{E}[\hat{\rho}_{id}] = \mathbb{E}_{\underline{k},l}[\hat{\rho}_{\underline{k},l}] = \rho_{id}$*. Thereafter we will use the notation* $\hat{\rho}_{id}$ *instead of* $\hat{\rho}_{\underline{k},l}$*.*

See Appendix E.1 for a proof of Theorem 2. Note that contrary to the conventional classical shadows protocol, the measurement E is performed on the states $\mathcal{G}_{\underline{k}} |0\rangle\langle 0|$ rather than directly on ρ. We replaced the large density matrices $\mathcal{G}_{\underline{k}} |0\rangle\langle 0|$ in Lemma 2 by a snapshot $\hat{\rho}_l = C_E^{-1}(E_l)$ which has a compact classical representation. The PEC shadows protocol goes as follows,

- draw a vector $\underline{k}$ according to $p_{\underline{k}}$ and store $\text{sign}(\gamma_{\underline{k}})$.
- draw a unitary $Q_j \in \mathcal{Q}$ uniformly at random.

- prepare the state $\mathcal{G}_{\underline{k}} |0\rangle\langle 0|$, apply the basis transformation Q_j, and measure in the computational basis. Store the outcome of the measurement $\underline{b}$ and the index j as $l = (j, \underline{b})$.
- the tuple $(\underline{k}, l)$ uniquely corresponds to the state $\hat{\rho}_{\underline{k},l} = \|\gamma\|_1 \text{sign}(\gamma_{\underline{k}}) C_E^{-1}(E_l)$
- repeat the process N_s times to generate a set of shadows $S(\rho_{id}, N_S)$ of the ideal state ρ_{id}.

While we showed that we can construct the shadows of the ideal state ρ_{id} from a noisy quantum device in Theorem 2, in practice we are more interested by properties of ρ_{id} rather than the full density matrix itself. A powerful feature of conventional shadows is that it allows one to simultaneously predict *multiple properties* of a state from very few snapshots [219]. Hence, as PEC shadows gives us shadows of the ideal state, we can recover ideal properties from a noisy quantum computer. In Section 6.2.2, we present rigorous performance guarantees of PEC shadows for the practical task of estimating expectation values and show that they consists in slight modifications of those of conventional shadows. Then, we will present post-processing algorithms to compute expectation values in different scenarios in Section 6.2.3

6.2.2 Rigorous performance guarantees

Given a PEC snapshot $\hat{\rho}_{id}$, the error-mitigated expectation value of an observable O can be computed via the estimator $\hat{o} = \text{Tr}[O\hat{\rho}_{id}]$. As explained in Section 5.1, using error mitigation introduces a sampling overhead. It is thus necessary to characterise the cost associated with PEC shadows to assess the practicality of our scheme.

Lemma 3 (Variance of linear properties)**.** *Given an observable O and a PEC snapshot $\hat{\rho}_{id}$, the variance of the estimator $\hat{o} = \text{Tr}[O\hat{\rho}_{id}]$ is bounded by,*

$$\text{Var}[\hat{o}] \leq \|\gamma\|_1^2 \|O\|_E^2, \tag{6.2}$$

with $\|\gamma\|_1$ defined in Definition 1 and $\|O\|_E^2$ is the shadow norm of O defined in Lemma 5 which also depends on the set of random basis transformations $\mathcal{Q}$.

The proof is given in Appendix E.1.3. For conventional shadows [219], we have $\mathrm{Var}[\hat{o}] \leq \|O\|_E^2$. Notice that the increase in variance due to the combination of shadows and PEC is similar to that of PEC itself (see Eq. (5.8)). Hence, in a similar way to PEC, PEC shadows are impractical when the circuit error rate $\xi \gg 1$. In the practical case in which $\mathcal{Q}$ corresponds to single-qubit Pauli basis transformation, and O is a q-local Pauli string then $\|O\|_E^2 = 3^q$ (remember that any observable O can be written as a linear combination of Pauli strings). A proof is given in Appendix E.1.2. While single-qubit random Pauli basis transformations are experimentally friendly, they yield a shadow norm that grows exponentially with the locality of O and thus leads to a variance that could blow up. Although other sets $\mathcal{Q}$ lead to more favorable scaling [219], we decide to stay within the random single-qubit Pauli transformations case and limit ourselves to observables of small locality. See [201] for further discussion on the shadows norms. For the simple readout error model presented in Section 5.2.2, it is straightforward to show that the shadow norm goes from $\|O\|_E^2 = 3^q$ to $\|O\|_E^2 = 3^q(1-2\alpha)^{-2q}$ (use the new snapshot defined in Section 5.2.2 in the proof presented in Appendix E.1.2).

In ref. [219], Huang et al. showed that using classical shadows we can simultaneously estimate *multiple properties* of a quantum state using very few measurements. This is done by first splitting the shadows $S(\rho_{id}, N_s)$ into independent batches, then computing the mean expectation value for each of them and returning the median of the independent means. Using the concentration properties of this estimator, they demonstrated that to compute the expectation values of M observables, we need to generate a collection of shadows of size N_s with N_s only growing logarithmically with M. Using a similar approach, my co-author Bálint Koczor derived an expression for the number of samples needed for PEC shadows to estimate the expectation values of M observables to an accuracy ε with failure probability δ.

Theorem 3 (Informal version of Theorem 4)**.** *Given the PEC shadows* $H \coloneqq \{(\gamma_{\underline{k}}, \mathcal{G}_{\underline{k}}, E_l)\}_{\underline{k},l}$ *from Theorem 2, we wish to simultaneously estimate the expectation values of* M *observables* $O_1, \ldots O_M$ *with respect to* ρ_{id}*. Using a median of means*

estimator, the number of shots needed to obtain an accuracy $\varepsilon \in [0, 1]$ *with failure probability* δ *is,*

$$N_s = 32\varepsilon^{-2} \log\left(\frac{M}{\delta}\right) \|\gamma\|_1^2 \max_{1\leq j\leq M} \|O_j\|_E^2. \tag{6.3}$$

A more formal version of Theorem 3 alongside its proof are given in Appendix E.1.3.

Shadows can also be used to estimate non-linear properties using the fact that any polynomial function in the state ρ_{id} (e.g. $\mathrm{Tr}[O\rho_{id}^m]$) can be expressed as a linear function in tensor product of the state [172, 219, 234]. For instance, $\mathrm{Tr}[O\rho_{id}^2] = \mathrm{Tr}[\bar{O}\rho_{id} \otimes \rho_{id}]$ where $\bar{O} = \mathrm{SWAP}O_1$ with $O_1 = O \otimes \mathbb{1}$ (the subsystem on which O acts is not relevant). In Appendix E.1.3, I detail the construction of my co-author based on U-statistics [235] for estimating non-linear properties using PEC shadows. It yields the following performances.

Lemma 4 (Variance of non-linear properties)**.** *Given an observable* O *and a PEC shadows* H*, we wish to estimate polynomial properties of degree* m *of the ideal state* ρ_{id}*. Using U-statistics, the number of shots needed to obtain an accuracy* ε *is given by,*

$$N_s = O\left(\frac{\|\gamma\|_1^{2m}}{\varepsilon^2}\right). \tag{6.4}$$

The proof can be found in Appendix E.1.3. It is worth noting that the number of shots now grows as $\|\gamma\|_1^{2m}$ for a tensor product of m states compared to $\|\gamma\|_1^2$ for a single state. Using the median of means estimator, we can here also estimate multiple non-local properties.

6.2.3 Post-processing algorithms computing expectation values

Now that we have shown that PEC shadows can be used in order to estimate noiseless properties of a quantum state in theory, let us delve into their practical implementation. Focusing on the practical case of q-local Pauli observables O, we give reconstruction algorithms for the estimation of linear and non-linear properties of the ideal state ρ_{id}.

Algorithm 1 (Expectation value of local Paulis). Given a PEC snapshot $\hat{\rho}_{id}$ and a q-local Pauli observable O, the expectation value $\mathrm{Tr}[O\hat{\rho}_{id}]$ is given by,

$$\mathrm{Tr}[O\hat{\rho}_{id}] = 3^q \|\gamma\|_1^2 \mathrm{sign}(\gamma_{\underline{k}}) f(O, \underline{b}, Q_j). \tag{6.5}$$

Splitting the shadows $S(\rho_{id}, N_s)$ into K batches , the algorithm computes $\mathrm{Tr}[O\hat{\rho}_{id}]$ for each snapshot and returns the value given by the median of means estimator. The quantity $f(O, \underline{b}, Q_j)$ is equal to ± 1 if O and Q_j are compatible (meaning that the basis transformation and the observable are the same, e.g. X basis rotation for an X observable) and 0 if they are not. The sign of $f(O, \underline{b}, Q_j)$ is given by $\underline{b}$. The runtime of the algorithm is $\mathcal{O}(qN_s)$. A more detailed derivation is given in Appendix E.2.

In the specific case of estimating expectation values of q-local Pauli observables, we can further reduce the variance of our estimator by using light cone arguments [236, 237]. In Fig. 6.2, we represent the light cone of an observable O in green and notice that not every gate of the circuit lies within it. The gates outside the light cone are thus not affecting the value of O. We can then remove them from our PEC protocol which ultimately reduces the number of circuit variants $\mathcal{G}_{\underline{k}}$. Indeed, the fact that these gates are noisy does not have an impact on the expectation value of O, we then do not need to consider their decomposition in Eq. (5.6). It is important to note that we considered local noise model in order to obtain similar light cones for every circuit variant $\mathcal{G}_{\underline{k}}$, but our method will also be valid for non-local ones using arguments from [236].

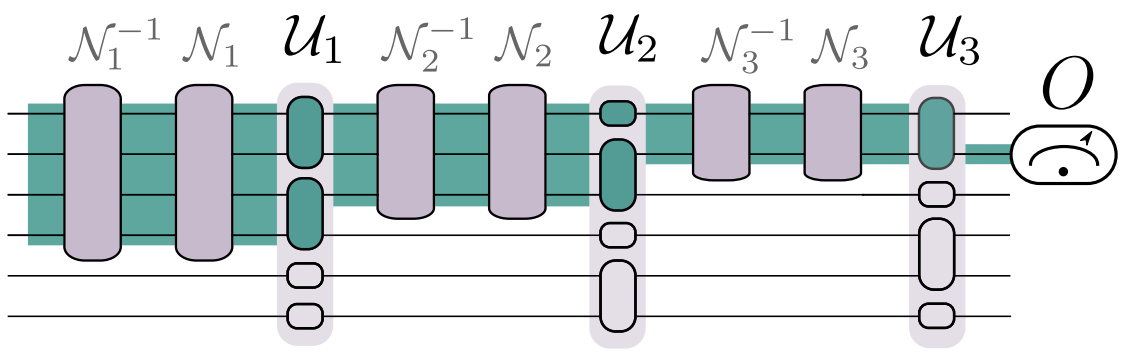

Figure 6.2: Illustration of the light cone of an observable O acting on the second qubit. The green area represents the light cone and we note $\mathcal{I}$ the set of qubits in the ideal circuit (green boxes) contained in it. For simplicity, we suppose that the noise channels $\mathcal{N}_k$ are local such that they do not extend the light cone to other qubits. Our results can be extended to non-local models using the results of [237]. Figure adapted from [201].

Algorithm 2 (Light cones). Given a q-local Pauli observable O, let $\mathcal{I}$ be the set of indices of gates lying within the observable's light cone such that $\mathcal{I} \coloneqq \{i \mid \mathcal{U}_i \text{ is in the light cone of } O\}$. Using Algorithm 1 with the new set of quasiprobabilities from Definition 1,

$$\|\tilde{\gamma}\|_1 = \prod_{i\in\mathcal{I}} \left\|\gamma^{(i)}\right\|_1 \quad \text{and} \quad \operatorname{sign}(\tilde{\gamma}_{\underline{k}}) = \prod_{i\in\mathcal{I}} \operatorname{sign}(\gamma^{(i)}_{k_i}). \tag{6.6}$$

still yields an unbiased estimator of $\operatorname{Tr}[O\rho_{id}]$. The complexity of this algorithm scales also as $\mathcal{O}(qN_s)$ and the preprocessing step to determine the set $\mathcal{I}$ for every observable O has a negligible cost. A more detailed derivation is given in Appendix E.2.

The new sampling cost given by $\|\tilde{\gamma}\|_1$ depends now on the number of gates in the light cone of the observable $|\mathcal{I}|$ rather than the number of gates in the circuit ν. For instance suppose that $\left\|\gamma^{(i)}\right\|_1 = (1+p)/(1-p)$, then $\|\tilde{\gamma}\|_1 = ((1+p)/(1-p))^{|\mathcal{I}|}$ compared to the worst-case $\|\gamma\|_1 = ((1+p)/(1-p))^{\nu}$. Using this modified algorithm we may drastically reduce the exponent. Note that we have not modified the experimental part of PEC shadows, the huge advantage of this trick is that we only have to modify the postprocessing step to obtain a reduced variance.

Finally, for non-local properties we will focus on Rényi entropies R_Q defined using the purities $\operatorname{Tr}[\rho^2_{id,Q}]$, such that $R_Q \coloneqq -\log \operatorname{Tr}[\rho^2_{id,Q}]$ with $\rho_{id,Q}$ being the ideal reduced density matrix on the subsystem Q.

Algorithm 3 (Local purities). Given a PEC snapshot $\hat{\rho}_{id}$ and a subsystem Q given as the set of qubits $\{q_1, \dots q_{|Q|}\}$, the local purity of $\rho_{id,Q}$ is given by,

$$\mathrm{Tr}[\rho_{id,Q}^2] = \|\gamma\|_1^2 \mathrm{sign}(\gamma_i)\mathrm{sign}(\gamma_j) f(i,j,Q). \tag{6.7}$$

with i, j being snapshot indices e.g. $i = (\underline{k}, l)$ and $i \neq j$. Splitting the shadows $S(\rho_{id}, N_s)$ into K batches, the algorithm computes $\mathrm{Tr}[\rho_{id,Q}^2]$ for each pair of snapshots and returns the value given by the median of means estimator. The quantity $f(i,j,Q)$ only depends on whether the measurement bases and outcome bitstrings are identical for the subsystem Q. The runtime of the algorithm is $\mathcal{O}(|Q|N_s^2)$. A more detailed derivation is given in Appendix E.2.

6.3 Further error mitigation techniques

In this section, we present how shadows can be combined with the two other error mitigation techniques presented in Section 5.1.

6.3.1 Error boosted shadows

In Section 5.1.2, I presented zero noise extrapolation, an error mitigation technique consisting in extrapolating the ideal expectation value $\mathrm{Tr}[O\rho_{id}]$ from noisy ones $\{\mathrm{Tr}[O\rho_{\xi_i}]\}_i$ with the $\{\xi_i\}_i$ being different circuit error rates such that if $i < j$, $\xi_i < \xi_j$. Here, we generate multiple sets of shadows at different circuit error rate $S(\rho_{\xi_1}, N_s), \dots, S(\rho_{\xi_n}, N_s)$, extract from them the $\{\mathrm{Tr}[O\rho(\xi_i)]\}_i$ and using a model function $f(\xi)$ for $\mathrm{Tr}[O\rho_\xi]$ extrapolate to $\xi = 0$.

Rather than using standard techniques to *boost* the error rate of gates [133, 205], we leverage the formalism introduced in PEC Shadows. In a similar fashion to the way in which we considered the quasiprobability decomposition of the ideal circuit $\mathcal{U}_{circ}$ in Definition 1, we can decompose the noise-boosted circuits as $\mathcal{G}_{circ}(p) = \sum_{\underline{k}} g_{\underline{k}}(p)\mathcal{G}_{\underline{k}}$ with positive probabilities $g_{\underline{k}}(p)$ and p being the per gate error rate. For instance, when boosting the error rate of a single-qubit depolarising channel from p_1 to p, the circuit variants $\mathcal{G}_{\underline{k}}$ correspond to randomly inserting Pauli X, Y or Z with probability $(p - p_1)/(1 - \frac{4}{3}p_1)$ after the gate (see Eq. (6.15) for more

details). This generalises to more complex circuits by repeating the process for each gate. The $g_{\underline{k}}(p)$ can be learned in the same way we learned the quasiprobabilities $\gamma_{\underline{k}}$ using Pauli-Lindblad learning [211]. We can now formalise the error-boosted shadows protocol as a corollary to Theorem 2.

Corollary 1 (Error-boosted shadows)**.** *Given a quasiprobability decomposition* Γ *of the noise-boosted circuit* $\mathcal{G}_{circ}(p)$ *with* $p > p_1$ *and a classical shadow protocol with the POVM measurement* E *as in Eq.* (5.18)*, the PEC shadows* $H \coloneqq \{(g_{\underline{k}}(p), \mathcal{G}_{\underline{k}}, E_l)\}_{\underline{k},l}$ *yields an unbiased estimator of* ρ_p*,*

$$\hat{\rho}_p \coloneqq \hat{\rho}_{\underline{k},l} = C_E^{-1}(E_l), \tag{6.8}$$

in that $\mathbb{E}[\hat{\rho}_p] = \mathbb{E}_{\underline{k},l}[\hat{\rho}_{\underline{k},l}] = \rho_{id}$*. Note that we obtain a simpler expression here, as* $\|g\|_1 = 1$ *and* $\text{sign}(g_{\underline{k}}) = +1$*.*

A big advantage of this error-boosting protocol is that, as the $g_{\underline{k}}(p)$ are positive, there is no sampling cost associated with it as $\|g\|_1 = 1$. However, there is a measurement cost associated with the use of zero noise extrapolation as explained in Section 5.1.2.

More generally, error-extrapolated shadows can be used beyond the estimation of linear properties of a quantum state. We could apply this technique directly to the states by reconstructing the matrices $\hat{\rho}_p$ for different error rates and extrapolating to $p = 0$ for each matrix entries. We could also apply it to estimate non-linear properties but we might need to adapt the model functions f for each situation.

Finally , it is important to remark that this technique yields a biased estimator contrary to PEC shadows.

6.3.2 Symmetry verified shadows

With symmetry verification, we mitigate the errors in the state preparation process by noting that prepared state should respect the symmetries of the ideal state. We can enforce these symmetry requirements by either post-selecting on states respecting the symmetries or by using post-processing as explained in Section 5.1.3.

While we can apply classical shadows on the post-selected state ρ_{symm} (at the cost of a modification of the circuit) as defined in Eq. (5.14) and obtain an error-mitigated estimate of $\mathrm{Tr}[O\rho_{symm}]$, a simpler and more elegant way to leverage shadows is to combine them with the post-processing way of enforcing symmetries. In Eq. (5.15), I showed that,

$$\mathrm{Tr}[O\rho_{symm}] = \frac{\mathrm{Tr}[O\Pi_{\mathbb{S}}\rho\Pi_{\mathbb{S}}]}{\mathrm{Tr}[\Pi_{\mathbb{S}}\rho]} = \frac{1}{|\mathbb{S}|}\frac{\sum_{S,S'\in\mathbb{S}}\mathrm{Tr}[SOS'\rho]}{\sum_{S\in\mathbb{S}}\mathrm{Tr}[S\rho]}.$$

In this form, the error mitigated expectation value $\mathrm{Tr}[O\rho_{symm}]$ can be computed from the estimate of multiple expectation values, which is exactly what shadows are good at. Hence, we can use shadows to compute $\mathrm{Tr}[SOS'\rho]$ and $\mathrm{Tr}[S\rho]$ for all $S, S' \in \mathbb{S}$ and combine them to estimate $\mathrm{Tr}[O\rho_{symm}]$ provided that the symmetries' localities are small enough. Suppose that they are supported by Pauli-strings of weight at most s and that O is a q-local Pauli operator, then SOS' is at most a $(2s+q)$-local observable. Recall that for the practical case of single-qubit Pauli basis transformation the shadow norm of an observable grows exponentially with its locality. Hence, symmetry verified shadows are only viable when $(2s+q)$ is small.

The sampling cost can be further reduced when O commutes with the operator $\Pi_{\mathcal{S}}$, which is usually the case in practice. In that case, $\Pi_{\mathcal{S}}O\Pi_{\mathcal{S}} = O\Pi_{\mathcal{S}}$, which leads to,

$$\mathrm{Tr}[O\rho_{symm}] = \frac{\mathrm{Tr}[O\Pi_{\mathbb{S}}\rho\Pi_{\mathbb{S}}]}{\mathrm{Tr}[\Pi_{\mathbb{S}}\rho]} = \frac{\mathrm{Tr}[O\Pi_{\mathbb{S}}\rho]}{\mathrm{Tr}[\Pi_{\mathbb{S}}\rho]} = \frac{\sum_{S\in\mathbb{S}}\mathrm{Tr}[OS\rho]}{\sum_{S\in\mathbb{S}}\mathrm{Tr}[S\rho]}. \tag{6.9}$$

Notice that now the observables with the largest weight are the SO for $S \in \mathbb{S}$ with a weight at most $(s+q)$. By reducing the largest weight, we reduce the shadow norm and thus the sampling complexity.

6.4 Applications

In this section, I present multiple applications of QEM shadows that could extend the reach of current noisy quantum computers. While in the fault-tolerant era we will be able to extract properties of states using coherent techniques such as quantum amplitude estimation (whose complexity is $\mathcal{O}(M/\varepsilon)$ and is proportional

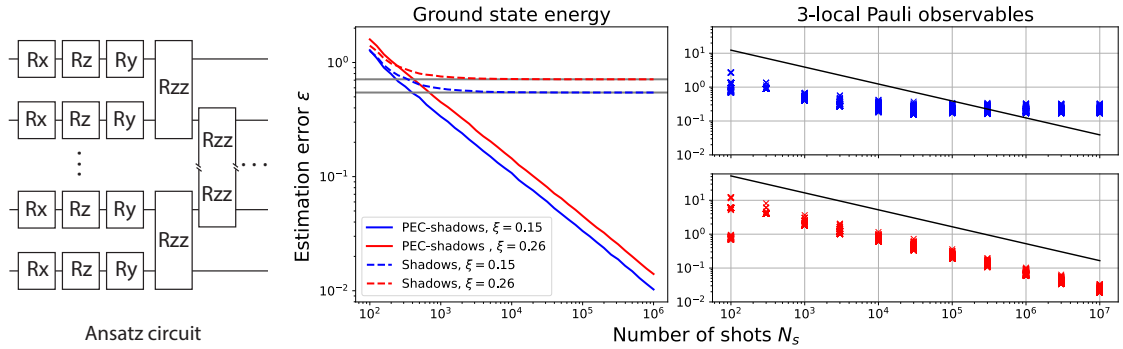

Figure 6.3: (left) A noisy variational Hamiltonian ansatz is used to prepare the ground state of Eq. (6.10) but our aim is to learn properties of the noise-free state. (middle) Energy estimation errors for different noise strengths with conventional shadows (dashed blue, dashed red) and with PEC shadows including readout error mitigation (solid blue, solid red). Bias (grey solid lines) is introduced when the ground state energy $\mathrm{Tr}[\rho\mathcal{H}]$ is directly estimated from the noisy quantum state ρ. Error mitigated shadows are unbiased as they estimate $\mathrm{Tr}[\rho_{id}\mathcal{H}]$. Increasing the circuit error rate ξ (blue vs. red) increases the bias in standard shadows (dashed blue vs. dashed red) and increases the variance of the error mitigated shadows (solid blue vs. solid red). Each data point is an average over 10^4 experiments of a fixed shot budget N_s. (right) Error in simultaneously estimating all 3-local Pauli strings without (blue) and with (red) PEC and readout error mitigation – only the 200 observables of the highest estimation error are shown and a circuit error rate $\xi \approx 0.72$ is assumed. Errors are significantly below our rigorous bounds from Theorem 3 (which also take into account the overhead due to readout errors) for PEC shadows but the errors for conventional shadows can be above their respective bounds from [219] due to bias and readout errors (right end of blue).

to the circuit's depth) [238], they are not implementable on current hardware due to their limited circuit depth. Conversely, to implement classical shadows one only needs to increase the depth by an amount that is independent of the state preparation circuit (e.g. a single layer of single-qubit gate in the case of random single-Pauli basis transformations). Combining this fact with their moderate sample complexity $\mathcal{O}(\log(M)/\varepsilon^2)$, classical shadows are notably relevant when the goal is to estimate a large number of properties M at once.

For instance, noisy quantum computers in either the late NISQ era or in the early fault-tolerance era may be used to propagate quantum states in time or to compute eigenstates [186, 226–228, 239–241]. Our protocol can then be used to accurately and efficiently extract a large number of properties of these states provided that the measurement overhead $\|\gamma\|_1$ remains reasonable, i.e., the circuit error rate is small. In these application areas a fixed precision, such as chemical

accuracy $\varepsilon \approx 10^{-3}$, is often sought [216, 226–228] which will set a lower bound on the number of measurements needed.

6.4.1 Ground-state preparation

We first consider a spin-ring Hamiltonian as

$$\mathcal{H} = \sum_{k \in \text{ring}(N)} \omega_k Z_k + J \vec{\sigma}_k \cdot \vec{\sigma}_{k+1}, \tag{6.10}$$

with coupling $J = 0.3$, on-site interaction strengths uniformly randomly generated in the range $-1 \leq \omega_k \leq 1$ and $\vec{\sigma}_k = (\sigma_k^x, \sigma_k^y, \sigma_k^z)^T$ is a vector of single-qubit Pauli matrices. This spin problem is relevant in condensed-matter physics in understanding many-body localisation [242] but is challenging to simulate classically for large N [243, 244]. A broad range of techniques are available in the literature for finding eigenstates of such quantum Hamiltonians using near-term or early fault-tolerant quantum computers [226–228, 239]. Here, I prepare the ground state of this model using a variational Hamiltonian ansatz in Fig. 6.3 (left) of $l = 5$ layers on 12 qubits using an emulated noisy quantum computer. I use an exact quantum state emulator [245] to simulate the noisy quantum circuits. Given that our protocol is effectively a Monte-Carlo sampling scheme, and to avoid the simulation of the full density matrices, I also simulate the effect of noise using a Monte-Carlo approach. For instance, a Pauli noise channel is simulated by randomly applying a Pauli operator according to its corresponding probability.

As recent experiments have shown that using Pauli twirling can guarantee that the noise model is well approximated by a Pauli noise model [131], I suppose that gate noise can be represented as a Pauli channel. Moreover, I assume that two-qubit gates are the dominant source of noise and that the readout error-model is the one given in Section 5.2.2 whereby each qubit has the same bit flip probability $\alpha = 0.01$ (which is consistent with current readout errors).

We thus have that each noisy quantum gate is affected by the following, biassed, local Pauli channel

$$\Phi_k(\rho) \coloneqq (1 - p_k)\rho + p_k \left(\frac{1 - \eta_k}{2} X\rho X + \frac{1 - \eta_k}{2} Y\rho Y + \eta_k Z\rho Z \right), \tag{6.11}$$

where η_k is a bias parameter. Furthermore, we consider that the error probability p_k is specific to each gate and we randomly draw their values from the distribution $\mathcal{N}(\mu = p, \sigma = p)$ and as we explain below we explore two different regimes $p = 10^{-3}$ and $p = 2 \times 10^{-3}$. We randomly choose probabilities to reflect the high variability of two-qubit error rates in typical superconducting devices [131]. Of course, this variability may be lower in ion-trap devices [246, 247]. Given in most platforms T_2^* relaxation timescales are significantly faster than T_1 relaxation timescales, we consider a biased noise channel, that is specific to each gate, by choosing the bias parameters η_k from the distribution $\mathcal{N}(\mu = 0.9, \sigma = 0.015)$. This corresponds to a noise biased towards dephasing.

We can straightforwardly find the inverse noise channel analytically as

$$\Phi_k(\rho)^{-1} \coloneqq \gamma_0 \rho + \gamma_1 X\rho X + \gamma_2 Y\rho Y + \gamma_3 Z\rho Z,$$

via the coefficients

$$\begin{aligned}
\gamma_0 &= \|\gamma\|_1 \frac{2 - p_k(p_k\eta_k^2 - p_k - 2\eta_k + 4)}{2 + 2p_k(\eta_k - 1)(1 + p_k + p_k\eta_k)}, \\
\gamma_1 = \gamma_2 &= \|\gamma\|_1 \frac{p_k(\eta_k - 1)(p_k\eta_k + p_k - 1)}{2 + 2p_k(\eta_k - 1)(1 + p_k + p_k\eta_k)}, \\
\gamma_3 &= \|\gamma\|_1 \frac{p_k(2\eta_k + p_k(\eta_k^2 - 1))}{2 + 2p_k(\eta_k - 1)(1 + p_k + p_k\eta_k)},
\end{aligned} \tag{6.12}$$

with $\|\gamma\|_1 = 1/2\,(1/(1 + 2p_k(\eta_k - 1)) - 2/(p_k + p_k\eta_k - 1))$. In the simulations, I thus randomly apply Pauli X, Y, Z or $\mathbb{1}$ gates to the qubits in the support of the relevant gate via probabilities defined by $\underline{\gamma}$. Note that, while I found the coefficients $\underline{\gamma}$ analytically, they can also be computed numerically at the preprocessing stage [211].

Ground state energy with PEC

I use classical shadows to estimate the individual energies for each term indexed by $k \in \text{ring}(N)$ in the Hamiltonian and sum them to obtain the ground state energy. Fig. 6.3 (middle) shows the error in the ground state energy estimated using conventional shadows (dashed blue and dashed red) and PEC shadows (solid blue and solid red) for an increasing number of shots N_s when averaging over 10^4

different experiments. I efficiently simulate this sampling task via the following resampling scheme: I generate a very large pool of 10^7 snapshots (an order of magnitude more than the largest simulated shot budget $N_s \leq 10^6$) using the above described quantum-mechanical simulations. I then estimate the ground-state energy by randomly choosing N_s snapshots each time from this large pool.

The ansatz circuit would ideally prepare the ground state ρ_{id} but due to gate noise and readout errors it actually prepares the noisy state ρ. Thus, conventional shadows (dashed blue, dashed red) converge to a plateau corresponding to the biased energy $\mathrm{Tr}[\mathcal{H}\rho]$ (solid grey). This bias is significantly increased as I increase the circuit error rate from $\xi \approx 0.15$ to $\xi \approx 0.26$ (dashed blue vs. dashed red), with $\xi = \sum_i p_i$ being the expected number of errors in the full circuit as explained in Section 5.1.1 (with $\nu = 60$ noisy two-qubit gates). In contrast, PEC shadows that include measurement-error mitigation (solid blue and solid red) converge to the true energy $\mathrm{Tr}[\mathcal{H}\rho_{id}]$ in standard shot-noise scaling $\mathcal{O}(1/\sqrt{N_s})$.

Local properties with PEC

Besides Hamiltonian energy estimation, which is one of the typical subroutines in quantum computing, there is also significant value in simultaneously determining many local observables' expectation values. For example, the rich information from classical shadows can be used to significantly improve parameter training or to directly estimate Hamiltonian energy gaps through the use of efficient classical post-processing [239, 240]. In Fig. 6.3(right), I plot errors when simultaneously estimating all 3-local Pauli operators for an increasing number of shots N_s. Fig. 6.3(red) shows that the errors in PEC shadows are always significantly below the theoretical bounds (black line) from Theorem 3 confirming looseness of the bounds (assuming success probability $\delta = 10^{-3}$, and $M = 3^3\binom{12}{3} = 5940$). Fig. 6.3(blue) shows the errors in conventional shadows are below their bounds (with $\|\gamma\|_1 = 1$) only for a small number of shots but then asymptotically reach a plateau due to circuit noise.

These bounds depend on the largest shadow norm and as my co-author showed in Lemma 5, the shadow norm of a q-local observable is equal to 3^q when using

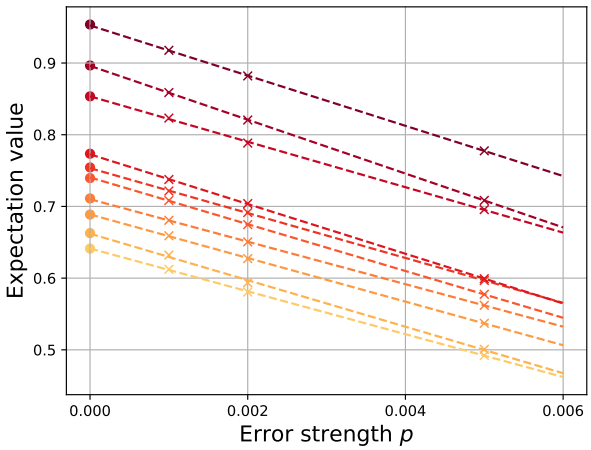

Figure 6.4: Simultaneously estimating all 3-local Pauli operators using error extrapolated shadows. I estimated noisy expected values (crosses) from shadows of size $N_s = 10^7$ by increasing the native depolarising error rate $p = 10^{-3}$ to higher levels $\{2 \times 10^{-3}, 5 \times 10^{-3}\}$ by randomly sampling noisy circuit variants. Using a linear model function, I then extrapolate to zero noise to obtain an error mitigated expectation value close to the ideal ones (disks).

ideal measurements. As I am taking readout error into account, the PEC snapshots are modified according to Eq. (5.22) and it is straightforward to show that the new shadow norm for a q-local observable is given by,

$$\|O\|_E^2 = \frac{3^q}{(1-2\alpha)^{2q}}. \tag{6.13}$$

By modifying the bounds from Theorem 3 accordingly, I obtain the black lines in Fig. 6.3 (right). Note that readout-error mitigation has an associated measurement overhead of $(1-2\alpha)^{-2q}$.

Local properties with extrapolation

I now consider the same task of simultaneously estimating expectation values of Pauli operators but using error extrapolation. Here, I start by generating shadows at different noise strengths (which leads to different circuit error rate) that I use to compute the noisy Pauli expectation values.

As the main objective is to demonstrate that extrapolation-based error mitigation is indeed compatible with classical shadows, I assume a simple error model whereby every two-qubit gate has the same error probability p that I can perfectly magnify

to λp. For instance, consider a single-qubit depolarising channel with probability p given by taking the asymmetry parameter $\eta_k = 1/3$ in Eq. (6.11),

$$\Phi_p(\rho) \coloneqq (1-p)\rho + p/3[X\rho X + Y\rho Y + Z\rho Z]. \tag{6.14}$$

where X, Y and Z are Pauli matrices. Note that $p \leq 3/4$ in this model and for $p = 3/4$ we obtain the completely mixed state (in practice we have $p < 3/4$).

As explained in Corollary 1, we can magnify the noise by randomly inserting Pauli gates after this channel, effectively adding another depolarising channel with probability $p^{'}$. One can show that combining these two channels yields a single depolarising channel $\Phi_q(\rho)$,

$$\Phi_q(\rho) = \Phi_{p'}(\rho)\Phi_p(\rho), \tag{6.15}$$

with $q = p + p^{'} - \frac{4}{3}pp^{'}$. Hence, by setting $p^{'} = (\lambda p - p)/(1 - \frac{4}{3}p)$ we boost the error rate from p to λp. In practice, boosting the error rates accurately is quite challenging but has been successfully performed in large-scale experiments [131]. The channel in Eq. (6.14) can be analytically inverted to obtain the inverse channel as

$$\Phi_p(\rho)^{-1} \coloneqq \gamma_0\rho + \gamma_1 X\rho X + \gamma_2 Y\rho Y + \gamma_3 Z\rho Z.$$

The explicit form of the coefficients follows as

$$\underline{\gamma} = \|\gamma\|_1 (\frac{3-p}{2p+3}, \frac{-p}{2p+3}, \frac{-p}{2p+3}, \frac{-p}{2p+3}),$$

and the norm is $\|\gamma\|_1 = (3+2p)/(3-4p)$. The same noise model will be used for Fig. 6.5.

Fig. 6.4 shows 10 examples of expected values (crosses) as a function of noise strength and the respective linear models I fit (dashed lines). The intercept of the fitted model (dashed lines) is the estimate of the ideal expected value (disks) and is indeed reasonably close in the example. While ZNE has been very effective and typically has a lower measurement overhead then PEC, it is generally biased.

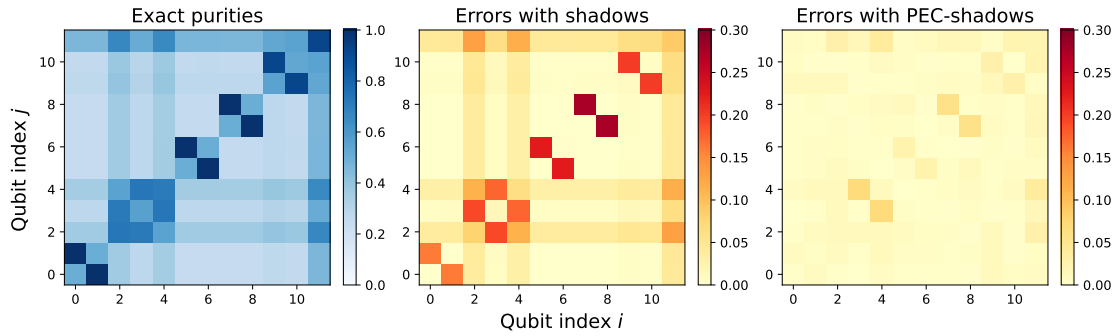

Figure 6.5: A noisy variational Hamiltonian ansatz is used to prepare the ground state of Eq. (6.16) whose ideal, noise-free Rényi entropies R_Q I can learn with PEC shadows. I plot purities $\mathrm{Tr}[\rho_Q^2]$ as a proxy for $R_Q := -\log \mathrm{Tr}\left[\rho_Q^2\right]$. (left) Purity heat map in the noiseless case and infinite shot limit. An increasing value indicates that the subsystem Q is less entangled with the remaining qubits. (middle-right) Absolute error in the purities due to gate noise for a circuit error rate $\xi = 0.6$ and due to finite repetition using $N_s = 10^5$. (middle) Although the entanglement pattern is approximately recovered with conventional shadows, in some instances I observe substantial errors, i.e., the largest error is 0.27. (right) Absolute errors with PEC shadows are significantly smaller, i.e., the largest error is 7×10^{-2} but this figure could be further reduced by increasing N_s.

6.4.2 Error mitigated estimation of entanglement entropies

Finally, I consider an application for which classical shadows are a primary enabler but for which error mitigation techniques have been less explored [133]. As opposed to studying entanglement properties or verifying the presence thereof in mixed quantum states [248–251], here my primary goal is to extend the reach of noisy quantum computers: I explore entanglement properties of ideally pure states which are prepared by quantum algorithms, such as phase estimation or VQE. For example, near-term quantum computers will enable us to prepare eigenstates [226–228, 239] of quantum Hamiltonians and error-mitigated entanglement measures can be used for, e.g., characterising phase transitions. Similarly, one could simulate the time evolution of a collision of two molecules with an early fault-tolerant quantum computer and investigate how entanglement builds up across the individual subsystems. Furthermore, efficiently characterising many local correlations in a state can be used to train DFT models for accurate classical simulations [252].

I consider the Heisenberg chain

$$\mathcal{H} = \sum_k J_k \vec{\sigma}_k \cdot \vec{\sigma}_{k+1}, \tag{6.16}$$

with uniformly random $-1 \leq J_k \leq 1$ and prepare its ground state with a variational Hamiltonian ansatz of $l = 8$ layers on 12 qubits. This system was used in ref. [219] to illustrate the power of classical shadows in predicting entanglement entropies. However, the ground state was approximated by a set of noise-free singlet states [253, 254] whereas here I assume a noisy quantum computer is used for state preparation.

I use PEC shadows to extract purities $\mathrm{Tr}[\rho_Q^2]$ for all single and two-qubit subsystems Q. These purities then define Rényi entropies as $R_Q \coloneqq -\log \mathrm{Tr}[\rho_Q^2]$. In Fig. 6.5, I plot the exact purities in the noiseless case – disjoint blocks involving two qubits confirm that the ground state could be approximated by a tensor product of noise-free singlet states.

Fig. 6.5 (middle) shows the errors in estimating local purities using shadows of size $N_s = 10^5$ for a circuit error rate $\xi = 0.6$. Even for this moderate error rate conventional shadows are significantly impacted by imperfections and result in errors as large as 0.27 – whereas for an increasing noise rate all purities converge to a constant value of $1/d$ where d is the subsystem dimension. In contrast, PEC shadows drastically improve the accuracy in Fig. 6.5 (right) and the largest error is approximately 7×10^{-2}.

6.4.3 Further applications

In this section, I only presented a few applications of QEM shadows. As PEC shadows yield an unbiased estimator of the ideal density matrix ρ_{id}, it could be applied in various settings. One such application could be the estimation of the fidelity of a state as explored in [219]. Using classical shadows generated from random Clifford measurement, Huang et al. showed that we can estimate fidelities of a state ρ_{id} with respect to a given state σ. However, suppose now that the state was generated by a noisy quantum device, then the whole protocol breaks down. The extracted fidelity corresponds to the fidelity between a noisy version of ρ_{id} and thus does not make sense. With PEC shadows this problem is now solved. This could allow one to check the structure of a variational quantum circuit by verifying that it prepares the state that it should, even in the presence of gate noise.

Similarly, in quantum metrology, the quantum Fisher information (QFI) can be estimated using classical shadows [255] and in certain situations, obtaining the noisy version of the QFI is not relevant [256].

Finally, while quantum error mitigation could help broaden the applicability of classical shadows, we note the converse can also be true with classical shadows used to improve error mitigation techniques. For instance, one could perform shadow tomography to reconstruct the noisy channels in Section 5.1.1 (assuming that the channels are local) and invert them to implement PEC. As it might be expensive to estimate the complete channel, one can decide to use learning-based error mitigation in which one wants to learn the quasiprobabilities $\gamma_{\underline{k}}$ rather than the channel itself. Learning-based error mitigation consists in running quantum circuits that can be simulated classically in order to compare the noisy and ideal expectation values (obtained classically) of observables [133, 209, 210, 212]. With conventional measurement protocols, a model is trained for a particular observable. With classical shadows, we can estimate a large number of expectation values and mitigate the errors in every local observables. Classical shadows could also be useful in learning sparse Pauli models [211] as it requires the measurement of multiple Pauli operators' expectation value.

6.5 Discussion

Classical shadows allows for multiple properties of a quantum state to be estimated at a low cost. However, the noise present in current quantum computers introduces errors during state preparation, resulting in corrupted properties when using classical shadows. Standard error mitigation techniques can offer a solution, allowing for the extraction of noiseless properties at the cost of an increased number of measurements. In the research describe here, my co-authors and I combine both techniques and show that quantum error mitigated classical shadows not only enable the recovery of noiseless properties, but more interestingly allows us to obtain an efficient representation of the ideal state itself.

While we explore various error mitigation techniques, we identify probabilistic error cancellation as the most natural fit with classical shadows. Consequently, we present a comprehensive theoretical study of PEC shadows showing that they yield an unbiased estimator of the ideal quantum state ρ_{id}. It is important to note that our scheme relies on a detailed knowledge of the noise model, which can be efficiently learned as discussed in the main text. This compact classical representation then allows us to extract multiple properties of the ideal state. We establish bounds on the number of samples required to achieve a given accuracy in predicting multiple linear and non-linear properties, demonstrating that the measurement overhead compared to unmitigated classical shadows is comparable to that of standard PEC. The cost introduced by error mitigation grows exponentially with the circuit error rate, albeit with the huge advantage of not depending on the number of qubits. Consequently, our approach is limited to scenarios where the circuit error rate is $\xi = \mathcal{O}(1)$, similar to other error mitigation techniques.

Finally, I showcased the usefulness of QEM shadows focusing on the practical case of random single-qubit Pauli basis transformations through numerical simulations. Note however that our scheme is completely general and can be applied to other basis transformations ensembles. First, I examined how PEC shadows allow for the recovery of noiseless ground state energy estimates even when using a noisy device. Then, I showed that more generally it allow us to extract numerous local properties of the ideal state using PEC shadows or the less costly error extrapolation. Obtaining accurate expectation values for many local observables can then be used to compute Hamiltonian energy gaps [240] or improve parameter training of variational circuits [239]. Finally, I studied the estimation of entanglement entropies for which the use of error mitigation has been less explored. QEM shadows allow for the computation of entanglement properties of the pure state an ideal device would produce in the absence of noise, which could be used for characterising phase transitions. Our results indicate that QEM shadows may be indispensable in practical applications of classical shadows and will extend the reach of NISQ-era and early fault-tolerant quantum computers.

7

Biased estimator channels for classical shadows

This chapter presents the research contained in the preprint [257]. The idea of biasing the estimator channel in the shadow protocol originated from Adrian Chapman during one of our discussions. Section 7.2 elaborates on the concept of biased shadows for which I participated in the development. My specific role in this project involved analytically deriving the worst- and best-case scenarios and study them numerically while implementing the application proposed by Zhenyu Cai, with feedback from Bálint Koczor. At times, I use text verbatim from [257] where I am an original author.

Contents

7.1 Introduction 127
7.2 Biased estimator channel 128
7.2.1 Classical shadows review 128
7.2.2 Biased local estimator channel 129
7.2.3 Optimal biasing strategy 129
7.3 Analysis of the worst- and best-case scenarios 131
7.3.1 Best-case scenario 131
7.3.2 Worst-case scenario 132
7.4 Practical application 135
7.5 Discussion 138

7.1 Introduction

In the previous chapter, we showed that through the combination of classical shadows and error mitigation, one could recover a noiseless representation of a quantum state even in the presence of noise at the cost of an increased number of measurements. This means that the remaining error in estimating properties of a quantum state is due to shot noise, as with conventional shadows. As such, we will suppose that no gate nor readout noise is present in this chapter and that the goal is to reduce the shot noise error.

In order to understand how to reduce this error, it is worth taking a step back and rethink the shadow protocol. If we disregard the intricacies of the classical shadow protocol, it essentially provides an unbiased estimator of a certain quantity (e.g. density matrix, expectation value). Leveraging standard statistics techniques, we can find ways to reduce shot noise. In particular, in the spirit of the well-known variance-bias tradeoff, one can introduce a bias to reduce the variance of the estimator and thus reduce shot noise. Here, we show that surprisingly it is always worth biasing the shadow estimator. Through a comprehensive analysis of our protocol, we find that there is always an optimal biasing point ε^* (that depends on the property we want to extract) which will yield a smaller error than conventional shadows.

Furthermore, another big advantage of our scheme is that it is simple to implement as it only requires a modification of the postprocessing step, where we scale down the estimator channel by a factor ε. We choose this factor to be dependent on the number of shots N_s, such that $\varepsilon = 0$ when N_s goes to infinity. This allows us to recover the unbiased estimator in the infinite sample limit. Our scheme is thus more general that conventional classical shadows [219–221, 232, 258–260] as they correspond to the special case $\varepsilon = 0$.

This chapter is divided as follows. First, I give a brief summary of classical shadows and introduce the biased estimator channel. Then, I analyse the best- and worst-case scenarios for our protocol applied to the estimation of expectation values. Finally, I explore a practical scenario in which biased shadows outperform unbiased shadows.

7.2 Biased estimator channel

7.2.1 Classical shadows review

In Section 5.2, I presented the conventional classical shadows protocol and showed that it is equivalent to applying the channel C_E to the state of interest ρ such that,

$$C_E(\rho) = \sum_{l=1}^{N_E} \mathrm{Tr}[E_l \rho] E_l, \tag{7.1}$$

with E_l the effects of the POVM E defined in Eq. (5.18). We wish to construct a map $\mathcal{M}$ – called the estimator channel – that returns an unbiased estimator for every input density matrix ρ. More precisely, $\mathcal{M}$ must verify the following conditions,

$$\forall \rho,\ \mathcal{M} \circ C_E(\rho) = \rho. \tag{7.2}$$

Eq. (7.2) ensures that $\mathcal{M} = C_E^{-1}$ and we retrieve the construction of the snapshot of Theorem 1, which for the outcome $l = (j, \underline{b})$ yields the unbiased estimator $\hat{\rho}_l$,

$$\hat{\rho}_l = \mathcal{M}(E_l) = C_E^{-1}(E_l). \tag{7.3}$$

To make things more concrete let us explore the practical case of random single-qubit Pauli basis transformations, while noting that our scheme can be applied to different unitary ensembles $\mathcal{Q}$ as defined in Section 5.2. In this case we can rewrite Eq. (7.1) as,

$$C_E(\rho) = \sum_{l=1}^{N_E} q_l \bigotimes_{l=1}^{N} \frac{1}{3} (Q_j^{(i)})^\dagger \left|b^{(i)}\right\rangle\!\left\langle b^{(i)}\right| Q_j^{(i)} = \sum_{l=1}^{N_E} q_l \bigotimes_{l=1}^{N} \frac{1}{3} E_l^{(i)}, \tag{7.4}$$

with $E_l^{(i)} = (Q_j^{(i)})^\dagger \left|b^{(i)}\right\rangle\!\left\langle b^{(i)}\right| Q_j^{(i)}$ and $q_l = \mathrm{Tr}[E_l \rho]$. Thus, the effect of C_E is to contract the single-qubit states $E_l^{(i)}$ by a factor 3 with probability q_l. Given the outcome E_l, the estimator channel yields the unbiased estimator Eq. (7.3) which can also be split in a tensor product as,

$$\begin{aligned} \hat{\rho}_l &= \bigotimes_{i=1}^{N} \left[3 (Q_j^{(i)})^\dagger \left|b^{(i)}\right\rangle\!\left\langle b^{(i)}\right| Q_j^{(i)} - \mathbb{1} \right] \\ &= \bigotimes_{i=1}^{N} \mathcal{M}_{local}(E_l^{(i)}). \end{aligned} \tag{7.5}$$

We notice that the effect of the estimator channel is to uniformly dilate each single-qubit states by a factor 3 to compensate for the contraction introduced by C_E. In the next subsection, we explore how adding bias to this channel helps reducing the variance.

7.2.2 Biased local estimator channel

Here, we relax the condition set in Eq. (7.2) by deliberately introducing a bias in the estimator channel. Instead of uniformly dilating the Bloch sphere by a factor 3, we dilate it by a factor $3\sqrt{1-\varepsilon}$, with ε being a tunable bias parameter.

Statement 1 (Biased shadow estimator)**.** *Given a bias parameter ε, the local biased shadow estimator defined in Eq.* (7.5) *is modified as follows,*

$$\mathcal{M}_{local,\varepsilon}(E_l^{(i)}) = 3\sqrt{1-\varepsilon}E_l^{(i)} + \frac{1}{2}\left(1 - 3\sqrt{1-\varepsilon}\right)\mathbb{1}. \tag{7.6}$$

Note that we retrieve the standard local channel for $\varepsilon = 0$. While we do not converge to the true state in the infinite-sample limit for $\varepsilon \neq 0$, tuning ε allows us to control shot noise as the variance is reduced by $(1-\varepsilon)$.

In practice, we will chose an ε that depends on the number of shots, such that we recover an unbiased estimator in the infinite-sample regime. Note that this implies that our estimator is consistent.

7.2.3 Optimal biasing strategy

In order to characterise the performance of our scheme for estimating properties of a quantum state, let us take a step back from the classical shadows formalism and remind ourselves that in the unbiased case classical shadows can be seen as a way to obtain the single-shot estimator $\hat{o} = \mathrm{Tr}[O\hat{\rho}_l]$ for a given observable O. In the biased case, the single-shot estimator is given by $\hat{o}_\varepsilon = \sqrt{1-\varepsilon}\hat{o}$. In practice, we will repeat the experiment N_s times and use the sample mean estimator $\bar{o}$ defined as,

$$\bar{o} = \frac{1}{N_s}\sum_{i=1}^{N_s}\hat{o}_i \tag{7.7}$$

with $\hat{o}_i$ being the outcome of the ith run. Similarly we define $\bar{o}_\varepsilon = \sqrt{1-\varepsilon}\bar{o}$. To compare the performance of both schemes we will use their respective Mean Squared Error (MSE) as defined in Eq. (D.1). We have,

$$\mathrm{MSE}[\bar{o}] = \frac{\mathrm{Var}[\hat{o}]}{N_s} \tag{7.8}$$

$$\mathrm{MSE}[\bar{o}_\varepsilon] = (1-\sqrt{1-\varepsilon})^2\mathbb{E}[\hat{o}] + (1-\varepsilon)\frac{\mathrm{Var}[\hat{o}]}{N_s}. \tag{7.9}$$

The first term in Eq. (7.9) represents the introduced bias (which is equal to zero when $\varepsilon = 0$) while the second represents the variance. In Appendix F.1 we minimise Eq. (7.9) and find that the optimal biasing point ε^* is not equal to zero for a fixed shot budget N_s– which means that the unbiased scheme is not optimal – and is given by,

$$\varepsilon^* = \frac{1+2\beta}{(1+\beta)^2}, \tag{7.10}$$

with β being the signal-to-noise ratio of the unbiased shadow estimator,

$$\beta \coloneqq \frac{\mathbb{E}[\bar{o}]}{\mathrm{MSE}[\bar{o}]} = \frac{\mathbb{E}[\hat{o}]}{\mathrm{Var}[\hat{o}]/N_s}. \tag{7.11}$$

The MSE of the optimised biased estimator now reads,

$$\mathrm{MSE}[\hat{o}_{\varepsilon^*}] = (1-\sqrt{1-\varepsilon^*})\mathbb{E}[\hat{o}]^2 = \frac{1}{1+\beta}\mathbb{E}[\hat{o}]^2, \tag{7.12}$$

and its SNR is given by,

$$\beta_{biased} = 1+\beta. \tag{7.13}$$

The optimised SNR β_{biased} is always larger than 1 for *all estimators*, which is not the case of β. While the gain obtained through biasing is marginal when the unbiased estimator has a large SNR i.e. $\beta \gg 1$, we can now always extract useful signal even when $\beta \ll 1$.

Let us now come back to the shadow protocol where $\mathbb{E}[\hat{o}] = \mathrm{Tr}[O\rho]$ and $\mathrm{Var}[\hat{o}] = 3^q - \mathrm{Tr}[O\rho]^2$ for a q-local observable O (see Appendix F.2 for the derivation). Note that to find ε^* one needs a perfect knowledge of $\mathbb{E}[\hat{o}]$, which is exactly the quantity we are trying to estimate here. While this seems problematic, I show in Section 7.4 that

an approximate estimate of $\mathbb{E}[\hat{o}]$ can be sufficient to obtain a significant reduction in the MSE. The factor of SNR gain β_{biased}/β is given by,

$$\frac{\beta_{biased}}{\beta} = 1 + \frac{3^q/\mathrm{Tr}[O\rho]^2 - 1}{N_s}. \tag{7.14}$$

Notice that there is no gain in the infinite-sample limit as $\lim_{N_s\to\infty} \beta_{biased}/\beta = 1$. This is consistent with the fact that we recover the conventional unbiased shadow estimator in this regime. This ratio suggests that our scheme works best for high-weight observable and we will see that it is indeed in the case in Section 7.4. From Eq. (7.14) we can also extract the worst- and best-case scenarios obtained when $\mathrm{Tr}[O\rho] = 1$ and $\mathrm{Tr}[O\rho] = 0$ (we go from an SNR $\beta = 0$ to $\beta_{biased} = 1$) respectively.

7.3 Analysis of the worst- and best-case scenarios

7.3.1 Best-case scenario

Using the factor of SNR gain in Eq. (7.14), we find that the best-case scenario occurs when $\mathrm{Tr}[O\rho] = 0$. However, there is a more intuitive reason why this is the most favorable situation for our protocol. As explained in Section 7.2.2, increasing the bias corresponds to decreasing the dilation factor of the Bloch sphere. As a consequence, increasing the bias will reduce the value of the estimated expectation value up to to the point where we always return 0 (the ideal value here) when $\varepsilon = 1$. Biasing is thus always advantageous. While we used the SNR to derive the worst- and best-case scenarios, I focus now on the MSE which is a more practical measure of the error.

For ease of presentation, I focus first on the single-qubit case before generalising to multiple qubits. Without loss of generality, let $\rho = |+\rangle\langle+|$ and $O = Z$ which thus satisfies $\mathrm{Tr}[O\rho] = 0$. For a given collection of shadows $S(\rho, N_s) = \{\hat{\rho}_1, \ldots, \hat{\rho}_{N_s}\}$, our estimate for the expectation value is given by the mean of each snapshot's expectation value,

$$\frac{1}{N_s}\sum_{i=1}^{N_s} \mathrm{Tr}[Z\hat{\rho}_i]. \tag{7.15}$$

Each $\hat{\rho}_i$ has been obtained by measuring in a randomly chosen single-qubit Pauli basis, and as a result $\mathrm{Tr}[Z\hat{\rho}_i]$ yields either $+1$ or -1 with equal probability (compatible measurement in the Pauli Z basis) or yields 0 (incompatible measurement via Pauli X and Y bases). Here, I aim to calculate the mean squared error made in computing the expectation value $\langle E_1^{\mathrm{b}}(\varepsilon)\rangle$ when using a bias parameter ε. Suppose that $\sum_{i=1}^{N_s}\mathrm{Tr}[Z\hat{\rho}_i] = k$, I define the error in the mean value as,

$$E_1^{\mathrm{b}}(\varepsilon) \coloneqq \left|\frac{1}{N_s}\sum_{i=1}^{N_s}\mathrm{Tr}[Z\hat{\rho}_i] - 0\right|^2 = \left|\frac{3\sqrt{1-\varepsilon}}{N_s}k\right|^2. \tag{7.16}$$

The average error is then given by,

$$\langle E_1^{\mathrm{b}}(\varepsilon)\rangle = \sum_{k=-N_s}^{N_s} P\left(\sum_{i=1}^{N_s}\mathrm{Tr}[Z\hat{\rho}_i] = k\right)\left|\frac{3\sqrt{1-\varepsilon}}{N_s}k\right|^2, \tag{7.17}$$

with $P(\mathrm{Tr}[Z\hat{\rho}_i] = 0) = \frac{2}{3}$ and $P(\mathrm{Tr}[Z\hat{\rho}_i] = 1) = P(\mathrm{Tr}[Z\hat{\rho}_i] = -1) = \frac{1}{6}$. Finding an analytic formula for $P\left(\sum_{i=1}^{N_s}\mathrm{Tr}[Z\hat{\rho}_i] = k\right)$ is quite challenging, but one can easily compute it numerically. In Fig. 7.1, I compute the relative mean squared error between the biased and unbiased cases by sampling Eq. (7.17) $N_{reps} = 10^6$ times for different values of N_s. The plot confirms our intuition that increasing the bias monotonically decreases the error.

The generalisation to the multiqubit case is straightforward as both the initial state and observable can be written as tensor products. Let $\rho = |+\rangle\langle+|^{\otimes N}$ and consider the q-local observable $O = Z^{\otimes q}$ with $\mathcal{I}$ being the set of qubit indices for which O is non-trivial. Noting that $\forall i, \hat{\rho}_i = \otimes_{j=1}^{N}\hat{\rho}_{i,j}$ the error reads,

$$\langle E_q^{\mathrm{b}}(\varepsilon)\rangle = \sum_{k=-N_s}^{N_s} P\left(\sum_{i=1}^{N_s}\left(\prod_{j\in\mathcal{I}}\mathrm{Tr}[Z\hat{\rho}_{i,j}]\right) = k\right)\left|\frac{3^q(1-\varepsilon)^{q/2}}{N_s}k\right|^2, \tag{7.18}$$

with the trace being equal to one when O acts trivially on ρ. Here again, one will need to numerically compute the error.

7.3.2 Worst-case scenario

The worst-case scenario happens when the studied quantum state ρ is an eigenstate of the Pauli observable of interest O such that $\mathrm{Tr}[O\rho] = \pm 1$. Without loss of generality, let $\rho = |+\rangle\langle+|$ and $O = X$ which satisfies $\mathrm{Tr}[O\rho] = 1$. Our estimator

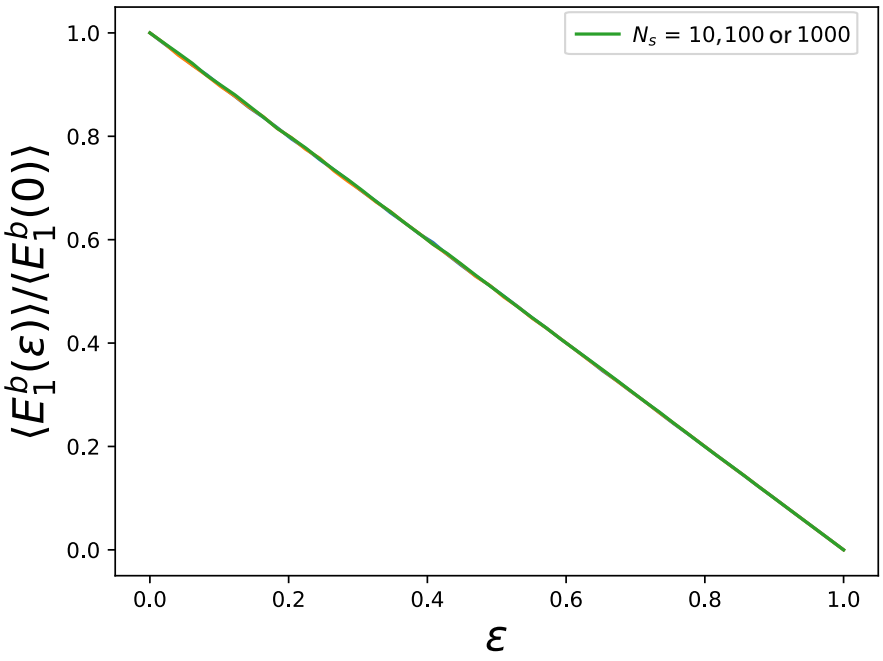

Figure 7.1: Numerical evaluation of the mean squared error $\langle E_1^{\mathrm{b}}(\varepsilon)\rangle$ of the biased scheme relative to the unbiased one when the aim is to estimate Pauli expected values $\mathrm{Tr}[O\rho]$ in the best-case scenario for an increasing number of shots N_s and bias parameters ε. As $\mathrm{Tr}[O\rho] = 0$ and biasing corresponds to shrinking the expectation value, we always reduce the relative error.

$\mathrm{Tr}[X\hat{\rho}_i]$ yields either $+1$ (when we measure in the Pauli X basis) or 0 (when we measure in the Y or Z bases). Suppose that k snapshots resulted in the outcome $+1$ while $N_s - k$ resulted in the outcome 0, I define the error in the mean value as,

$$E_1^{\mathrm{w}}(\varepsilon) \coloneqq \left| \frac{1}{N_s} \sum_{i=1}^{N_s} \mathrm{Tr}[X\hat{\rho}_i] - 1 \right|^2 = \left| \frac{3\sqrt{1-\varepsilon}}{N_s} k - 1 \right|^2 . \tag{7.19}$$

The average error is then given by,

$$\langle E_1^{\mathrm{w}}(\varepsilon)\rangle = \sum_{k=0}^{N_s} P\left(\sum_{i=1}^{N_s} \mathrm{Tr}[X\hat{\rho}_i] = k \right) \left| \frac{3\sqrt{1-\varepsilon}}{N_s} k - 1 \right|^2 ,$$

where $\sum_{i=1}^{N_s} \mathrm{Tr}[X\hat{\rho}_i]$ is simply a random variable following a binomial distribution $B(N_s, p)$ with $p = 1/3$. I can then obtain an analytic expression for $\langle E_1^{\mathrm{w}}(\varepsilon)\rangle$,

$$\langle E_1^{\mathrm{w}}(\varepsilon)\rangle = \sum_{k=0}^{N_s} \binom{N_s}{k} p^k (1-p)^{(N_s-k)} \left| \frac{3\sqrt{1-\varepsilon}}{N_s} k - 1 \right|^2 . \tag{7.20}$$

For the multiqubit case, let $\rho = |+\rangle\langle+|^{\otimes N}$ and $O = X^{\otimes q}$, the error in the

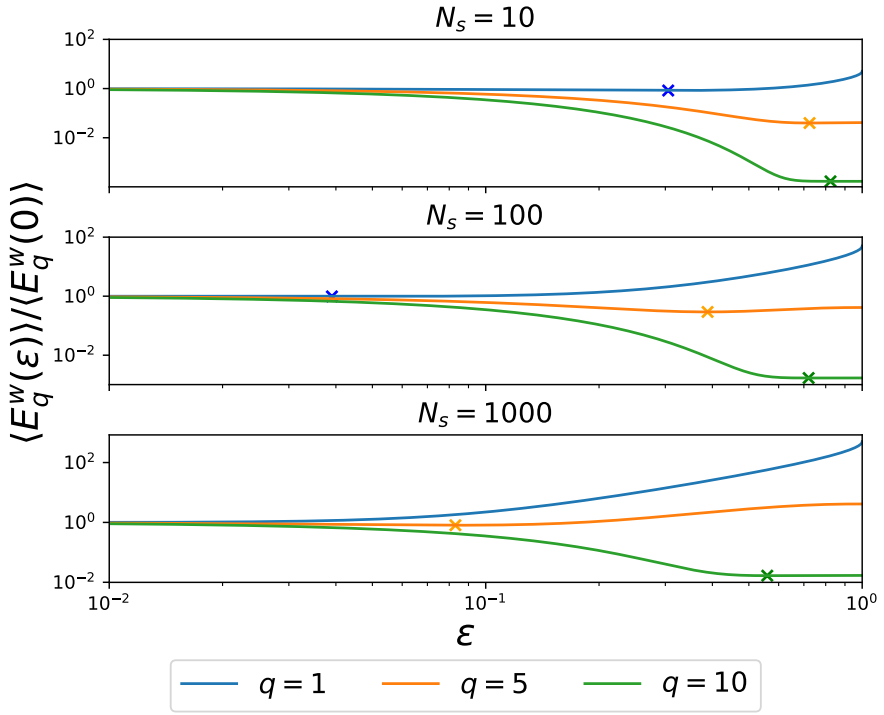

Figure 7.2: Mean squared error $\langle E_q^{\mathrm{w}}(\varepsilon)\rangle$ of the biased scheme relative to the unbiased one when the aim is to estimate Pauli expected values $\mathrm{Tr}[O\rho]$ in the worst case. I plot the analytical expression for an increasing number of shots, weight of the Pauli string q and bias parameters ε. There is an optimal ε (crosses at the minimum of the curves) for which the relative error is minimised and is smaller than 1. When $q = 1$ and $N_s = 1000$ the minimum is attained outside the plotted region.

worst-case now reads,

$$\langle E_q^{\mathrm{w}}(\varepsilon)\rangle = \sum_{k=0}^{N_s} P\left(\sum_{i=1}^{N_s}\left(\prod_{j\in\mathcal{I}}\mathrm{Tr}[X\hat{\rho}_{i,j}]\right) = k\right)\left|\frac{3^q(1-\varepsilon)^{q/2}}{N_s}k - 1\right|^2,$$

with the product of traces taken on qubit indices where O acts non-trivially. Here $\sum_{i=1}^{N_s}\left(\prod_{j\in\mathcal{I}}\mathrm{Tr}[X\hat{\rho}_{i,j}]\right) \sim B(N_s, p_q)$ with $p_q = 3^{-q}$. Hence,

$$\langle E_q^{\mathrm{w}}(\varepsilon)\rangle = \sum_{k=0}^{N_s}\binom{N_s}{k}p_q^k(1-p_q)^{(N_s-k)}\left|\frac{3^q(1-\varepsilon)^{q/2}}{N_s}k - 1\right|^2. \tag{7.21}$$

In Fig. 7.2, I plot the ratio between the biased and unbiased MSE $\langle E_q^{\mathrm{w}}(\varepsilon)\rangle/\langle E_q^{\mathrm{w}}(0)\rangle$ for different scenarios. From this plot we can draw several conclusions. First, it is *always worth* biasing as there is always an ε (crosses) for which the ratio is smaller than one (which confirms the results of Section 7.2.3). Then, the advantage of our method grows exponentially with the locality of the observable confirming the

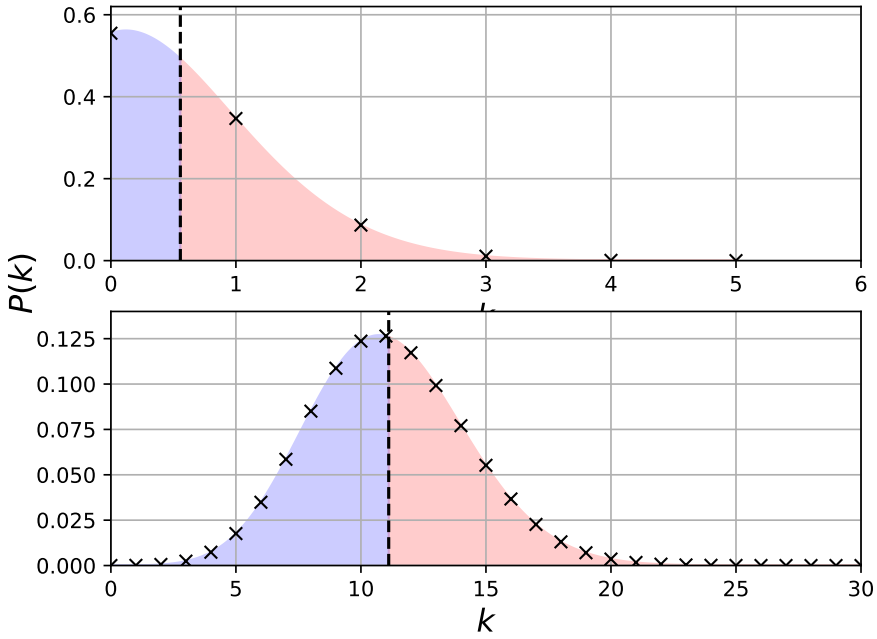

Figure 7.3: Binomial distribution for $p = 1/9$, $N_s = 5$ (top) and $N_s = 100$ (bottom). For a small N_s, it is more likely to obtain an outcome larger than the mean (black dashed line) which leads to a high mean-squared error which can be reduced by biasing the estimator. When N_s gets larger, biasing doesn't improve our estimates as the distribution becomes symmetric around the mean.

prediction of Eq. (7.14). This is due to the fact that the variance grows as 3^q in the unbiased case while it grows as $3^q(1-\epsilon)^q$ when introducing a bias. Finally, as one increases the number of shots N_s the biasing scheme becomes less and less favorable. This can be explained simply by noting that our estimator follows a binomial distribution $B(N_s, p_q)$. As showed in Fig. 7.3, the binomial distribution is not symmetric around the mean for small values of N_s. The MSE on the right-hand side of the mean is then larger than on the left-hand side. It is thus always worth biasing by rescaling our estimate. However, when N_s gets larger, the binomial distribution becomes symmetric and rescaling is not advantageous anymore.

7.4 Practical application

As an application of our scheme, let us consider the following situation. We want to compute the energy of a Hamiltonian H composed of only low-weight Pauli observables. In practice, one starts by preparing the ground state ρ and then

generating a collection of shadows $S(\rho, N_s) = \{\hat{\rho}_1, \ldots, \hat{\rho}_{N_s}\}$ from which we can estimate the ground state's energy E in postprocessing using the estimator $\bar{h}$,

$$E \approx \bar{h} = \frac{1}{N_s} \sum_i \mathrm{Tr}[H\hat{\rho}_i]. \tag{7.22}$$

Suppose now that we want to add a perturbative correction to the Hamiltonian in the form of a high-weight Pauli string P, as when computing a first-order correction to the energy in perturbation theory. While the collection of shadows we already generated allows us to estimate the expectation value of Pauli strings (of reasonable weight) beyond the ones present in the Hamiltonian, we cannot reliably estimate the correction to the energy due to the large variance of the shadow estimator for P. Including this observable will thus include more noise than signal, which leads to a SNR < 1. It is thus more sensible to not include $\bar{p}$ (defined in a similar way as $\bar{h}$) in our estimate of the energy and only use $\bar{h}$. The MSE of our estimate is then given by,

$$\mathrm{MSE}[\bar{h}] = \mathrm{Var}[\bar{h}] + \underbrace{\mathbb{E}[\bar{p}]^2}_{bias} \tag{7.23}$$

However, we saw in Eq. (7.13) that by biasing we can always improve the SNR of an estimator such that it contains more signal than noise. By optimally biasing the perturbation and introducing the new estimator $\bar{e} = \bar{h} + (1 - \sqrt{1 - \varepsilon^*})\bar{p}$ the new MSE is given by,

$$\begin{aligned} \mathrm{MSE}[\bar{e}] &= \mathrm{Var}[\bar{h}] + \mathrm{MSE}[(1 - \sqrt{1 - \varepsilon^*})\bar{p}] \\ &= \mathrm{Var}[\bar{h}] + \frac{1}{1+\beta}\mathbb{E}[\bar{p}]^2, \end{aligned} \tag{7.24}$$

such that $\forall \beta > 0, \mathrm{MSE}[\bar{e}] < \mathrm{MSE}[\bar{h}]$.

As a concrete example we consider a spin-ring Hamiltonian as

$$H = \sum_{k \in \mathrm{ring}(N)} \omega_k Z_k + J\vec{\sigma}_k \cdot \vec{\sigma}_{k+1}, \tag{7.25}$$

with coupling $J = 0.3$, on-site interaction strengths uniformly randomly generated in the range $-1 \leq \omega_k \leq 1$ and $\vec{\sigma}_k = (X_k, Y_k, Z_k)^T$ is a vector of single-qubit Pauli matrices. We prepare the ground state ρ of a 12-qubit Hamiltonian through a variational Hamiltonian ansatz of $l = 5$ layers. We generate a collection of shadows

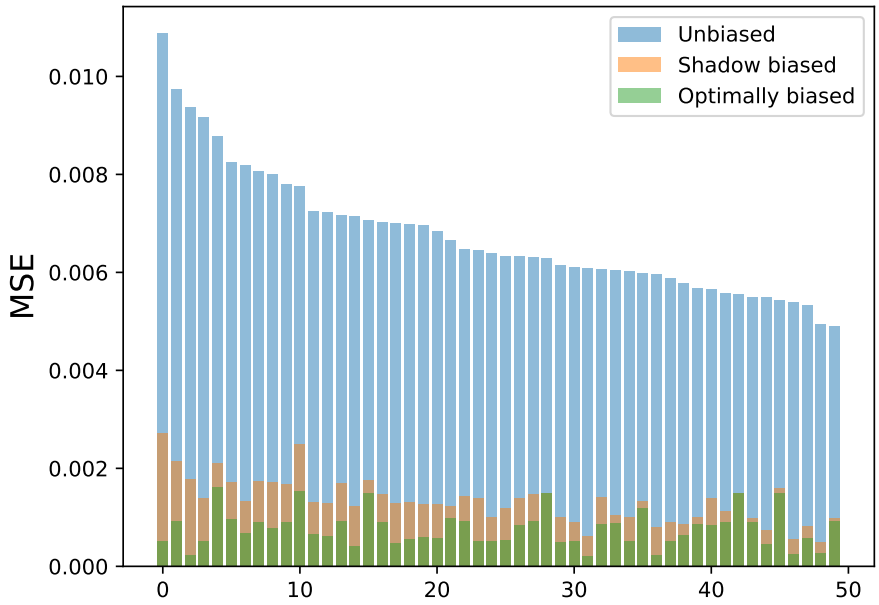

Figure 7.4: Illustration of the impact of different biasing strategies. In blue, I plot the MSE of fifty 8-local observables (whose SNR is smaller than 1) sorted by how large the error is. While analytically choosing the optimal bias according to Eq. (7.10) drastically reduces the error (green), this assumes previous knowledge of the exact expectation values, which is not available in practice. Through estimating the optimal bias from experimental data (orange), one still obtains a significant error reduction.

$S(N_s, \rho)$ with $N_s = 10^6$ and estimate the expected value $\mathrm{Tr}[(H + P)\rho]$ with respect to corrections P as Pauli observables of weight $q = 8$.

In Fig. 7.4, I plot the mean-squared error obtained by averaging over 10^5 repetitions and consider different 8-local observables with and without biasing (blue). I consider two different biasing strategies. First, I analytically choose the optimal bias parameter ε^* according to Eq. (7.10) (green) assuming direct access to the exact expected values (which one does not have access to in practice). Second, I use the experimentally estimated expected values to estimate the optimal biasing point in Eq. (7.10). While the latter deviates from the exact ε^* due to shot noise, Fig. 7.4 (orange) clearly demonstrates that the MSE is still significantly reduced compared to the unbiased scheme (see Appendix F.3 for a discussion on the critical biasing point). This demonstrates robustness against errors in determining the optimal bias parameter, i.e., the approach can still be used even when the optimal biasing point can only be estimated from the experimental data.

7.5 Discussion

In this chapter, I explored the benefits of deliberately introducing a bias in the classical shadows protocol in order to reduce statistical uncertainties due to the finite number of snapshots. It led to a simple and elegant scheme where one only needs to modify the estimator channel at the classical postprocessing step; no additional quantum resources are needed.

Through a rigorous analytical study, I derived the best- and worst-case scenarios and showed that even in the latter biasing is always advantageous. More precisely, we showed that given the knowledge of the optimal biasing parameter we can increase the SNR of any estimator. As knowing the optimal biasing parameter perfectly would be equivalent to knowing the expectation value, we do not have access to it. However, I showed for a practically motivated numerical simulation that an approximate estimate for ε^* obtained from experimental data was sufficient to reduce the MSE.

Our biasing scheme performs best for a small number of shots N_s and for the estimation of high-weight Pauli observables. This makes our technique well-suited for practical applications where one only has access to a relatively small number of samples. For instance, when estimating gradients [186, 187, 261, 262] or covariances [239] when training variational circuits or when estimating time-dependent properties [240]. One could also think about merging quantum error mitigated shadows [201, 263] and the biasing scheme developed here. While I only presented results related to the estimation of expectation values, it is worth noting that our scheme can be applied to shadow tomography where the aim is to recover a description of the full state [257]. Furthermore, my co-author showed that biasing also reduces the error of shadow tomography in the average-case.

This work represents the first attempt of using biased estimator to increase the performance of classical shadows. More generally, as information from quantum system can only be extracted through statistical estimators, developing biased estimators may be relevant for techniques beyond classical shadows [264, 265]. We believe that our work will spark interest in developing applications with minimal sample requirements due to more efficient biased estimators.

8
Conclusion

As the quality of quantum hardware continues to improve, it appears we could soon enter an era where quantum computers become practical. However, I believe additional theoretical advances will be essential to fully exploit the potential of near-term devices. The aim of this thesis is to address this challenge by exploring various components of the quantum computing stack, and developing practical techniques to enhance the capabilities of near-term silicon-based quantum computers.

Focusing first on hardware in Chapter 2, I developed a tool for characterising exchange coupling, a crucial parameter in the study of two-qubit gates, for arbitrary device geometries. This tool can be used to accelerate the development of silicon-based quantum computers by preventing bad designs from being sent to fabrication. Additionally, it enables the study of two-qubit gates' robustness against charge noise, which is considered to be one of the most detrimental sources of noise. To enhance this robustness, the tool facilitates the development of robust pulse sequences designed to increase the fidelity of two-qubit operations.

Assuming the capability to engineer multiple qubits on a single chip, I then address the challenge of global qubit control in Chapter 3. For Loss-DiVincenzo qubits, single-qubit gates are expected to be performed using a global magnetic field, inevitably leading to crosstalk between qubits. Through frequency homogenisation enabled by the exchange interaction or shuttling, I demonstrated that we can significantly increase the fidelity of such operations.

With both single- and two-qubit gates addressed, we can now consider potential architectures. In the quest for fault-tolerant quantum computers, it has become clear that an intermediate stage involving modular quantum computers will emerge. Without interlinks between them, the measurement overhead (shot noise) can be reduced through parallelisation. With interlinks, powerful quantum error mitigation techniques can be implemented, as discussed in Chapter 4. In this chapter, we propose a multicore architecture based on the distribution of Bell pairs between cores, and I demonstrated that Bell pair fidelities of approximately 99.5% can be achieved using entanglement purification. Even Bell pairs with this level of noise can be employed to implement error suppression by derangement, significantly improving the accuracy of expectation value estimates.

In the second part of the thesis, I focused on near-term applications of quantum computers, abstracting away from a specific type of qubit. In Chapter 6, I explored the recently introduced classical shadows framework and investigated how this protocol can be modified for practical use. By combining standard quantum error mitigation techniques and classical shadows, my co-authors and I showed that multiple noiseless properties of a quantum state can be recovered efficiently even when using noisy quantum devices. This protocol can be easily parallelised using a multicore architecture without interlinks. This approach is thus perfectly suited for silicon-based quantum computers where multiple independent cores could be fabricated on a single chip.

There are also other methods to reduce the number of measurements or equivalently improve accuracy for a given shot budget. In Chapter 7, I examined how the bias-variance tradeoff can be exploited to reduce shot noise in the context of classical shadows, making them even more practical. In this chapter, I assume ideal state preparation and thus access to noiseless properties of the prepared state, which is equivalent to the use of QEM shadows on noisy quantum computers.

Given the recent advancements in error correction, I am optimistic about the development of large-scale fault-tolerant quantum computers. However, I believe that the current leading platforms will face significant challenges as they scale.

For instance, superconducting qubits will need to be miniaturised, while ion traps and neutral atoms must improve their speed.

Despite being a relatively new platform, I am convinced that silicon-based qubits will play a crucial role in achieving fault-tolerance due to their scalability, small footprint, and fast operations. However, challenges remain, mainly due to the variability in qubit properties within silicon. As I started to explore in this thesis, co-designing both architectures and applications will be key to unlocking the full potential of silicon spin qubits. Platform-specific techniques should be leveraged more to design new fault-tolerant architectures. For instance, qubit shuttling, which has already proven to be essential in the development of logical qubits with neutral atoms, is likely to be equally invaluable for silicon spin qubits. I am excited to see what the future holds for this field, and I hope that this thesis has made a meaningful contribution to its progress.

Appendices

A

Appendix for Ab-initio modeling of quantum dot qubits

Contents

A.1 Device specifications 144
A.2 Abrupt transition 145
A.3 Charge trapped outside of the dot line 145
A.4 Fluctuating charges 146
A.5 Convergence versus basis size 148

A.1 Device specifications

For the device we consider (see Fig. 2.1), we used gates and oxide dimensions consistent with designs and expected fabricated dimensions from the IMEC SiMOS process. The silicon substrate is separated from the three gate layers by 8 nm thick silicon dioxide. Each gate layer is modelled as 30 nm thick Schottky contacts with a work function of 4.1 eV consistent with n++ doped polysilicon with 5 nm silicon dioxide separating each layer. The first layer (the closest to the silicon surface) is used to define the channel through confinement gates with a separation of 64 nm. The second layer is used for plunger gates to form our potential wells, they are separated by 30 nm and have a length of 54 nm (due to the oxide layer between gate layers) and width of 40 nm making an area of 2160 nm^2. The third layer incorporate

source and drain gates which yields lateral confinement to our channel, and a barrier gate which sits between both plunger gates, with length 44 nm and width of 20 nm. The plunger gates were set to 0.3 V and the barrier was swept from -0.1 to 0.4 V, all other gates were set to 0 V. The potential landscape 2D slice was taken 2 nm below the Si/SiO2 interface. This potential was generated by James Williams.

A.2 Abrupt transition

The grid-based method allows us to study any type of evolution of the potential in time. In the main text, I focus on the more realistic scenario where the barrier is smoothly lowered and raised to perform a $\sqrt{\text{SWAP}}$ gate. The intuition behind this approach is that by adiabatically modifying the potential, I can achieve a higher-quality two-qubit gate. To validate this choice, I compare it to the extreme case where the coupling is abruptly switched on. In Fig. A.1, it is clear that this abrupt change significantly increases p_{loss} (up to 1%). More precisely, by averaging the results over the remaining time after the gate, I find that $p_{loss} = 0.73\%$ and $F = 97.06\%$. This highlights the need for careful considerations of the control pulse.

A.3 Charge trapped outside of the dot line

As mentioned in the main body, I am not limited to studying charges located in the dot line, I can place them anywhere. In this subsection, I examine the performance of the $\sqrt{\text{SWAP}}$ gate for three different charge positions. To ensure a fair comparison, I position the charge at a fixed distance of 100 nm from the DQD, but at different angles: 0 (in the dot line), 45 and 90 degrees.

In Fig. A.2, I plot the evolution of the qubits' probability and p_{loss} for a given voltage profile (optimised for the noiseless case), and report the final fidelities and loss probabilities in Table A.1. The results correspond to our expectations, as the worst fidelity is observed when the charge is positioned along the dot line. It is worth noting that when the charge is placed perpendicularly to the dot line (black), the probability to find the electrons in the $|01\rangle$ and $|10\rangle$ does not intersect for this

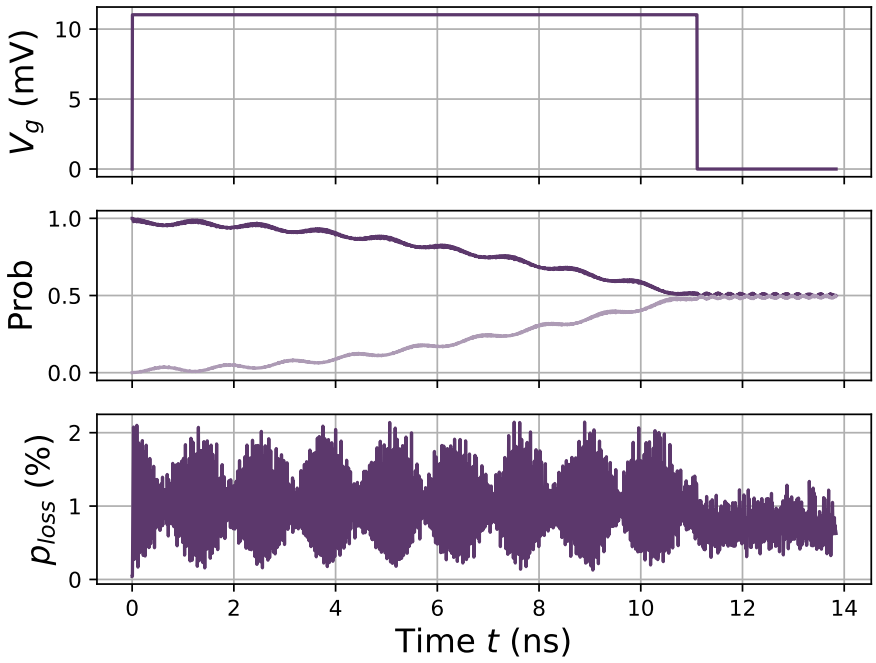

Figure A.1: Abrupt $\sqrt{\text{SWAP}}$ pulse. I abruptly switch the exchange interaction on and off and observe beatings due to the excitation of higher energy charge states. As expected the loss probability is highly increased with this method and can be as large as 1%.

Charge position (°)	0	45	90
p_{loss} (%)	0.05	0.05	0.04
F (%)	99.10	99.96	99.87

Table A.1: Loss probability and fidelity for different charge position. Each charge is placed at a fixed distance of 100 nm from the center of the DQD. Their position can thus be characterised by their angle with respect to the dot line (0° corresponds to having a charge in the dot line).

voltage profile, unlike the other charge environments studied in this chapter. This occurs because placing the charge as such pushes the trapped electrons apart from one another, rather than bringing them closer, which reduces the exchange coupling.

A.4 Fluctuating charges

In this section, I explore various scenarios involving multiple charge fluctuations between different positions, quantifying their impact on fidelity and loss probability, as illustrated in Fig. A.3. Employing a similar voltage control profile to that used in Fig. 2.6, I start with a charge placed at 100 nm of the DQD in the dot line,

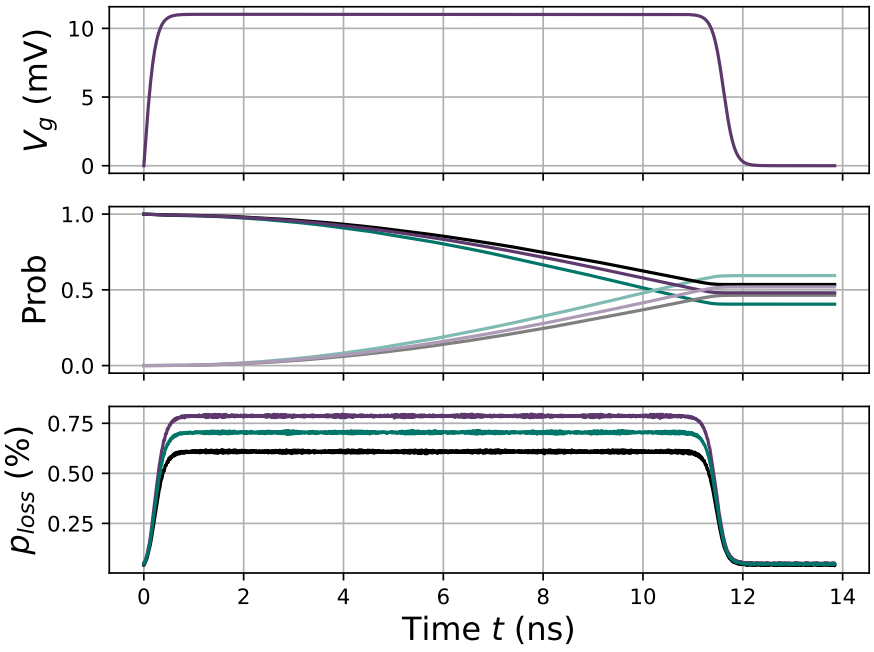

Figure A.2: Smooth $\sqrt{\text{SWAP}}$ gate for charges respectively positioned at $(100, 0, 2)$ nm (green), $(0, 100, 2)$ nm (black) and $(50\sqrt{2}, 50\sqrt{2}, 2)$ nm (purple).

leading to an initial $p_{loss} = 0.05\%$. I then perform a $\sqrt{\text{SWAP}}$ pulse and let the charge oscillates between this initial position and additional charge traps located at 90 nm, 75 nm, and 50 nm at the top, middle, and bottom of Fig. A.3, respectively.

The impact on fidelity is similar to that of the static case, with reductions attributed to timing errors, and the magnitude of the effect amplifying as the charge approaches the DQD. Notably, a significant departure from the static scenario lies in the loss of adiabaticity during the gate operation due to the dynamic potential changes. This is evident in the evolution of p_{loss}, exhibiting pronounced shifts whenever a fluctuation occurs. Consequently, there is a significant increase in p_{loss} at the end of the gate, constituting the most detrimental type of error as it cannot be corrected. Similar to fidelity, the closer the charge is to the DQD the more damaged the gate's performance. For oscillations between 100 and 50 nm, loss probabilities reach the order of the percent, rendering the gate impractical for use.

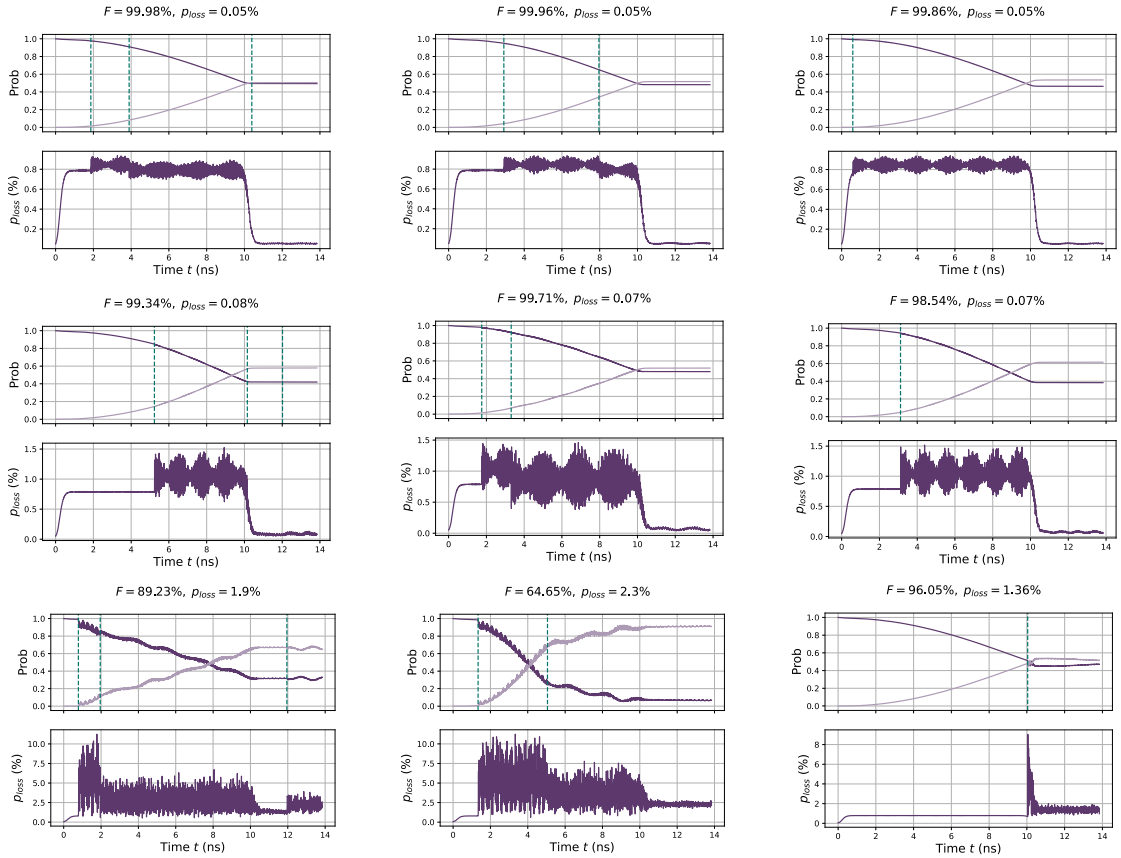

Figure A.3: Evolution of the probability to find the qubits in the states $|01\rangle$ and $|10\rangle$ and the loss probability for different fluctuations scenario. Initially the charge is located at $(100, 0, 2)$ nm and switches to another position at each green dashed lines. The voltage profile has been optimised for the $(100, 0, 2)$ nm case to yield a perfect $\sqrt{\text{SWAP}}$ gate. Oscillation between $(100, 0, 2)$ nm and $(90, 0, 2)$ nm at the top, $(100, 0, 2)$ nm and $(75, 0, 2)$ nm in the middle (bottom) $(100, 0, 2)$ nm and $(50, 0, 2)$ nm at the bottom. The fidelity and loss probability are obtained by averaging over the time following the end of the gate.

A.5 Convergence versus basis size

In the main body of the chapter, I show how the exchange coupling evolves with respect to the number of basis states in Fig. 2.3(a). I notice that when using 16 basis states (per particle, per dimension), the resulting couplings differ significantly from those at higher resolutions. However, there is very little difference between the 32 and 64 cases. Since our tool is computationally demanding, identifying the smallest resolution that still yields accurate results is crucial. To perform this task, Simon Benjamin modified his code to allow for an arbitrary number of basis states. The results plotted in Fig. A.4 show that for basis sizes greater than 22,

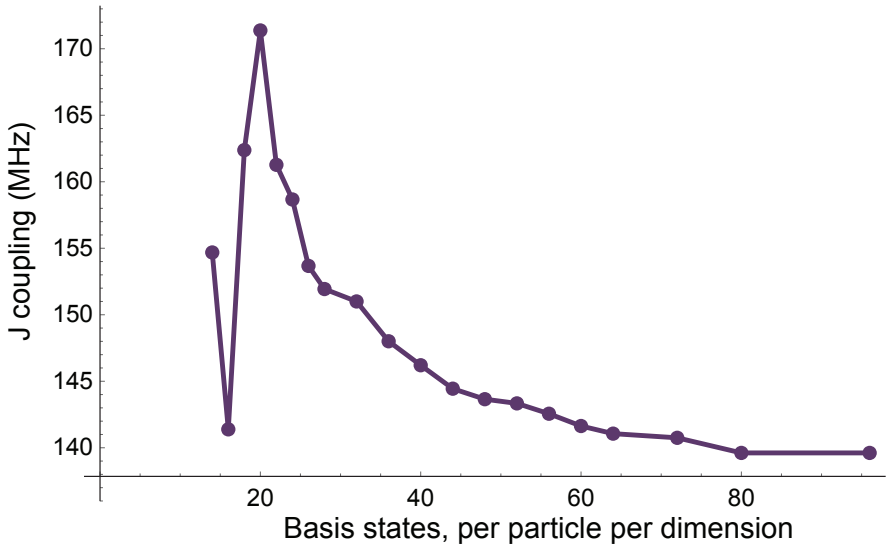

Figure A.4: Evolution of the J coupling for different basis set sizes with $V_g = 52\,\mathrm{mV}$. The exchange coupling converges to ≈ 140 MHz. The figure is extracted from [76] and shows results from my Simon Benjamin's code.

the exchange coupling steadily converges to ≈ 140 MHz. Below 22, the coupling is unstable but remains within 20% of the converged value.

B

Appendices for Harnessing transport for global spin qubit control

Contents

B.1 Binning 150

B.1 Binning

On top of the clustering methods introduced in Chapter 3, we can further enhance the fidelity of the global control operation by optimising the g-factor landscape. By placing our effective qubits in target and non-target bins using Stark shift, the idea is to maximise the separation between each qubit's frequency to reduce crosstalk. Given a set of effective g-factors $g_i^{'}$, the effective qubits are placed into bins as follows,

- If it is a target, shift its g-factor $g_i^{'}$ to $g_i^{''} = k\delta g$ with k the closest even number to $g_i^{'}/\delta g$
- If it is a non-target, shift its g-factor $g_i^{'}$ to $g_i^{''} = k\delta g$ with k the closest odd number to $g_i^{'}/\delta g$

It is important to emphasise that this protocol places the *effective* qubits into bins and not the physical qubits. After having modified the g-factors according to the

above protocol, the physical qubits will have distinct g-factors equal to $g_i + g_i^{''} - g_i^{'}$. It is only after performing the clustering protocols (exchange- or shuttling-based) that the g-factors can become equal. We then set the drive frequencies ω_i to match the final frequencies of the target qubits corresponding to their g-factors $g_i^{''}$.

Another advantage of using binning is that we can further reduce crosstalk by synchronising the rotations such that non-target qubits (approximately) undergo a $2p\pi$ rotation ($p \in \mathbb{N}$) while the target qubits receive an X gate (π rotation) [109]. This is done by choosing Ω such that,

$$\Omega = \frac{\delta g B_0}{2m},\ p \in \mathbb{N}. \tag{B.1}$$

Note that p is a free parameter and can thus be adjusted to account for realistic values of $\delta g, B_0$ and Ω.

C

Appendices for Multicore quantum computing

Contents

C.1 Parameter regime 152
C.2 Lindblad master equation 154
C.3 Time evolution 155
C.4 State prior to measurement 155
C.5 Alternative purification protocol 157
C.6 Average generation rate 157

C.1 Parameter regime

In my modelling I have explicitly—and numerically exactly—simulated the unitary dynamics generated by the Hamiltonain H in Eq. (4.3) in a resonant parameter regime, i.e., I did not use a rotating frame approximation. To ensure numerical stability I numerically exactly exponentiated the matrix $-itH$ after projecting the bosonic creation and annihilitation operators into the subspace spanned by the first seven Fock states (eigenstates with a fixed number of photons). Furthermore, I computed the matrix $-itH$ explicitly assuming a set of dimensionless parameters and the resulting optimised values are reported in Table C.1.

Recall that given a Hamiltonian $H = aH_1 + bH_2$ in terms of the dimensionless parameters a and b we can map the dynamics under the dimensionless time t to

description	symbol	unitless	physical value
magnetic field	B	100	10^4 MHz
transverse magnetic field	$b_{x,1}$	5.26	526 MHz
transverse magnetic field	$b_{x,2}$	1.74	174 MHz
charge cavity coupling	$g_{c,1}$	3.9	390 MHz
charge cavity coupling	$g_{c,2}$	1.4	140 MHz
DQD detuning	ϵ_1	0	0 MHz
DQD detuning	ϵ_2	0	0 MHz
optimal time	t	1.5	15 ns

Table C.1: Parameters used in the cavity model. These parameters were obtained from a gradient-based optimisation. B is the parallel magnetic field used to lift the spin degeneracy, $b_{x,1}$ ($b_{x,2}$) is the transverse magnetic field above the first (second) DQD which hybridises the spin and charge degrees of freedom, $g_{c,1}$ ($g_{c,2}$) is the dipole coupling constant between the cavity and the first (second) DQD and ϵ_1 (ϵ_2) the detuning.

any physical set of parameters via the symmetry

$$e^{-it(aH_1+bH_2)} = e^{-it/\lambda(\lambda aH_1+\lambda bH_2)} = e^{-it'(a'H_1+b'H_2)}.$$

Here I have introduced a new set of parameters $a' = \lambda a$, $b' = \lambda b$ and $t' = t/\lambda$, and the *arbitrary* scaling factor λ. I choose a possible scaling factor $\lambda := 100$ MHz such that the resulting optimal physical parameters fall into a regime that can be realised with near-term technology (see [266] for the possibility of reaching large charge-cavity couplings). Note, however, that the optimal unitary dynamics can be mapped to any other regime, e.g., choosing $\lambda := 200$ MHz allows a faster generation of Bell pairs ($t' = 7.5$ ns) but requires the fabrication of twice as large couplings (and all other dynamical parameters twice as large).

While the arbitrary choice of λ is an exact symmetry of the unitary dynamics, I can introduce a pseudosymmetry. As discussed above, I explicitly take into account and simulate the effect of the magnetic field B which is by orders of magnitude larger than all other dynamical parameters. Given that one could make the rotating frame approximation and remove this interaction without significantly affecting the dynamics, in numerical simulations I have confirmed that indeed increasing or decreasing the value of B by a small factor, e.g., 2 or 3, does only affect the

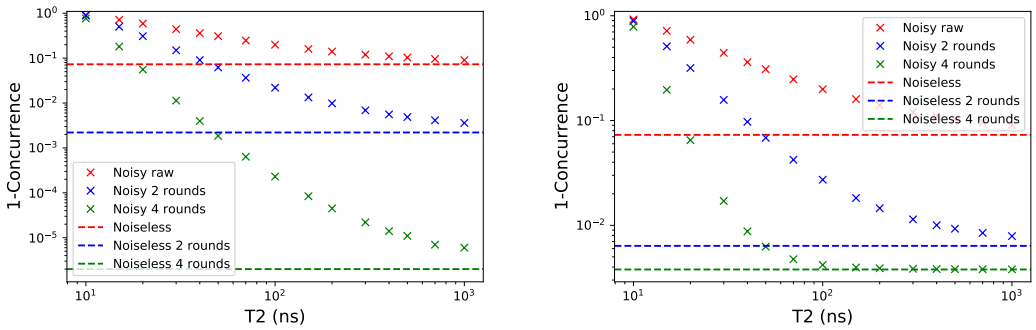

Figure C.1: Concurrence (measure of entanglement) as a function of charge decoherence for different levels of purification. Left (right) without (with) errors in the local quantum gates in the purification process. The dashed red line represents a perfect unitary evolution (no decoherence) and perfect charge measurements and the only source of imperfections is the fact that our fast, resonant dynamics in Eq. (4.3) cannot exactly generate a Bell pair.

resulting dynamics to a very small extent. This ensures us that an experimental system is robust to the choice of the B as long as the resonant condition is satisfied.

C.2 Lindblad master equation

The evolution of an d-dimensional quantum system in the state ρ with time is given by the Lindblad master equation,

$$\begin{aligned} \frac{\partial \rho(t)}{\partial t} &= \mathcal{L}\rho \\ &= -i\left[H,\rho\right] + \sum_i \gamma_i \left(A_i \rho A_i^\dagger - \frac{1}{2}\left\{ A_i^\dagger A_i, \rho \right\} \right), \end{aligned} \tag{C.1}$$

where γ_i are positive rates, and A_i Lindblad terms. Given the Hamiltonian is time independent I can exponentiate the superoperator $\mathcal{L}$, and the above equation can be rewritten as

$$\rho(t) = e^{\mathcal{L}t}\rho(0). \tag{C.2}$$

I can use standard numerical techniques for the exponentiation [267] if I represent ρ and $\mathcal{L}$ as a vector and a matrix respectively. To do so, I need to flatten ρ, meaning that I will stack the different rows one after the other. ρ becomes then a $1 \times d^2$ vector.

The superoperator $\mathcal{L}$ becomes a $d^2 \times d^2$ matrix, by applying the following rules,

$$\begin{aligned} A\rho &\rightarrow \mathbb{I} \otimes A\rho \\ \rho B &\rightarrow B^T \otimes \mathbb{I}\rho \\ A\rho B &\rightarrow B^T \otimes A\rho, \end{aligned} \tag{C.3}$$

where on the left-hand side I have the operation in the operator formalism and on the right-hand side the equivalent operation in the matrix-vector formalism.

Eq. (C.1) becomes then,

$$\begin{aligned} \frac{\partial \rho(t)}{\partial t} = -i \left(\mathbb{I} \otimes H - H^T \otimes \mathbb{I} \right) \rho + \sum_i \gamma_i \times \\ \left(\left(A_i^\dagger \right)^T \otimes A_i - \frac{1}{2} \mathbb{I} \otimes A_i^\dagger A_i - \frac{1}{2} \left(A_i^\dagger A_i \right)^T \otimes \mathbb{I} \right) \rho. \end{aligned} \tag{C.4}$$

C.3 Time evolution

There are two ways to numerically simulate Eq. (C.2). One way consists in recomputing the exponential $e^{\mathcal{L}t}$ for each time t and apply it to $\rho(0)$. I discarded this method as exponentiating $\mathcal{L}$, which is represented by a (12544×12544) matrix, is time-consuming. The other way relies on time discretization. In this case, as I only compute once the exponential for a small time step dt, and apply it sequentially to propagate the state, the simulation is much faster. For large time t, the first method is more accurate. However, in our case both techniques give similar results, reinforcing our choice to use the second method.

C.4 State prior to measurement

I start the dynamics with the initial quantum state as $|\psi_{init}\rangle = |{-}{-}\rangle_{ch} |{\uparrow}{\uparrow}\rangle_{sp} |0\rangle_{cav}$, such that the charge degrees of freedom are in the separable state $|{-}{-}\rangle_{ch}$, the spin degrees of freedom are in the separable state $|{\uparrow}{\uparrow}\rangle_{sp}$ and the cavity is separable in the vacuum state $|0\rangle_{cav}$.

Given we assume a resonant condition with the parameters $B_1 = B_2 = \omega_r \equiv B$ as well as $2t_{c,1} = 2t_{c,2} = 2t_c \equiv B$ with no detuning $\epsilon = 0$, the expected energy of

this initial state is exactly $E = 0$. Furthermore, for any other state we obtain from this initial state by pairs of "flips", e.g., $|-+\rangle_{ch}|\uparrow\downarrow\rangle|0\rangle_{cav}$, we obtain an identical energy $E = 0$ given any flip represents an energy quantum of $\pm B$. Assuming no decoherence, at our optimal evolution time in Fig. 4.3 I obtain the quantum state

$$|\psi\rangle = \sqrt{F}\Big(|\Psi^-\rangle_{ch}|\Psi^-\rangle_{sp}|0\rangle_{cav}\Big) + \sqrt{1-F}|\psi_{err}\rangle, \tag{C.5}$$

where with a high fidelity $F \approx 0.89$ I obtain our desired entangled state with a Bell pair $|\Psi^-\rangle = (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)/\sqrt{2}$ between the two spins up to some coherent error $|\psi_{err}\rangle$.

As we discussed in Section 4.3.2, this coherent error is present because we have fewer parameters of control in the Hamiltonian than the number of constraints the states need to satisfy in order be usefully entangled. Given the desired state above has energy $E = 0$, as it can formally be obtained by an even number of energy flips from the initial state, the coherent error state $|\psi_{err}\rangle$ must also have energy $E = 0$ due to energy conservation under a unitary evolution. While $|\psi_{err}\rangle$ is a complex quantum state that is a superposition of nearly all charge and spin configurations, we find that the dominant contribution in $|\psi_{err}\rangle$ is the state

$$|\Phi^-\rangle_{ch}|\downarrow\downarrow\rangle_{sp}|1\rangle_{cav}, \tag{C.6}$$

which occurs with a probability $F' \approx 0.05$ as the fidelity with respect to the full state $|\psi\rangle$. Indeed this state has energy $E = 0$, however, given the charge is entangled as $|\Phi^-\rangle_{ch} = (|LL\rangle - |RR\rangle)/\sqrt{2}$, energy conservation forces the spins into the separable state $|\downarrow\downarrow\rangle_{sp}$ while the cavity also obtains a photon.

This motivates our measurement-based protocol: I measure the cavity state and only accept the outcomes $|LR\rangle$ or $|RL\rangle$ which does not affect the ideal contribution of the state in Eq. (C.5), however, it eliminates the dominant coherent error contribution from Eq. (C.6). This necessarily increases entanglement between the spins given the spins are separable in the error contribution in Eq. (C.6) that I eliminate. Furthermore, the probability of an empty cavity is also increased given the state I project out has a non-empty cavity.

C.5 Alternative purification protocol

By adding single S gates to the three last rounds of purification, one can obtain a better purification scheme than the one presented in the main text for the case of the cavity-based approach. In Fig. C.1, I note that the concurrence is slightly improved in the ideal case. However, in the noisy case, I increase the noise (and thus lower the concurrence) by adding these new gates. With better hardware, one would transition from the scheme presented in the main text to this one.

C.6 Average generation rate

Assuming a raw Bell pair is generated with time $T_{raw} = 15$ ns and I neglect gate and measurement times in the purification process, the time to generate a purified Bell pair is $T_{raw}N_{avg}$ where the average number of required raw Bell pairs is N_{avg}. More realistically, the average time for generating a single purified Bell pair is

$$T_{avg} = T_{raw}\, N_{avg} + T_{1qb}\, \frac{N_{1qb,avg}}{4} + T_{2qb}\, \frac{N_{2qb,avg}}{2} + T_{meas}\, \frac{N_{meas,avg}}{2},$$

which depends on the number of single-qubit gates $N_{1qb,avg} = 13.6$, on the number of two-qubit gates and on measurements $N_{2qb,avg} = N_{meas,avg} = 6.8$. In these figures, I assumed $T^*_{2,c} = 400$ ns and I have divided $N_{1qubit,avg}$ by 4 and $N_{2qubit,avg}, N_{meas,avg}$ by 2 as the four single-qubit gates, the two two-qubit gates and the two measurements used in a purification round can be performed in parallel. The generation rate is then given by $g_{Bell} = [T_{avg}]^{-1}$. Upon substituting typical state-of-the-art values for one-, two-qubit gate and measurement times as $T_{1qubit} = 1$ µs, $T_{2qubit} = 100$ ns, $T_{meas} = 1$ µs [46, 268–270], I obtain an average generation rate of $g_{Bell} \approx 0.14$ MHz.

As I discussed in the main text, for smaller values of $T^*_{2,c}$, one needs to perform four rounds of purification. On average, this corresponds to $N_{avg} \approx 25.4$ and 33.0 raw Bell pairs to successfully go through the four rounds, leading to average generation rates of $g_{Bell} \approx 2.6$ MHz and $g_{Bell} \approx 2.0$ MHz ($g_{Bell} \approx 24.9$ KHz

and $g_{Bell} \approx 20.8$ KHz by incorporating gate time and measurement time) for $T^*_{2,c} = 100$ ns and $T^*_{2,c} = 50$ ns respectively.

D

Appendices for Introduction to quantum error mitigation and classical shadows

Contents

D.1 Cost of error mitigation 159
D.2 Details on probabilistic error cancellation 160
D.3 Details on classical shadows 161

D.1 Cost of error mitigation

Suppose we want to compute the expectation value of an observable O with respect to the state ρ_{id} on a quantum computer. In practice, we only have access to the noisy state ρ. We characterise the error made in evaluating the expectation value of O using a noisy device using the mean squared error (MSE). As quantum measurement is a stochastic process, each time we run the quantum computer, we will obtain a different value for the expectation value of the observable. Suppose that we run the computation N_s times and let $\hat{o}$ be the value obtained from an individual run. Let $\bar{o}$ be the mean estimator of $\mathrm{Tr}[O\rho]$ constructed from taking the average of the $\hat{o}$ obtained at each run and $o_{id} = \mathrm{Tr}[O\rho_{id}]$ the ideal expectation

value, the MSE of $\bar{o}$ is given by,

$$\begin{aligned}\mathrm{MSE}[\bar{o}] &= \mathbb{E}[(\bar{o} - o_{id})^2] \\ &= (\mathbb{E}[\bar{o}] - o_{id})^2 + \mathrm{Var}[\bar{o}] \\ &= (\mathbb{E}[\bar{o}] - o_{id})^2 + \frac{\mathrm{Var}[\hat{o}]}{N_s}\end{aligned} \tag{D.1}$$

with $\mathbb{E}[\bar{o}] - o_{id}$ being the bias of our estimator. Note that the variance part of the MSE (also called shot noise) can be made arbitrarily small by increasing the number of measurement. The error bottleneck resides then in the bias. Error mitigation's goal is to reduce this bias by constructing a new estimator $\bar{o}_{em}$ using clever post-processing tricks. However, this usually comes at the cost of an increased variance $\mathrm{Var}(\bar{o}_{em})$.

To estimate the sampling overhead or cost C_{em} introduced by error mitigation, we compare the number of shots needed for both estimators to reach the same level of shot noise ε. It is important to note that while we will be ε-close to the ideal value for the error-mitigated estimator, we will be ε-close to the bias for the naive estimator. Let $\hat{o}_{em}$ be the *single-shot* error-mitigated observable, defined as $\hat{o}_{em} = N_s \bar{o}_{em}$. We have $\mathrm{Var}[\hat{o}] = N_s \mathrm{Var}[\bar{o}]$ and $\mathbb{E}[\hat{o}] = \mathbb{E}[\bar{o}]$. To obtain a shot noise level ε for $\hat{o}$, one needs $N_s^{\varepsilon}(\hat{o}) = \mathrm{Var}[\hat{o}]/\varepsilon^2$ shots. The sampling overhead is thus given by,

$$C_{\mathrm{em}} = \frac{N_s^{\varepsilon}(\hat{o}_{em})}{N_s^{\varepsilon}(\hat{o})} = \frac{\mathrm{Var}[\hat{o}_{em}]}{\mathrm{Var}[\hat{o}]} \tag{D.2}$$

D.2 Details on probabilistic error cancellation

Proof of Lemma 2. Let $\hat{\rho}_{id}$ be the single-shot estimator of the state ρ_{id} as defined in Lemma 2. Then,

$$\begin{aligned}\mathbb{E}[\hat{\rho}_{id}] = \mathop{\mathbb{E}}_{\underline{k}}[\hat{\rho}_{\underline{k}}] &= \sum_{\underline{k}} p_{\underline{k}} \|\gamma\|_1 \mathrm{sign}(\gamma_{\underline{k}}) \mathcal{G}_{\underline{k}} \left|0\right\rangle\!\left\langle 0\right| \\ &= \sum_{\underline{k}} \frac{|\gamma_{\underline{k}}|}{\|\gamma\|_1} \|\gamma\|_1 \mathrm{sign}(\gamma_{\underline{k}}) \mathcal{G}_{\underline{k}} \left|0\right\rangle\!\left\langle 0\right| \\ &= \left(\sum_{\underline{k}} \gamma_{\underline{k}} \mathcal{G}_{\underline{k}} \right) \left|0\right\rangle\!\left\langle 0\right| \\ &= \mathcal{U}_{circ} \left|0\right\rangle\!\left\langle 0\right| \\ &= \rho_{id}\end{aligned} \tag{D.3}$$

where we used $|\gamma_{\underline{k}}|\text{sign}(\gamma_{\underline{k}}) = \gamma_{\underline{k}}$ in the third line. □

D.3 Details on classical shadows

Proof of Theorem 1. Let $\hat{\rho}_l$ be the single-shot estimator of the state ρ as defined in Theorem 1. Then,

$$\begin{aligned}
\mathop{\mathbb{E}}_{l}[\hat{\rho}_l] &= \sum_l q_l C_E^{-1}(E_l) \\
&= \sum_l \text{Tr}(E_l\rho) C_E^{-1}(E_l) \\
&= C_E^{-1}\left(\sum_l \text{Tr}(E_l\rho) E_l\right) \\
&= C_E^{-1}\left(C_E(\rho)\right) \\
&= \rho
\end{aligned} \tag{D.4}$$

Where we used the linearity of C_E^{-1} in the second line. □

E

Appendices for Quantum error mitigated shadows

Contents

E.1 Details on probabilistic error cancelled shadows 163

E.1.1 Unbiased estimator 163

E.1.2 Shadow norm 163

E.1.3 Performance guarantees 165

E.2 Details on classical post-processing algorithms 170

E.2.1 Expectation value of local Paulis 170

E.2.2 Improved algorithm using light cones 171

E.2.3 Local purities 171

E.1 Details on probabilistic error cancelled shadows

E.1.1 Unbiased estimator

Proof of Theorem 2. Let $\hat{\rho}_{id}$ be the single-shot estimator of the state ρ_{id} as defined in Theorem 2. Then,

$$\begin{aligned}
\mathbb{E}[\hat{\rho}_{id}] = \mathop{\mathbb{E}}_{\underline{k},l}\left[\hat{\rho}_{\underline{k},l}\right] &= \sum_{\underline{k}}\sum_{l} q_l p_{\underline{k}} \hat{\rho}_{\underline{k},l} \\
&= \sum_{\underline{k}}\sum_{l} \mathrm{Tr}\left[\mathcal{G}_{\underline{k}}\left|0\right\rangle\!\left\langle 0\right| E_l\right] \frac{|\gamma_{\underline{k}}|}{\|\gamma\|_1} \|\gamma\|_1 \mathrm{sign}(\gamma_{\underline{k}}) C_E^{-1}(E_l) \\
&= \sum_{l} \mathrm{Tr}\left[\sum_{\underline{k}} |\gamma_{\underline{k}}| \mathrm{sign}(\gamma_{\underline{k}}) \mathcal{G}_{\underline{k}}\left|0\right\rangle\!\left\langle 0\right| E_l\right] C_E^{-1}(E_l) \\
&= \sum_{l} \mathrm{Tr}\left[\left(\sum_{\underline{k}} \gamma_{\underline{k}} \mathcal{G}_{\underline{k}}\right)\left|0\right\rangle\!\left\langle 0\right| E_l\right] C_E^{-1}(E_l) \\
&= \sum_{l} \mathrm{Tr}\left[\mathcal{U}_{circ}\left|0\right\rangle\!\left\langle 0\right| E_l\right] C_E^{-1}(E_l) \\
&= \sum_{l} \mathrm{Tr}\left[\rho_{id} E_l\right] C_E^{-1}(E_l) \\
&= C_E^{-1}\left(\sum_{l} \mathrm{Tr}\left[\rho_{id} E_l\right] E_l\right) \\
&= C_E^{-1}\left(C_E(\rho_{id})\right) \\
&= \rho_{id}
\end{aligned} \tag{E.1}$$

where we used the linearity of the trace and of C_E^{-1} in the third and the seventh line respectively. □

E.1.2 Shadow norm

In this section we define the shadow norm of a q-local observable O and derive an analytic formula for the experimentally friendly case of random single-qubit Pauli basis transformation.

Lemma 5 (Shadow norm)**.** *Given a POVM E and a q-local observable O, the shadow norm of O with respect to E is given by,*

$$\|O\|_E^2 \coloneqq \left\| \sum_{l=1}^{N_E} \mathrm{Tr}[O\hat{\rho}_l]^2 E_l \right\|_\infty , \tag{E.2}$$

with $\|\cdot\|_\infty$ *returning the maximal eigenvalue of the operator and* $\hat{\rho}_l = C_E^{-1}(E_l)$. *In the case of uniformly random single-qubit Pauli basis transformation we have* $\|O\|_E^2 = 3^q$.

Proof. Ref. [220] showed that when using uniformly random single-qubit Pauli basis transformations, the effect E_l (with $l = (j, \underline{b})$) for a single qubit is given by,

$$E_l = \frac{1}{3} Q_j^\dagger \left|b\right\rangle\!\left\langle b\right| Q_j, \tag{E.3}$$

with $b \in \{0, 1\}$ and Q_j is a unitary matrix corresponding to changing the measurement basis to the X, Y or Z basis. Note that in this case $N_E = 6$. We can then rewrite Eq. (E.3) as,

$$E_l = \frac{1}{3} \left|t^\pm\right\rangle\!\left\langle t^\pm\right|, \tag{E.4}$$

with $|t^\pm\rangle$ being the eigenvector corresponding to the ± 1 eigenvalue of the Pauli t operator with $t \in \{X, Y, Z\}$. Furthermore, ref. [220] shows that due to the symmetry of the measurement,

$$\hat{\rho}_l = 9E_l - \mathbb{1}. \tag{E.5}$$

Suppose that O is a single-qubit Pauli observable,

$$\begin{aligned}
\|O\|_E^2 &:= \left\| \sum_{l=1}^{6} \mathrm{Tr}[O\hat{\rho}_l]^2 E_l \right\|_\infty \\
&= \left\| \sum_{l=1}^{6} (9\mathrm{Tr}[OE_l] - \mathrm{Tr}[O])^2 E_l \right\|_\infty \\
&= \left\| \sum_{t^\pm} 3 \left(\mathrm{Tr}[O \left|t^\pm\right\rangle\!\left\langle t^\pm\right|] \right)^2 \left|t^\pm\right\rangle\!\left\langle t^\pm\right| \right\|_\infty \\
&= 3 \left\| \sum_{t^\pm} \left\langle t^\pm\right| O \left|t^\pm\right\rangle^2 \left|t^\pm\right\rangle\!\left\langle t^\pm\right| \right\|_\infty
\end{aligned} \tag{E.6}$$

where we used the fact that $\mathrm{Tr}[O] = 0$ as O is a Pauli operator in the third line. As $O \in \{X, Y, Z\}$, there is a unique t_0 such that $O = \left|t_0^+\right\rangle\!\left\langle t_0^+\right| - \left|t_0^-\right\rangle\!\left\langle t_0^-\right|$ with $|t_0^\pm\rangle$ being the normalised eigenstates of O. This means that only two terms remain in $\sum_{t^\pm} \langle t^\pm| O |t^\pm\rangle^2 |t^\pm\rangle\langle t^\pm|$ due to the orthogonality of the eigenstates such that,

$$\sum_{t^\pm} \left\langle t^\pm\right| O \left|t^\pm\right\rangle^2 \left|t^\pm\right\rangle\!\left\langle t^\pm\right| = \left|t_0^+\right\rangle\!\left\langle t_0^+\right| + \left|t_0^-\right\rangle\!\left\langle t_0^-\right| = \mathbb{1}. \tag{E.7}$$

The shadow norm is thus equal to,

$$\|O\|_E^2 = 3\|\mathbb{1}\|_\infty = 3 \tag{E.8}$$

for a single-qubit Pauli observable. Consider now a N-qubit state ρ and a q-local observable O. O can then be written as a tensor product of q single-qubit Pauli observables and $N-q$ single-qubit identities as $\bigotimes_{j=1}^N O_j$. Using single-qubit measurements, the effects E_l can also be written as a tensor product and as a consequence, the shadow norm factorises as $\|O\|_E^2 = \prod_{j=1}^N \|O_j\|_E^2$. Finally, using Eq. (E.8) and the fact that if $O = \mathbb{1}$, then $\|O\|_E^2 = 1$, we get $\|O\|_E^2 = 3^q$. □

E.1.3 Performance guarantees

Variance of linear properties

Proof of Lemma 3. Let $\hat{\rho}_{id}$ be the unbiased estimator of the state ρ_{id} as defined in Theorem 2 and O an observable. We wish to bound the variance of the estimator $\hat{o} = \mathrm{Tr}[O\hat{\rho}_{id}]$. First, note that,

$$\begin{aligned}
\mathrm{Var}[\hat{o}] &= \mathbb{E}[(\hat{o} - \mathbb{E}[\hat{o}])^2] \\
&= \mathbb{E}[\hat{o}^2] - o^2 \\
&\leq \mathbb{E}[\hat{o}^2]
\end{aligned} \tag{E.9}$$

where we use the fact that $\hat{o}$ is an unbiased estimator of the ideal expectation value $o = \mathrm{Tr}[O\rho_{id}]$. We thus only need to bound $\mathbb{E}[\hat{o}^2]$,

$$\begin{aligned}
\mathbb{E}[\hat{o}^2] &= \mathop{\mathbb{E}}_{\underline{k},l}[\mathrm{Tr}[O\hat{\rho}_{\underline{k},l}]^2] \\
&= \mathop{\mathbb{E}}_{\underline{k},l}[\mathrm{Tr}[O\|\gamma\|_1 \mathrm{sign}(\gamma_{\underline{k}})\hat{\rho}_l]^2] \\
&= \|\gamma\|_1^2 \mathop{\mathbb{E}}_{\underline{k},l}[\mathrm{Tr}[O\hat{\rho}_l]^2] \\
&= \|\gamma\|_1^2 \sum_{\underline{k}} \sum_l q_l p_{\underline{k}} \mathrm{Tr}[O\hat{\rho}_l]^2 \\
&= \|\gamma\|_1^2 \sum_{\underline{k}} \sum_l \mathrm{Tr}[\mathcal{G}_{\underline{k}}(|0\rangle\langle 0|)E_l] \frac{|\gamma_{\underline{k}}|}{\|\gamma\|_1} \mathrm{Tr}[O\hat{\rho}_l]^2 \\
&= \|\gamma\|_1^2 \sum_l \mathrm{Tr}[\Omega(|0\rangle\langle 0|)E_l] \mathrm{Tr}[O\hat{\rho}_l]^2,
\end{aligned} \tag{E.10}$$

where we introduced the quantum channel $\Omega \coloneqq \|\gamma\|^{-1} \sum_{\underline{k}} |\gamma_{\underline{k}}| \mathcal{G}_{\underline{k}}$. It is indeed a valid quantum channel as Ω is defined as a convex combination of valid quantum channels. We can then upper bound $\mathbb{E}[\hat{o}^2]$, by replacing the initial state $|0\rangle\langle 0|$ by a maximisation over all states σ,

$$\begin{aligned}\mathbb{E}[\hat{o}^2] &\leq \|\gamma\|_1^2 \max_{\sigma} \sum_l \mathrm{Tr}[\Omega(\sigma) E_l] \mathrm{Tr}[O\hat{\rho}_l]^2 \\ &= \|\gamma\|_1^2 \max_{\sigma} \mathrm{Tr}\left[\Omega(\sigma) \sum_l \mathrm{Tr}[O\hat{\rho}_l]^2 E_l\right] \\ &= \|\gamma\|_1^2 \max_{\sigma} \mathrm{Tr}\left[\Omega(\sigma)\Theta\right], \end{aligned} \tag{E.11}$$

where we introduced $\Theta = \sum_l \mathrm{Tr}[O\hat{\rho}_l]^2 E_l$. Finally, as $\Omega(\sigma)$ is a valid density matrix, $\mathrm{Tr}[\Omega(\sigma)\Theta]$ can be upper bounded by the operator norm $\|\Theta\|_\infty$,

$$\begin{aligned}\mathbb{E}[\hat{o}^2] &\leq \|\gamma\|_1^2 \|\Theta\|_\infty \\ &= \|\gamma\|_1^2 \|O\|_E^2, \end{aligned} \tag{E.12}$$

where we used the definition of the shadow norm from Lemma 5. □

Sample complexity for predicting linear properties

We use the median of means estimator to predict the expectation values of M independent observables $O_1, \ldots, O_M$. This consists in dividing the classical shadows $S(\rho_{id}, N_s)$ into K batches $\{\mathcal{B}_1, \ldots \mathcal{B}_K\}$ of equal size N_{batch}, and taking as estimate for $\mathrm{Tr}[O_j \rho_{id}]$ the median of the empirical means $\{\hat{\mu}_i(O_j) = N_{batch}^{-1} \sum_{k \in \mathcal{B}_i} \mathrm{Tr}[O(\hat{\rho}_{id})_k]\}_{1 \leq i \leq K}$. We note

$$\hat{\mu}_K(O_j) = \mathrm{median}\{\hat{\mu}_1(O_j), \ldots, \hat{\mu}_K(O_j)\}, \tag{E.13}$$

the median of means estimator. While requiring more snapshots, this method is more resilient to outliers. For $\hat{\mu}_K(O_j)$ to deviate more than ε from $\mathrm{Tr}[O_j \rho_{id}]$, more than $K/2$ of the individual empirical means $\hat{\mu}_i(O_j)$ must deviate by more than ε. Using the concentration property of the estimator we can show that the probability of such an event is exponentially reduced [271, 272],

$$\mathrm{Prob}[|\hat{\mu}_K(O_j) - \mathrm{Tr}[O_j \rho_{id}]| > \frac{2\sigma}{\sqrt{N_{batch}}}] \leq \exp\left(\frac{-K}{8}\right), \tag{E.14}$$

with σ the standard deviation.

We now introduce a more formal version of Theorem 3

Theorem 4 (Formal version of Theorem 3)**.** *Given the PEC shadows* $H \coloneqq \{(\gamma_{\underline{k}}, \mathcal{G}_{\underline{k}}, E_l)\}_{\underline{k},l}$ *from Theorem 2, we wish to simultaneously estimate the expectation values of* M *observables* $O_1, \dots O_M$ *with respect to* ρ_{id}*. For performance parameters* $\varepsilon, \delta \in [0,1]$*, let*

$$N_{batch} = \frac{4\|\gamma\|_1^2}{\varepsilon^2} \max_{1\leq j\leq M} \|O_j\|_E^2 \quad \textit{and} \quad K = 8\log\Big(\frac{M}{\delta}\Big). \tag{E.15}$$

Using a collection of shadows of size $N_s = KN_{batch}$*, we can predict the ideal expectation values of the* M *observables using the median of means estimator with an accuracy* ε *and a probability of failure* δ *such that,*

$$\text{Prob}[\forall\ 1 \leq j \leq M,\ |\hat{\mu}_K(O_j) - \text{Tr}[O_j\rho_{id}]| \leq \varepsilon] \geq 1 - \delta. \tag{E.16}$$

Proof. The proof is derived from the concentration property of the median of means estimator coupled to the variance bound from Lemma 3. Setting $\varepsilon = 2\sigma/\sqrt{N_{batch}}$ in Eq. (E.14), we obtain $N_{batch} = 4\sigma^2/\varepsilon^2$. As $\sigma^2 = \text{Var}[\hat{\mu}_K(O_j)] \leq \|\gamma\|_1^2\|O_j\|_E^2$, we take $N_{batch} = 4\|\gamma\|_1^2\|O_j\|_E^2/\varepsilon^2$. To account for the fact that we are considering M different observables which require different values of N_{batch}, we set $N_{batch} = 4\|\gamma\|_1^2/\varepsilon^2 \max_{1\leq j\leq M}\|O_j\|_E^2$. Using the concentration property we have,

$$\forall\ 1 \leq j \leq M,\ \text{Prob}[|\hat{\mu}_K(O_j) - \text{Tr}[O_j\rho_{id}]| > \varepsilon] \leq \exp\Big(\frac{-K}{8}\Big). \tag{E.17}$$

Hence, by choosing $K = 8\log(M/\delta)$ and using the union bound we get,

$$\begin{aligned}
\text{Prob}[\exists\ 1 \leq j \leq M,\ |\hat{\mu}_K(O_j) - \text{Tr}[O_j\rho_{id}]| > \varepsilon] &= \text{Prob}[\bigcup_{j=1}^{M} |\hat{\mu}_K(O_j) - \text{Tr}[O_j\rho_{id}]| > \varepsilon] \\
&\leq \sum_{j=1}^{M} \text{Prob}[|\hat{\mu}_K(O_j) - \text{Tr}[O_j\rho_{id}]| > \varepsilon] \\
&\leq M\frac{\delta}{M} = \delta,
\end{aligned} \tag{E.18}$$

which leads to,

$$\text{Prob}[\forall\ 1 \leq j \leq M,\ |\hat{\mu}_K(O_j) - \text{Tr}[O_j\rho_{id}]| \leq \varepsilon] \geq 1 - \delta. \tag{E.19}$$

□

Sample complexity for predicting non-linear properties

Proof of Lemma 4. First, recall that any polynomial function in a quantum state ρ can be written as a linear function in tensor products of ρ. In the main text, we took the example $\mathrm{Tr}[O\rho^2] = \mathrm{Tr}[\bar{O}\rho \otimes \rho]$ where $\bar{O} = \mathrm{SWAP}O_1$ with $O_1 = O \otimes \mathbb{1}$ (the subsystem on which O acts is not relevant). More generally, for a given observable O we have,

$$\mathrm{Tr}[O\rho^m] = \mathrm{Tr}[\bar{O}\rho^{\otimes m}], \tag{E.20}$$

with $\bar{O} = C^{(m)}O_i$ and $C^{(m)}$ is the cyclic permutation operator,

$$C^{(m)}(|\psi_1\rangle |\psi_2\rangle \dots |\psi_m\rangle) = |\psi_m\rangle |\psi_1\rangle \dots |\psi_{m-1}\rangle \tag{E.21}$$

with $O_i = \mathbb{1} \otimes \cdots \otimes O \otimes \cdots \otimes \mathbb{1}$ representing the observable O acting the ith.

We now introduce U-statistics which gives a uniformly minimum variance unbiased estimator for non-linear polynomial functions. Given N snapshots $\hat{\rho}_1, \dots \hat{\rho}_N$ and a polynomial function in these snapshots f such that $\mathbb{E}[f(\hat{\rho}_1, \dots \hat{\rho}_m)] = \mathrm{Tr}[\bar{O}\rho^{\otimes m}]$ (note that we used the abbreviation $\hat{\rho}_i = (\hat{\rho}_{id})_i = \hat{\rho}_{\underline{k}_i, l_i}$), the U-statistics [235] of order m is then given by,

$$U_N \coloneqq \binom{N}{m}^{-1} \sum_{\mathcal{C}_{N,m}} f(\hat{\rho}_{i_1}, ..., \hat{\rho}_{i_m}), \tag{E.22}$$

with $C_{N,m}$ representing the set of all lists of m *distinct* elements one can construct from the N snapshots. Ref. [235] gives us a closed form for the variance of this estimator as,

$$\mathrm{Var}[U_N] = \frac{1}{\binom{N}{m}} \sum_{d=1}^{m} \binom{m}{d} \binom{N-m}{m-d} \mathrm{Var}[f^{(d)}(\hat{\rho}_1, ..., \hat{\rho}_d)], \tag{E.23}$$

with,

$$f^{(d)}(\hat{\rho}_1, ..., \hat{\rho}_d) \coloneqq \mathop{\mathbb{E}}_{\hat{\rho}_{d+1}, \dots, \hat{\rho}_m} [f(\hat{\rho}_1, ..., \hat{\rho}_d, \hat{\rho}_{d+1}, ..., \hat{\rho}_m)], \tag{E.24}$$

and $f(\hat{\rho}_1, ..., \hat{\rho}_m) = \mathrm{Tr}[\bar{O}\hat{\rho}_1, ..., \hat{\rho}_m]$. As we want to bound the variance in Eq. (E.23), we can focus first on bounding a single term $\mathrm{Var}[f^{(d)}(\hat{\rho}_1, ..., \hat{\rho}_d)]$ where,

$$f^{(d)}(\hat{\rho}_1, ..., \hat{\rho}_k) = \mathrm{Tr}[\bar{O}\hat{\rho}_1 \otimes \cdots \otimes \hat{\rho}_d \otimes \rho^{\otimes(m-d)}], \tag{E.25}$$

with the convention $\rho^{\otimes 0} = 1 \in \mathbb{C}$. Let $\hat{\rho}_{\underline{l}} = \hat{\rho}_{l_1} \otimes \cdots \otimes \hat{\rho}_{l_d} \otimes \rho^{\otimes(m-d)}$ with $\underline{l} = (l_1, \ldots, l_d)$. Using the fact that the estimator is unbiased we have,

$$\begin{aligned}
\mathrm{Var}[f^{(d)}(\hat{\rho}_1, ..., \hat{\rho}_d)] &= \mathbb{E}[f^{(d)}(\hat{\rho}_1, ..., \hat{\rho}_d)^2] \\
&= \mathbb{E}[\mathrm{Tr}[\bar{O}\hat{\rho}_1 \otimes \cdots \otimes \hat{\rho}_k \otimes \rho^{\otimes p}]^2] \\
&= \|\gamma\|_1^{2d} \mathop{\mathbb{E}}_{\underline{k}_1, l_1} \cdots \mathop{\mathbb{E}}_{\underline{k}_d, l_d} \mathrm{Tr}[\bar{O}\hat{\rho}_{\underline{l}}]^2 \\
&= \|\gamma\|_1^{2d} \sum_{\underline{k}_1, ..., \underline{k}_d} \sum_{l_1, ..., l_d} \left(\prod_{j=1}^{d} p_{\underline{k}_j} p(l_j | \underline{k}_j) \right) \mathrm{Tr}[\bar{O}\hat{\rho}_{\underline{l}}]^2 \\
&= \|\gamma\|_1^{2d} \sum_{\underline{k}_1, ..., \underline{k}_d} \sum_{l_1, ..., l_d} \left(\prod_{j=1}^{d} \frac{|\gamma_{\underline{k}_j}|}{\|\gamma\|_1} \mathrm{Tr}[\mathcal{G}_{\underline{k}_j}(|0\rangle\langle 0|) E_{l_j})] \right) \mathrm{Tr}[\bar{O}\hat{\rho}_{\underline{l}}]^2 \\
&= \|\gamma\|_1^{2d} \sum_{l_1, ..., l_d} \left(\prod_{j=1}^{d} \mathrm{Tr}\left[\frac{1}{\|\gamma\|_1} \sum_{\underline{k}_j} |\gamma_{\underline{k}_j}| \mathcal{G}_{\underline{k}_j}(|0\rangle\langle 0|) E_{l_j} \right] \right) \mathrm{Tr}[\bar{O}\hat{\rho}_{\underline{l}}]^2.
\end{aligned}$$

As in the proof of Lemma 3 we introduce the operator Ω such that $\forall 1 \leq j \leq d,\ \Omega \coloneqq \|\gamma\|^{-1} \sum_{\underline{k}_j} |\gamma_{\underline{k}_j}| \mathcal{G}_{\underline{k}_j}$ and we note $E_{\underline{l}} = \bigotimes_{l_j} E_{l_j}$. This yields,

$$\begin{aligned}
\mathrm{Var}[f^{(d)}(\hat{\rho}_1, ..., \hat{\rho}_d)] &= \|\gamma\|_1^{2d} \sum_{l_1, ..., l_d} \left(\prod_{j=1}^{d} \mathrm{Tr}[\Omega(|0\rangle\langle 0|) E_{l_j})] \right) \mathrm{Tr}[\bar{O}\hat{\rho}_{\underline{l}}]^2 \\
&= \|\gamma\|_1^{2d} \sum_{\underline{l}} \mathrm{Tr}[\Omega(|0\rangle\langle 0|)^{\otimes d} E_{\underline{l}})] \mathrm{Tr}[\bar{O}\hat{\rho}_{\underline{l}}]^2,
\end{aligned}$$

where used the fact that the trace of a tensor product is equal to the product of the individual traces. Using the similar trick of maximising over all potential states (as in Eq. (E.11)) and introducing $\Theta = \sum_{\underline{l}} \mathrm{Tr}[\bar{O}\hat{\rho}_{\underline{l}}]^2 E_{\underline{l}}$ we get,

$$\begin{aligned}
\mathrm{Var}[f^{(d)}(\hat{\rho}_1, ..., \hat{\rho}_d)] &\leq \|\gamma\|_1^{2d} \max_{\sigma} \mathrm{Tr}\left[\Omega(\sigma)^{\otimes d} \sum_{\underline{l}} \mathrm{Tr}[\bar{O}\hat{\rho}_{\underline{l}}] E_{\underline{l}} \right] \\
&\leq \|\gamma\|_1^{2d} \max_{\sigma} \mathrm{Tr}\left[\Omega(\sigma)^{\otimes d} \Theta \right] \\
&= \|\gamma\|_1^{2d} \|\Theta\|_\infty. \qquad \text{(E.26)}
\end{aligned}$$

We can now upper bound $\mathrm{Var}[U_N]$,

$$\begin{aligned}
\mathrm{Var}[U_N] &\leq \|\gamma\|_1^{2m} \|\Theta\|_\infty \frac{1}{\binom{N}{m}} \sum_{d=1}^{m} \binom{m}{d} \binom{N-m}{m-d} \\
&= \|\gamma\|_1^{2m} \|\Theta\|_\infty \left(1 - \frac{((N-m)!)^2}{N!(N-2m)!} \right) \\
&\leq \|\gamma\|_1^{2m} \|\Theta\|_\infty. \qquad \text{(E.27)}
\end{aligned}$$

where we used $\|\gamma\|_1^{2d} \leq \|\gamma\|_1^{2m}$ (as $\|\gamma\|_1 > 1$) in the first line and the fact that the expression in the second line is bounded by 1. To achieve an accuracy ε in estimating polynomial functions of degree m we thus need a number of samples N_s that scales as $\mathcal{O}(\|\gamma\|_1^{2m}/\varepsilon^2)$. □

E.2 Details on classical post-processing algorithms

E.2.1 Expectation value of local Paulis

Suppose we want to compute the expectation value of a q-local observable $O = \bigotimes_{i\in Q} O^{(i)}$ for a given shadow snapshot $\hat{\rho}_{id}$,

$$\mathrm{Tr}[O\hat{\rho}_{id}] = \mathrm{Tr}[O\hat{\rho}_{\underline{k},l}] = \|\gamma\|_1^2 \mathrm{sign}(\gamma_{\underline{k}})\mathrm{Tr}[OC_E^{-1}(E_l)], \tag{E.28}$$

with Q is the set of qubits on which O acts non-trivially and $l = (j, \underline{b})$. In the case of single-qubit random Pauli basis transformation, the effects can be split too as in Eq. (5.21) and we get,

$$\begin{aligned}\mathrm{Tr}[O\hat{\rho}_{\underline{k},l}] &= \|\gamma\|_1^2 \mathrm{sign}(\gamma_{\underline{k}}) \prod_{i\in Q} \mathrm{Tr}\left[O^{(i)}\left(3(Q_j^{(i)})^\dagger \left|b^{(i)}\right\rangle\!\left\langle b^{(i)}\right| Q_j^{(i)} - \mathbb{1}\right)\right] \\ &= 3^q \|\gamma\|_1^2 \mathrm{sign}(\gamma_{\underline{k}}) f(O, \underline{b}, Q_j)\end{aligned} \tag{E.29}$$

with $f(O, b, Q_j) = \prod_{i\in Q} \mathrm{Tr}\left[O^{(i)}\left((Q_j^{(i)})^\dagger \left|b^{(i)}\right\rangle\!\left\langle b^{(i)}\right| Q_j^{(i)} - \mathbb{1}\right)\right]$ and where we used the fact that when $O^{(i)} = \mathbb{1}$, then $\mathrm{Tr}\left[3(Q_j^{(i)})^\dagger \left|b^{(i)}\right\rangle\!\left\langle b^{(i)}\right| Q_j^{(i)} - \mathbb{1}\right] = 1$. $f(O, b, Q_j)$ in Eq. (E.29) evaluates to ± 1 or 0 as follows,

- if O and Q_j are compatible, meaning that $\forall i \in Q$, $Q_j^{(i)}$ basis rotation corresponds to $O^{(i)}$ (e.g. Q_j is a X basis rotation and $O = X$) then $f(O, b, Q_j) = \pm 1$. If the Hamming weight of $\underline{b}$ is odd then $f(O, \underline{b}, Q_j) = -1$ and $f(O, \underline{b}, Q_j) = +1$ otherwise.
- if O and Q_j are incompatible then $f(O, \underline{b}, Q_j) = 0$.

The algorithm takes as input a PEC shadows H, computes $\mathrm{Tr}[O\hat{\rho}_{\underline{k},l}]$ as in Eq. (E.29) for each snapshot and returns their average. The runtime is $\mathcal{O}(qN_s)$.

E.2.2 Improved algorithm using light cones

For ease of presentation, suppose we have a circuit composed of a single noisy gate which quasiprobability decomposition is given by $|\gamma_1|\mathcal{G}_1 - |\gamma_2|\mathcal{G}_2$. Moreover, suppose that we wish to compute the expectation value of an observable O such that this gate is not contained in its lightcone. As explained in Algorithm 2, we can use Algorithm 1 with the modified quasiprobability $\text{sign}(\gamma_k) = 1$ and modified norm $\|\gamma\|_1 = 1$ to reduce the sample complexity of PEC shadows.

The expectation value we compute is thus equal to,

$$\begin{aligned} |\gamma_1|\text{Tr}[O\mathcal{G}_1(\rho)] + |\gamma_2|\text{Tr}[O\mathcal{G}_2(\rho)] &= |\gamma_1|\text{Tr}[O\mathcal{U}(\rho)] + |\gamma_2|\text{Tr}[O\mathcal{U}(\rho)] \\ &= (|\gamma_1| + |\gamma_2|)\,\text{Tr}[O\mathcal{U}(\rho)] = \text{Tr}[O\mathcal{U}(\rho)] \end{aligned} \tag{E.30}$$

where we used the fact that $\text{Tr}[O\mathcal{G}_1(\rho)] = \text{Tr}[O\mathcal{G}_2(\rho)] = \text{Tr}[O\mathcal{U}(\rho)]$ as the gate is not in the light cone of O, and $|\gamma_1| + |\gamma_2| = 1$.

We can perform the same transformation for each gate in a more complex circuit. Providing that the noise is local, we can define the new set of quasiprobabilities $\tilde{\gamma}_k$ such that

$$\|\tilde{\gamma}\|_1 = \prod_{i\in\mathcal{I}} \left\|\gamma^{(i)}\right\|_1 \quad \text{and} \quad \text{sign}(\tilde{\gamma}_{\underline{k}}) = \prod_{i\in\mathcal{I}} \text{sign}(\gamma^{(i)}_{k_i}). \tag{E.31}$$

with $\mathcal{I}$ being the set of qubits in the light cone of O. In practice, the PEC shadows protocol remains the same in that the circuit variants are generated as if each gate was noisy. One only has to modify the post-processing step by using the new set of quasiprobabilities as in Eq. (E.31) to get a reduction of the sample complexity.

E.2.3 Local purities

Given a collection of shadows $\hat{\rho}_1, \dots \hat{\rho}_{N_s}$ and a set of qubits $Q = \{q_1, \dots, q_{|Q|}\}$, we wish to predict the purity of the ideal reduced density matrix $\text{Tr}[\rho^2_{id,Q}]$. Note that we used the abbreviation $\hat{\rho}_n = (\hat{\rho}_{id})_n = \hat{\rho}_{\underline{k}_n,l_n}$ with $l_n = (j_n, \underline{b}_n)$ for the snapshots. Using the results of Appendix E.1.3 an unbiased estimator of the purity is given by,

$$\begin{aligned} \text{Tr}[\text{SWAP}_{Q,Q'}\hat{\rho}_n \otimes \hat{\rho}_m] &= \|\gamma\|_1^2 \text{sign}(\gamma_{\underline{k}_n})\text{sign}(\gamma_{\underline{k}_m})\text{Tr}[\text{SWAP}_{Q,Q'}\hat{\rho}_{l_n} \otimes \hat{\rho}_{l_m}] \\ &= \|\gamma\|_1^2 \text{sign}(\gamma_{\underline{k}_n})\text{sign}(\gamma_{\underline{k}_m}) f(n,m,Q) \end{aligned} \tag{E.32}$$

with $n \neq m$ and $\text{SWAP}_{Q,Q'}$ is an operation that swaps every pairs of qubits (q_r, q_{r+N}) in the $2N$ qubit state $\hat{\rho}_{l_n} \otimes \hat{\rho}_{l_m}$. Here again we use the fact that the snapshots can be written as single-qubit tensor products when using single-qubit random bases transformation,

$$f(n, m, Q) = \prod_{q_r \in Q} \text{Tr}[\text{SWAP} \hat{\rho}_{l_n}^{(q_n)} \otimes \hat{\rho}_{l_m}^{(q_n)}], \tag{E.33}$$

with SWAP being the standard 2-qubit SWAP operator. As $f(n, m, Q)$ is a product of traces involving only 2 qubits, we can compute it analytically as follows by focusing on its components first,

- if $Q_{j_n}^{(q_r)} \neq Q_{j_m}^{(q_r)}$ then $\text{Tr}[\text{SWAP} \hat{\rho}_{l_n}^{(q_n)} \otimes \hat{\rho}_{l_m}^{(q_n)}] = \frac{1}{2}$.
- if $Q_{j_n}^{(q_r)} = Q_{j_m}^{(q_r)}$ and $b_n^{(q_r)} = b_m^{(q_r)}$ then $\text{Tr}[\text{SWAP} \hat{\rho}_{l_n}^{(q_n)} \otimes \hat{\rho}_{l_m}^{(q_n)}] = 5$
- otherwise $\text{Tr}[\text{SWAP} \hat{\rho}_{l_n}^{(q_n)} \otimes \hat{\rho}_{l_m}^{(q_n)}] = -4$.

$f(n, m, Q)$ is then given by the product of these quantities. The algorithm goes over all distinct pairs of snapshots $\hat{\rho}_n$ and $\hat{\rho}_m$, computes $f(n, m, Q)$ and evaluates the purity as in Eq. (E.32). The final estimate is then given by using the median of means estimator of the previously computed estimates. The runtime of the algorithm is $\mathcal{O}(|Q| N_s^2)$.

F

Appendices for Biased Estimator for Classical Shadows

Contents

F.1 Derivation of the optimal biasing point 173
F.2 SNR when using classical shadows 174
F.3 Critical biasing 175

F.1 Derivation of the optimal biasing point

Let $\hat{o}$ be an unbiased estimator. We introduce the biased estimator $\hat{o}_\alpha$ by rescaling $\hat{o}$ as $\hat{o}_\alpha = (1-\alpha)\hat{o}$. We adopt this formalism as it will simplify the derivations, while noting that we retrieve the rescaling used in the main text by setting $\alpha = 1 - \sqrt{1-\varepsilon}$. Furthermore, we can associate to $\hat{o}$ and $\hat{o}_\alpha$ the sample mean estimators $\bar{o}$ and $\bar{o}_\alpha$ such that,

$$\mathrm{MSE}[\bar{o}] = \frac{\mathrm{Var}[\hat{o}]}{N_s} \tag{F.1}$$

$$\begin{aligned}\mathrm{MSE}[\bar{o}_\alpha] = \mathrm{MSE}[(1-\alpha)\bar{o}] &= \underbrace{\left(\mathbb{E}[(1-\alpha)\hat{o}] - \mathbb{E}[\hat{o}]\right)^2}_{\text{bias}} + \underbrace{\frac{\mathrm{Var}[(1-\alpha)\hat{o}]}{N_s}}_{\text{variance}} \\ &= \alpha^2\mathbb{E}[\hat{o}]^2 + (1-\alpha)^2\frac{\mathrm{Var}[\hat{o}]}{N_s}.\end{aligned} \tag{F.2}$$

In order to find the optimal biasing point α^* which minimises the MSE we take the derivative of Eq. (F.2) with respect to α and set it to zero,

$$2\alpha^*\mathbb{E}[\hat{o}]^2 - 2\alpha^*\frac{\mathrm{Var}[\hat{o}]}{N_s} = 0,$$

which leads to,

$$\alpha^* = \frac{\mathrm{Var}[\hat{o}]/N_s}{\mathbb{E}[\hat{o}] + \mathrm{Var}[\hat{o}]/N_s} = \frac{1}{1+\beta}, \tag{F.3}$$

with $\beta = \frac{\mathbb{E}[\hat{o}]}{\mathrm{Var}[\hat{o}]/N_s}$ the signal to noise ratio. Replacing α^* by $1 - \sqrt{1-\varepsilon^*}$ we find,

$$\begin{aligned}
\varepsilon^* &= 1 - (1-\alpha^*)^2 \\
&= 1 - (1 - \frac{1}{1+\beta})^2 \\
&= 1 - \frac{\beta^2}{(1+\beta)^2} \\
&= \frac{1+2\beta}{(1+\beta)^2}.
\end{aligned} \tag{F.4}$$

F.2 SNR when using classical shadows

In the previous section, we saw that the SNR β is given for general estimators by,

$$\beta = \frac{\mathbb{E}[\hat{o}]}{\mathrm{Var}[\hat{o}]/N_s} \tag{F.5}$$

Here, we will derive it for the estimator $\hat{o}$ obtained from classical shadows when the goal is to estimate expectation values. In this case, recall that $\hat{o} = \mathrm{Tr}[O\hat{\rho}_l]$ such that,

$$\hat{o} = \mathrm{Tr}[O\hat{\rho}_l] = \mathrm{Tr}\left[O\left(\bigotimes_{i=1}^{N}\left[3(Q_j^{(i)})^\dagger \left|b^{(i)}\right\rangle\!\left\langle b^{(i)}\right| Q_j^{(i)} - \mathbb{1}\right]\right)\right]. \tag{F.6}$$

Let $\mathcal{I}$ (with $|\mathcal{I}| = q$) be the set of indexes where O is non-trivial, we then have,

$$\begin{aligned}
\hat{o} &= \prod_{i\in\mathcal{I}} 3\mathrm{Tr}\left[O^{(i)}(Q_j^{(i)})^\dagger \left|b^{(i)}\right\rangle\!\left\langle b^{(i)}\right| Q_j^{(i)} - O^{(i)}\right] \\
&= 3^q \prod_{i\in\mathcal{I}} \mathrm{Tr}\left[O^{(i)}(Q_j^{(i)})^\dagger \left|b^{(i)}\right\rangle\!\left\langle b^{(i)}\right| Q_j^{(i)}\right] \\
&= \begin{cases} +3^q & \text{if } O \text{ and } Q_j \text{ are compatible and the Hamming weight of } \underline{b} \text{ is even} \\ -3^q & \text{if } O \text{ and } Q_j \text{ are compatible and the Hamming weight of } \underline{b} \text{ is odd} \\ 0 & \text{otherwise} \end{cases}
\end{aligned} \tag{F.7}$$

where we used the fact that the trace is equal to one where O acts trivially and $\mathrm{Tr}[O^{(i)}] = 0$ as $O^{(i)}$ is a Pauli matrix. As the probability for O and Q_j to be compatible (as defined in Appendix E.2) is equal to 3^{-q} we obtain,

$$\mathbb{E}[\hat{o}^2] = \frac{1}{3^q} 3^{2q} + (1 - \frac{1}{3^q}) 0^2 = 3^q, \tag{F.8}$$

which leads to,

$$\begin{aligned} \mathrm{Var}[\hat{o}] &= \mathbb{E}[\hat{o}^2] - \mathbb{E}[\hat{o}]^2 \\ &= 3^q - \mathrm{Tr}[O\rho]^2. \end{aligned} \tag{F.9}$$

Finally, we find that the SNR for classical shadows reads,

$$\beta = \frac{\mathbb{E}[\hat{o}]}{\mathrm{Var}[\hat{o}]/N_s} = \frac{\mathrm{Tr}[O\rho]}{(3^q - \mathrm{Tr}[O\rho]^2)/N_s} = N_s \frac{1}{3^q/\mathrm{Tr}[O\rho]^2 - 1}, \tag{F.10}$$

and as when biasing we obtain $\beta_{biased} = 1 + \beta$ the factor of SNR gain is given by,

$$\frac{\beta_{biased}}{\beta} = 1 + \frac{1}{\beta} = 1 + \frac{3^q/\mathrm{Tr}[O\rho]^2 - 1}{N_s}. \tag{F.11}$$

F.3 Critical biasing

In this section we derive the range of values for which biased shadows outperform the conventional ones. Similarly as in the previous section, we use the rescaling factor $(1 - \alpha)$ for ease of presentation. We have,

$$\begin{aligned} \mathrm{MSE}[(1-\alpha)\bar{o}] &\leq \mathrm{MSE}[\bar{o}] \\ \alpha^2 \mathbb{E}[\hat{o}]^2 + (1-\alpha)^2 \frac{\mathrm{Var}[\hat{o}]}{N_s} &\leq \frac{\mathrm{Var}[\hat{o}]}{N_s} \\ \alpha^2 \mathbb{E}[\hat{o}]^2 &\leq \alpha(2-\alpha) \frac{\mathrm{Var}[\hat{o}]}{N_s}. \end{aligned} \tag{F.12}$$

The solution $\alpha = 0$ corresponds to the unbiased case and we will suppose thereafter that $\alpha \neq 0$,

$$\begin{aligned} \alpha \mathbb{E}[\hat{o}]^2 &\leq (2-\alpha)\mathrm{Var}[\hat{o}]/N_s \\ \alpha &\leq \frac{2\mathrm{Var}[\hat{o}]/N_s}{\mathbb{E}[\hat{o}]^2 + \mathrm{Var}[\hat{o}]/N_s} = \alpha_c = 2\alpha^*, \end{aligned} \tag{F.13}$$

which leads to $\varepsilon_c = 1 - (1 - \alpha_c)^2 = 4\beta/(1+\beta)^2$. This could explain why we still obtain a reduction of the MSE when using an approximation of ε in Section 7.4.

Bibliography

[1] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information*, Vol. 10th Anniversary Edition (Cambridge University Press, 2010).

[2] H. J. Briegel, D. E. Browne, W. Dür, R. Raussendorf, and M. Van den Nest, "Measurement-based quantum computation", Nature Physics **5**, 19 (2009).

[3] D. Aharonov, W. van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev, *Adiabatic quantum computation is equivalent to standard quantum computation*, 2005.

[4] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, and J. M. Martinis, "Quantum supremacy using a programmable superconducting processor", Nature **574**, 505 (2019).

[5] H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu, P. Hu, X.-Y. Yang, W.-J. Zhang, H. Li, Y. Li, X. Jiang, L. Gan, G. Yang, L. You, Z. Wang, L. Li, N.-L. Liu, C.-Y. Lu, and J.-W. Pan, "Quantum computational advantage using photons", en, Science **370**, 1460 (2020).

[6] Y. Wu, W.-S. Bao, S. Cao, F. Chen, M.-C. Chen, X. Chen, T.-H. Chung, H. Deng, Y. Du, D. Fan, M. Gong, C. Guo, C. Guo, S. Guo, L. Han, L. Hong, H.-L. Huang, Y.-H. Huo, L. Li, N. Li, S. Li, Y. Li, F. Liang, C. Lin, J. Lin, H. Qian, D. Qiao, H. Rong, H. Su, L. Sun, L. Wang, S. Wang, D. Wu, Y. Xu, K. Yan, W. Yang, Y. Yang, Y. Ye, J. Yin, C. Ying, J. Yu, C. Zha, C. Zhang, H. Zhang, K. Zhang, Y. Zhang, H. Zhao, Y. Zhao, L. Zhou, Q. Zhu, C.-Y. Lu, C.-Z. Peng, X. Zhu, and J.-W. Pan, "Strong Quantum Computational Advantage Using a Superconducting Quantum Processor", Physical Review Letters **127**, 180501 (2021).

[7] H.-S. Zhong, Y.-H. Deng, J. Qin, H. Wang, M.-C. Chen, L.-C. Peng, Y.-H. Luo, D. Wu, S.-Q. Gong, H. Su, Y. Hu, P. Hu, X.-Y. Yang, W.-J. Zhang, H. Li, Y. Li, X. Jiang, L. Gan, G. Yang, L. You, Z. Wang, L. Li, N.-L. Liu, J. J. Renema, C.-Y. Lu, and J.-W. Pan, "Phase-Programmable Gaussian Boson Sampling Using Stimulated Squeezed Light", Physical Review Letters **127**, 180502 (2021).

[8] D. Bluvstein, S. J. Evered, A. A. Geim, S. H. Li, H. Zhou, T. Manovitz, S. Ebadi, M. Cain, M. Kalinowski, D. Hangleiter, J. P. B. Ataides, N. Maskara, I. Cong, X. Gao, P. S. Rodriguez, T. Karolyshyn, G. Semeghini, M. J. Gullans, M. Greiner, V. Vuletić, and M. D. Lukin, "Logical quantum processor based on reconfigurable atom arrays", Nature, 1 (2023).

[9] Y. Kim, A. Eddins, S. Anand, K. X. Wei, E. van den Berg, S. Rosenblatt, H. Nayfeh, Y. Wu, M. Zaletel, K. Temme, and A. Kandala, "Evidence for the utility of quantum computing before fault tolerance", Nature **618**, 500 (2023).

[10] R. Acharya, I. Aleiner, R. Allen, T. I. Andersen, M. Ansmann, F. Arute, K. Arya, A. Asfaw, J. Atalaya, R. Babbush, D. Bacon, J. C. Bardin, J. Basso, A. Bengtsson, S. Boixo, G. Bortoli, A. Bourassa, J. Bovaird, L. Brill, M. Broughton, B. B. Buckley, D. A. Buell, T. Burger, B. Burkett, N. Bushnell, Y. Chen, Z. Chen, B. Chiaro, J. Cogan, R. Collins, P. Conner, W. Courtney, A. L. Crook, B. Curtin, D. M. Debroy, A. Del Toro Barba, S. Demura, A. Dunsworth, D. Eppens, C. Erickson, L. Faoro, E. Farhi, R. Fatemi, L. Flores Burgos, E. Forati, A. G. Fowler, B. Foxen, W. Giang, C. Gidney, D. Gilboa, M. Giustina, A. Grajales Dau, J. A. Gross, S. Habegger, M. C. Hamilton, M. P. Harrigan, S. D. Harrington, O. Higgott, J. Hilton, M. Hoffmann, S. Hong, T. Huang, A. Huff, W. J. Huggins, L. B. Ioffe, S. V. Isakov, J. Iveland, E. Jeffrey, Z. Jiang, C. Jones, P. Juhas, D. Kafri, K. Kechedzhi, J. Kelly, T. Khattar, M. Khezri, M. Kieferová, S. Kim, A. Kitaev, P. V. Klimov, A. R. Klots, A. N. Korotkov, F. Kostritsa, J. M. Kreikebaum, D. Landhuis, P. Laptev, K.-M. Lau, L. Laws, J. Lee, K. Lee, B. J. Lester, A. Lill, W. Liu, A. Locharla, E. Lucero, F. D. Malone, J. Marshall, O. Martin, J. R. McClean, T. McCourt, M. McEwen, A. Megrant, B. Meurer Costa, X. Mi, K. C. Miao, M. Mohseni, S. Montazeri, A. Morvan, E. Mount, W. Mruczkiewicz, O. Naaman, M. Neeley, C. Neill, A. Nersisyan, H. Neven, M. Newman, J. H. Ng, A. Nguyen, M. Nguyen, M. Y. Niu, T. E. O'Brien, A. Opremcak, J. Platt, A. Petukhov, R. Potter, L. P. Pryadko, C. Quintana, P. Roushan, N. C. Rubin, N. Saei, D. Sank, K. Sankaragomathi, K. J. Satzinger, H. F. Schurkus, C. Schuster, M. J. Shearn, A. Shorter, V. Shvarts, J. Skruzny, V. Smelyanskiy, W. C. Smith, G. Sterling, D. Strain, M. Szalay, A. Torres, G. Vidal, B. Villalonga, C. Vollgraff Heidweiller, T. White, C. Xing, Z. J. Yao, P. Yeh, J. Yoo, G. Young, A. Zalcman, Y. Zhang, N. Zhu, and Google Quantum AI, "Suppressing quantum errors by scaling a surface code logical qubit", Nature **614**, 676 (2023).

[11] D. P. DiVincenzo, "The Physical Implementation of Quantum Computation", en, Fortschritte der Physik **48**, 771 (2000).

[12] M. Kjaergaard, M. E. Schwartz, J. Braumüller, P. Krantz, J. I.-J. Wang, S. Gustavsson, and W. D. Oliver, "Superconducting Qubits: Current State of Play", en, Annual Review of Condensed Matter Physics **11**, 369 (2020).

[13] H.-L. Huang, D. Wu, D. Fan, and X. Zhu, "Superconducting Quantum Computing: A Review", arXiv:2006.10433 [quant-ph], arXiv: 2006.10433 (2020).

[14] C. D. Bruzewicz, J. Chiaverini, R. McConnell, and J. M. Sage, "Trapped-ion quantum computing: Progress and challenges", Applied Physics Reviews **6**, 021314 (2019).

[15] K. Wintersperger, F. Dommert, T. Ehmer, A. Hoursanov, J. Klepsch, W. Mauerer, G. Reuber, T. Strohm, M. Yin, and S. Luber, "Neutral atom quantum computing hardware: performance and end-user perspective", EPJ Quantum Technology **10**, 1 (2023).

[16] N. Maring, A. Fyrillas, M. Pont, E. Ivanov, P. Stepanov, N. Margaria, W. Hease, A. Pishchagin, A. Lemaître, I. Sagnes, T. H. Au, S. Boissier, E. Bertasi, A. Baert, M. Valdivia, M. Billard, O. Acar, A. Brieussel, R. Mezher, S. C. Wein, A. Salavrakos, P. Sinnott, D. A. Fioretto, P.-E. Emeriau, N. Belabas, S. Mansfield, P. Senellart, J. Senellart, and N. Somaschi, "A versatile single-photon-based quantum computing platform", Nature Photonics **18**, 603 (2024).

[17] J. Romero and G. Milburn, *Photonic quantum computing*, 2024.

[18] A. Chatterjee, P. Stevenson, S. De Franceschi, A. Morello, N. P. de Leon, and F. Kuemmeth, "Semiconductor qubits in practice", en, Nature Reviews Physics **3**, 157 (2021).

[19] G. Burkard, T. D. Ladd, A. Pan, J. M. Nichol, and J. R. Petta, "Semiconductor spin qubits", Rev. Mod. Phys. **95**, 025003 (2023).

[20] M. P. da Silva, C. Ryan-Anderson, J. M. Bello-Rivas, A. Chernoguzov, J. M. Dreiling, C. Foltz, F. Frachon, J. P. Gaebler, T. M. Gatterman, L. Grans-Samuelsson, D. Hayes, N. Hewitt, J. Johansen, D. Lucchetti, M. Mills, S. A. Moses, B. Neyenhuis, A. Paz, J. Pino, P. Siegfried, J. Strabley, A. Sundaram, D. Tom, S. J. Wernli, M. Zanner, R. P. Stutz, and K. M. Svore, *Demonstration of logical qubits and repeated error correction with better-than-physical error rates*, 2024.

[21] D. Loss and D. P. DiVincenzo, "Quantum computation with quantum dots", Phys. Rev. A **57**, 120 (1998).

[22] B. E. Kane, "A silicon-based nuclear spin quantum computer", en, Nature **393**, 133 (1998).

[23] M. D. Michielis, E. Ferraro, E. Prati, L. Hutin, B. Bertrand, E. Charbon, D. J. Ibberson, and M. F. Gonzalez-Zalba, "Silicon spin qubits from laboratory to industry", Journal of Physics D: Applied Physics **56**, 363001 (2023).

[24] A. M. J. Zwerver, T. Krähenmann, T. F. Watson, L. Lampert, H. C. George, R. Pillarisetty, S. A. Bojarski, P. Amin, S. V. Amitonov, J. M. Boter, R. Caudillo, D. Correas-Serrano, J. P. Dehollain, G. Droulers, E. M. Henry, R. Kotlyar, M. Lodari, F. Lüthi, D. J. Michalak, B. K. Mueller, S. Neyens, J. Roberts, N. Samkharadze, G. Zheng, O. K. Zietz, G. Scappucci, M. Veldhorst, L. M. K. Vandersypen, and J. S. Clarke, "Qubits made by advanced semiconductor manufacturing", Nature Electronics **5**, 184 (2022).

[25] M. Veldhorst, H. G. J. Eenink, C. H. Yang, and A. S. Dzurak, "Silicon CMOS architecture for a spin-based quantum computer", Nature Communications **8**, 1766 (2017).

[26] R. Maurand, X. Jehl, D. Kotekar-Patil, A. Corna, H. Bohuslavskyi, R. Laviéville, L. Hutin, S. Barraud, M. Vinet, M. Sanquer, and S. De Franceschi, "A CMOS silicon spin qubit", Nature Communications **7**, 13575 (2016).

[27] L. Hutin, "From Silicon CMOS Technology to Spin Qubits", in 2019 IEEE 11th International Memory Workshop (IMW), ISSN: 2573-7503 (May 2019), pp. 1–72.

[28] M. F. Gonzalez-Zalba, S. de Franceschi, E. Charbon, T. Meunier, M. Vinet, and A. S. Dzurak, "Scaling silicon-based quantum computing using CMOS technology", Nature Electronics **4**, 872 (2021).

[29] R. Hanson, L. P. Kouwenhoven, J. R. Petta, S. Tarucha, and L. M. K. Vandersypen, "Spins in few-electron quantum dots", Reviews of Modern Physics **79**, 1217 (2007).

[30] D. P. DiVincenzo, D. Bacon, J. Kempe, G. Burkard, and K. B. Whaley, "Universal quantum computation with the exchange interaction", Nature **408**, 339 (2000).

[31] J. Levy, "Universal Quantum Computation with Spin-$1/2$ Pairs and Heisenberg Exchange", Physical Review Letters **89**, 147902 (2002).

[32] M. Russ and G. Burkard, "Three-electron spin qubits", Journal of Physics: Condensed Matter **29**, 393001 (2017).

[33] C. Boudreault, H. Eleuch, M. Hilke, and R. MacKenzie, "Decoherence-free universal quantum computation with symmetric qubit clusters", arXiv:2106.00754 [cond-mat, physics:quant-ph], arXiv: 2106.00754 (2021).

[34] F. A. Zwanenburg, A. S. Dzurak, A. Morello, M. Y. Simmons, L. C. L. Hollenberg, G. Klimeck, S. Rogge, S. N. Coppersmith, and M. A. Eriksson, "Silicon quantum electronics", Reviews of Modern Physics **85**, 961 (2013).

[35] M. Pioro-Ladrière, T. Obata, Y. Tokura, Y.-S. Shin, T. Kubo, K. Yoshida, T. Taniyama, and S. Tarucha, "Electrically driven single-electron spin resonance in a slanting Zeeman field", en, Nature Physics **4**, 776 (2008).

[36] J. Yoneda, T. Otsuka, T. Nakajima, T. Takakura, T. Obata, M. Pioro-Ladrière, H. Lu, C. J. Palmstrøm, A. C. Gossard, and S. Tarucha, "Fast electrical control of single electron spins in quantum dots with vanishing influence from nuclear spins", Phys. Rev. Lett. **113**, 267601 (2014).

[37] R. Li, L. Petit, D. P. Franke, J. P. Dehollain, J. Helsen, M. Steudtner, N. K. Thomas, Z. R. Yoscovits, K. J. Singh, S. Wehner, L. M. K. Vandersypen, J. S. Clarke, and M. Veldhorst, "A crossbar network for silicon quantum dot qubits", Science Advances **4**, eaar3960 (2018).

[38] T. Struck, A. Hollmann, F. Schauer, O. Fedorets, A. Schmidbauer, K. Sawano, H. Riemann, N. V. Abrosimov, Ł. Cywiński, D. Bougeard, and L. R. Schreiber, "Low-frequency spin qubit energy splitting noise in highly purified 28Si/SiGe", npj Quantum Information **6**, 1 (2020).

[39] E. Vahapoglu, J. P. Slack-Smith, R. C. C. Leon, W. H. Lim, F. E. Hudson, T. Day, J. D. Cifuentes, T. Tanttu, C. H. Yang, A. Saraiva, N. V. Abrosimov, H.-J. Pohl, M. L. W. Thewalt, A. Laucht, A. S. Dzurak, and J. J. Pla, "Coherent control of electron spin qubits in silicon using a global field", npj Quantum Information **8**, 1 (2022).

[40] D. C. McKay, C. J. Wood, S. Sheldon, J. M. Chow, and J. M. Gambetta, "Efficient Z gates for quantum computing", Phys. Rev. A **96**, 022330 (2017).

[41] J. van Dijk, E. Kawakami, R. Schouten, M. Veldhorst, L. Vandersypen, M. Babaie, E. Charbon, and F. Sebastiano, "Impact of classical control electronics on qubit fidelity", Phys. Rev. Appl. **12**, 044054 (2019).

[42] J. Yoneda, K. Takeda, T. Otsuka, T. Nakajima, M. R. Delbecq, G. Allison, T. Honda, T. Kodera, S. Oda, Y. Hoshi, N. Usami, K. M. Itoh, and S. Tarucha, "A quantum-dot spin qubit with coherence limited by charge noise and fidelity higher than 99.9%", Nature Nanotechnology **13**, 102 (2018).

[43] C. H. Yang, K. W. Chan, R. Harper, W. Huang, T. Evans, J. C. C. Hwang, B. Hensen, A. Laucht, T. Tanttu, F. E. Hudson, S. T. Flammia, K. M. Itoh, A. Morello, S. D. Bartlett, and A. S. Dzurak, "Silicon qubit fidelities approaching incoherent noise limits via pulse engineering", Nature Electronics **2**, 151 (2019).

[44] X. Xue, M. Russ, N. Samkharadze, B. Undseth, A. Sammak, G. Scappucci, and L. M. K. Vandersypen, "Quantum logic with spin qubits crossing the surface code threshold", Nature **601**, 343 (2022).

[45] A. R. Mills, C. R. Guinn, M. J. Gullans, A. J. Sigillito, M. M. Feldman, E. Nielsen, and J. R. Petta, "Two-qubit silicon quantum processor with operation fidelity exceeding 99%", Science Advances **8**, eabn5130 (2022).

[46] A. Noiri, K. Takeda, T. Nakajima, T. Kobayashi, A. Sammak, G. Scappucci, and S. Tarucha, "Fast universal quantum gate above the fault-tolerance threshold in silicon", Nature **601**, 338 (2022).

[47] T. Meunier, V. E. Calado, and L. M. K. Vandersypen, "Efficient controlled-phase gate for single-spin qubits in quantum dots", Phys. Rev. B **83**, 121403 (2011).

[48] M. Veldhorst, C. H. Yang, J. C. C. Hwang, W. Huang, J. P. Dehollain, J. T. Muhonen, S. Simmons, A. Laucht, F. E. Hudson, K. M. Itoh, A. Morello, and A. S. Dzurak, "A two-qubit logic gate in silicon", en, Nature **526**, 410 (2015).

[49] J. R. Petta, A. C. Johnson, J. M. Taylor, E. A. Laird, A. Yacoby, M. D. Lukin, C. M. Marcus, M. P. Hanson, and A. C. Gossard, "Coherent Manipulation of Coupled Electron Spins in Semiconductor Quantum Dots", Science **309**, 2180 (2005).

[50] G. Burkard, D. Loss, D. P. DiVincenzo, and J. A. Smolin, "Physical optimization of quantum error correction circuits", Phys. Rev. B **60**, 11404 (1999).

[51] G. Burkard, D. Loss, and D. P. DiVincenzo, "Coupled quantum dots as quantum gates", Phys. Rev. B **59**, 2070 (1999).

[52] T. F. Watson, S. G. J. Philips, E. Kawakami, D. R. Ward, P. Scarlino, M. Veldhorst, D. E. Savage, M. G. Lagally, M. Friesen, S. N. Coppersmith, M. A. Eriksson, and L. M. K. Vandersypen, "A programmable two-qubit quantum processor in silicon", en, Nature **555**, 633 (2018).

[53] D. M. Zajac, A. J. Sigillito, M. Russ, F. Borjans, J. M. Taylor, G. Burkard, and J. R. Petta, "Resonantly driven CNOT gate for electron spins", en, Science **359**, 439 (2018).

[54] M. Russ, D. M. Zajac, A. J. Sigillito, F. Borjans, J. M. Taylor, J. R. Petta, and G. Burkard, "High-fidelity quantum gates in si/sige double quantum dots", Phys. Rev. B **97**, 085421 (2018).

[55] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter, "Elementary gates for quantum computation", Phys. Rev. A **52**, 3457 (1995).

[56] V. N. Ciriano-Tejel, M. A. Fogarty, S. Schaal, L. Hutin, B. Bertrand, L. Ibberson, M. F. Gonzalez-Zalba, J. Li, Y.-M. Niquet, M. Vinet, and J. J. Morton, "Spin readout of a cmos quantum dot by gate reflectometry and spin-dependent tunneling", PRX Quantum **2**, 010353 (2021).

[57] C. H. Yang, A. Rossi, R. Ruskov, N. S. Lai, F. A. Mohiyaddin, S. Lee, C. Tahan, G. Klimeck, A. Morello, and A. S. Dzurak, "Spin-valley lifetimes in a silicon quantum dot with tunable valley splitting", Nature Communications **4**, 2069 (2013).

[58] A. Hollmann, T. Struck, V. Langrock, A. Schmidbauer, F. Schauer, T. Leonhardt, K. Sawano, H. Riemann, N. V. Abrosimov, D. Bougeard, and L. R. Schreiber, "Large, tunable valley splitting and single-spin relaxation mechanisms in a Si/Si_xGe_{1-x} quantum dot", Phys. Rev. Appl. **13**, 034068 (2020).

[59] W. A. Coish and D. Loss, "Hyperfine interaction in a quantum dot: non-markovian electron spin dynamics", Phys. Rev. B **70**, 195340 (2004).

[60] T. Bonsen, P. Harvey-Collard, M. Russ, J. Dijkema, A. Sammak, G. Scappucci, and L. M. K. Vandersypen, "Probing the jaynes-cummings ladder with spin circuit quantum electrodynamics", Phys. Rev. Lett. **130**, 137001 (2023).

[61] P. Harvey-Collard, J. Dijkema, G. Zheng, A. Sammak, G. Scappucci, and L. M. K. Vandersypen, "Coherent Spin-Spin Coupling Mediated by Virtual Microwave Photons", Physical Review X **12**, 021026 (2022).

[62] N. Samkharadze, G. Zheng, N. Kalhor, D. Brousse, A. Sammak, U. C. Mendes, A. Blais, G. Scappucci, and L. M. K. Vandersypen, "Strong spin-photon coupling in silicon", Science **359**, 1123 (2018).

[63] M. Veldhorst, J. C. C. Hwang, C. H. Yang, A. W. Leenstra, B. de Ronde, J. P. Dehollain, J. T. Muhonen, F. E. Hudson, K. M. Itoh, A. Morello, and A. S. Dzurak, "An addressable quantum dot qubit with fault-tolerant control-fidelity", en, Nature Nanotechnology **9**, 981 (2014).

[64] I. Hansen, A. E. Seedhouse, K. W. Chan, F. E. Hudson, K. M. Itoh, A. Laucht, A. Saraiva, C. H. Yang, and A. S. Dzurak, "Implementation of an advanced dressing protocol for global qubit control in silicon", Applied Physics Reviews **9**, 031409 (2022).

[65] L. Viola and S. Lloyd, "Dynamical suppression of decoherence in two-state quantum systems", Phys. Rev. A **58**, 2733 (1998).

[66] G. T. Genov, D. Schraft, N. V. Vitanov, and T. Halfmann, "Arbitrarily accurate pulse sequences for robust dynamical decoupling", Phys. Rev. Lett. **118**, 133202 (2017).

[67] K. W. Chan, W. Huang, C. H. Yang, J. C. C. Hwang, B. Hensen, T. Tanttu, F. E. Hudson, K. M. Itoh, A. Laucht, A. Morello, and A. S. Dzurak, "Assessment of a silicon quantum dot spin qubit environment via noise spectroscopy", Physical Review Applied **10**, `10.1103/physrevapplied.10.044017` (2018).

[68] J. Kerckhoff, B. Sun, B. Fong, C. Jones, A. Kiselev, D. Barnes, R. Noah, E. Acuna, M. Akmal, S. Ha, J. Wright, B. Thomas, C. Jackson, L. Edge, K. Eng, R. Ross, and T. Ladd, "Magnetic gradient fluctuations from quadrupolar ^{73}Ge in Si/SiGe exchange-only qubits", PRX Quantum **2**, 010347 (2021).

[69] P. Stano and D. Loss, "Review of performance metrics of spin qubits in gated semiconducting nanostructures", Nature Reviews Physics **4**, 672 (2022).

[70] D. J. Reilly, C. M. Marcus, M. P. Hanson, and A. C. Gossard, "Fast single-charge sensing with a rf quantum point contact", Applied Physics Letters **91**, 162101 (2007).

[71] C. Barthel, D. J. Reilly, C. M. Marcus, M. P. Hanson, and A. C. Gossard, "Rapid single-shot measurement of a singlet-triplet qubit", Phys. Rev. Lett. **103**, 160503 (2009).

[72] G. A. Oakes, V. N. Ciriano-Tejel, D. F. Wise, M. A. Fogarty, T. Lundberg, C. Lainé, S. Schaal, F. Martins, D. J. Ibberson, L. Hutin, B. Bertrand, N. Stelmashenko, J. W. A. Robinson, L. Ibberson, A. Hashim, I. Siddiqi, A. Lee, M. Vinet, C. G. Smith, J. J. L. Morton, and M. F. Gonzalez-Zalba, "Fast high-fidelity single-shot readout of spins in silicon using a single-electron box", Phys. Rev. X **13**, 011023 (2023).

[73] M. F. Gonzalez-Zalba, S. Barraud, A. J. Ferguson, and A. C. Betz, "Probing the limits of gate-based charge sensing", en, Nature Communications **6**, 6084 (2015).

[74] J. M. Elzerman, R. Hanson, L. H. Willems van Beveren, B. Witkamp, L. M. K. Vandersypen, and L. P. Kouwenhoven, "Single-shot read-out of an individual electron spin in a quantum dot", en, Nature **430**, 431 (2004).

[75] A. E. Seedhouse, T. Tanttu, R. C. Leon, R. Zhao, K. Y. Tan, B. Hensen, F. E. Hudson, K. M. Itoh, J. Yoneda, C. H. Yang, A. Morello, A. Laucht, S. N. Coppersmith, A. Saraiva, and A. S. Dzurak, "Pauli blockade in silicon quantum dots with spin-orbit control", PRX Quantum **2**, 010303 (2021).

[76] H. Jnane and S. C. Benjamin, *Ab initio modelling of quantum dot qubits: coupling, gate dynamics and robustness versus charge noise*, 2024.

[77] A. L. Saraiva, M. J. Calderón, and B. Koiller, "Reliability of the heitler-london approach for the exchange coupling between electrons in semiconductor nanostructures", Phys. Rev. B **76**, 233302 (2007).

[78] C. R. Anderson, M. F. Gyure, S. Quinn, A. Pan, R. S. Ross, and A. A. Kiselev, "High-precision real-space simulation of electrostatically confined few-electron states", AIP Advances **12**, 065123 (2022).

[79] H. E. Ercan, S. N. Coppersmith, and M. Friesen, "Strong electron-electron interactions in si/sige quantum dots", Phys. Rev. B **104**, 235302 (2021).

[80] M. J. Calderón, B. Koiller, and S. Das Sarma, "Exchange coupling in semiconductor nanostructures: validity and limitations of the heitler-london approach", Phys. Rev. B **74**, 045310 (2006).

[81] M. D. Reed, B. M. Maune, R. W. Andrews, M. G. Borselli, K. Eng, M. P. Jura, A. A. Kiselev, T. D. Ladd, S. T. Merkel, I. Milosavljevic, E. J. Pritchett, M. T. Rakher, R. S. Ross, A. E. Schmitz, A. Smith, J. A. Wright, M. F. Gyure, and A. T. Hunter, "Reduced sensitivity to charge noise in semiconductor spin qubits via symmetric operation", Phys. Rev. Lett. **116**, 110402 (2016).

[82] D. Culcer, X. Hu, and S. Das Sarma, "Dephasing of Si spin qubits due to charge noise", Applied Physics Letters **95**, 073102 (2009).

[83] E. Nielsen, R. W. Young, R. P. Muller, and M. S. Carroll, "Implications of simultaneous requirements for low-noise exchange gates in double quantum dots", Phys. Rev. B **82**, 075319 (2010).

[84] J. D. Cifuentes, P. Y. Mai, F. Schlattner, H. E. Ercan, M. Feng, C. C. Escott, A. S. Dzurak, and A. Saraiva, "Path integral simulation of exchange interactions in cmos spin qubits", arXiv (2023).

[85] B. Buonacorsi, M. Korkusinski, B. Khromets, and J. Baugh, "Optimizing lateral quantum dot geometries for reduced exchange noise", arXiv (2020).

[86] M. M. E. K. Shehata, G. Simion, R. Li, F. A. Mohiyaddin, D. Wan, M. Mongillo, B. Govoreanu, I. Radu, K. D. Greve, and P. V. Dorpe, "Modelling semiconductor spin qubits and their charge noise environment for quantum gate fidelity estimation", arXiv (2022).

[87] A. Elsayed, M. Shehata, C. Godfrin, S. Kubicek, S. Massar, Y. Canvel, J. Jussot, G. Simion, M. Mongillo, D. Wan, B. Govoreanu, I. P. Radu, R. Li, P. V. Dorpe, and K. D. Greve, *Low charge noise quantum dots with industrial cmos manufacturing*, 2022.

[88] M. H. Levitt, "Composite pulses", Progress in nuclear magnetic resonance spectroscopy **18**, 61 (1986).

[89] K. R. Brown, A. W. Harrow, and I. L. Chuang, "Arbitrarily accurate composite pulse sequences", Phys. Rev. A **70**, 052318 (2004).

[90] J. A. Jones, *Composite pulses in nmr quantum computation*, 2009.

[91] J. Jones, "Systematic errors and nmr composite pulses (teaching notes, oxford physics, online)",

[92] J. C. Light and T. C. Jr., "Discrete-variable representations and their utilization", Adv. Chem. Phys. **114**, 263 (2000).

[93] M. F. J. A. Fleck Jr. J. Morris, "Time-dependent propagation of high energy laser beams through the atmosphere", Appl. Phys. **10**, 129 (1976).

[94] R. Babbush, N. Wiebe, J. McClean, J. McClain, H. Neven, and G. K.-L. Chan, "Low-depth quantum simulation of materials", Phys. Rev. X **8**, 011044 (2018).

[95] H. H. S. Chan, R. Meister, T. Jones, D. P. Tew, and S. C. Benjamin, "Grid-based methods for chemistry simulations on a quantum computer", Science Advances **9**, eabo7484 (2023).

[96] A. Saraiva, W. H. Lim, C. H. Yang, C. C. Escott, A. Laucht, and A. S. Dzurak, "Materials for silicon quantum dots and their impact on electron spin qubits", Advanced Functional Materials **32**, 2105488 (2022).

[97] L. Lehtovaara, J. Toivanen, and J. Eloranta, "Solution of time-independent schrödinger equation by the imaginary time propagation method", Journal of Computational Physics **221**, 148 (2007).

[98] E. J. Connors, J. Nelson, L. F. Edge, and J. M. Nichol, "Charge-noise spectroscopy of Si/SiGe quantum dots via dynamically-decoupled exchange oscillations", Nature Communications **13**, 940 (2022).

[99] B. Paquelet Wuetz, D. Degli Esposti, A.-M. J. Zwerver, S. V. Amitonov, M. Botifoll, J. Arbiol, A. Sammak, L. M. K. Vandersypen, M. Russ, and G. Scappucci, "Reducing charge noise in quantum dots by using thin silicon quantum wells", Nature Communications **14**, 1385 (2023).

[100] J. Rojas-Arias, A. Noiri, P. Stano, T. Nakajima, J. Yoneda, K. Takeda, T. Kobayashi, A. Sammak, G. Scappucci, D. Loss, and S. Tarucha, "Spatial noise correlations beyond nearest neighbors in ^{28}Si/si-ge spin qubits", Phys. Rev. Appl. **20**, 054024 (2023).

[101] M. Kpa, N. Focke, Ł. Cywiński, and J. A. Krzywda, "Simulation of 1/f charge noise affecting a quantum dot in a Si/SiGe structure", Applied Physics Letters **123**, 034005 (2023).

[102] D. L. Mickelsen, H. M. Carruzzo, and C. C. Yu, "Interacting two-level systems as a source of $1/f$ charge noise in quantum dot qubits", Phys. Rev. B **108**, 195307 (2023).

[103] L. H. Pedersen, N. M. Møller, and K. Mølmer, "Fidelity of quantum operations", Physics Letters A **367**, 47 (2007).

[104] H. Jnane, A. Siegel, and F. Gonzalez-Zalba, *Harnessing electron motion for global spin qubit control*, 2025.

[105] S. G. J. Philips, M. T. Mądzik, S. V. Amitonov, S. L. De Snoo, M. Russ, N. Kalhor, C. Volk, W. I. L. Lawrie, D. Brousse, L. Tryputen, B. P. Wuetz, A. Sammak, M. Veldhorst, G. Scappucci, and L. M. K. Vandersypen, "Universal control of a six-qubit quantum processor in silicon", Nature **609**, 919 (2022).

[106] J. D. Cifuentes, T. Tanttu, W. Gilbert, J. Y. Huang, E. Vahapoglu, R. C. C. Leon, S. Serrano, D. Otter, D. Dunmore, P. Y. Mai, F. Schlattner, M. Feng, K. Itoh, N. Abrosimov, H.-J. Pohl, M. Thewalt, A. Laucht, C. H. Yang, C. C. Escott, W. H. Lim, F. E. Hudson, R. Rahman, A. S. Dzurak, and A. Saraiva, *Bounds to electron spin qubit variability for scalable cmos architectures*, 2023.

[107] I. Hansen, A. E. Seedhouse, S. Serrano, A. Nickl, M. Feng, J. Y. Huang, T. Tanttu, N. Dumoulin Stuyck, W. H. Lim, F. E. Hudson, K. M. Itoh, A. Saraiva, A. Laucht, A. S. Dzurak, and C. H. Yang, "Entangling gates on degenerate spin qubits dressed by a global field", en, Nature Communications **15**, 7656 (2024).

[108] S. M. Patomäki, M. F. Gonzalez-Zalba, M. A. Fogarty, Z. Cai, S. C. Benjamin, and J. J. L. Morton, *Pipeline quantum processor architecture for silicon spin qubits*, 2023.

[109] A. Fayyaz and J. P. Kestner, "Enhanced local addressability of a spin array with local exchange pulses and global microwave driving", Phys. Rev. B **108**, L081303 (2023).

[110] J. F. Chittock-Wood, R. C. C. Leon, M. A. Fogarty, T. Murphy, S. M. Patomäki, G. A. Oakes, F.-E. von Horstig, N. Johnson, J. Jussot, S. Kubicek, B. Govoreanu, D. F. Wise, M. F. Gonzalez-Zalba, and J. J. L. Morton, *Exchange control in a mos double quantum dot made using a 300 mm wafer process*, 2024.

[111] R. Ferdous, K. W. Chan, M. Veldhorst, J. C. C. Hwang, C. H. Yang, H. Sahasrabudhe, G. Klimeck, A. Morello, A. S. Dzurak, and R. Rahman, "Interface-induced spin-orbit interaction in silicon quantum dots and prospects for scalability", Phys. Rev. B **97**, 241401 (2018).

[112] M. A. Fogarty, *Silicon edge-dot architecture for quantum computing with global control and integrated trimming*, 2022.

[113] J. C. C. Hwang, C. H. Yang, M. Veldhorst, N. Hendrickx, M. A. Fogarty, W. Huang, F. E. Hudson, A. Morello, and A. S. Dzurak, "Impact of g-factors and valleys on spin qubits in a silicon double quantum dot", Phys. Rev. B **96**, 045302 (2017).

[114] A. Siegel, A. Strikis, and M. Fogarty, *Towards early fault tolerance on a $2\times n$ array of qubits equipped with shuttling*, 2024.

[115] J. Yoneda, W. Huang, M. Feng, C. H. Yang, K. W. Chan, T. Tanttu, W. Gilbert, R. C. C. Leon, F. E. Hudson, K. M. Itoh, A. Morello, S. D. Bartlett, A. Laucht, A. Saraiva, and A. S. Dzurak, "Coherent spin qubit transport in silicon", Nature Communications **12**, 4114 (2021).

[116] R. Xue, M. Beer, I. Seidler, S. Humpohl, J.-S. Tu, S. Trellenkamp, T. Struck, H. Bluhm, and L. R. Schreiber, *Si/sige qubus for single electron information-processing devices with memory and micron-scale connectivity function*, 2023.

[117] I. Seidler, T. Struck, R. Xue, N. Focke, S. Trellenkamp, H. Bluhm, and L. R. Schreiber, "Conveyor-mode single-electron shuttling in Si/SiGe for a scalable quantum computing architecture", npj Quantum Information **8**, 1 (2022).

[118] V. Langrock, J. A. Krzywda, N. Focke, I. Seidler, L. R. Schreiber, and Ł. Cywi ński, "Blueprint of a scalable spin qubit shuttle device for coherent mid-range qubit transfer in disordered Si/SiGe/SiO_2", PRX Quantum **4**, 020305 (2023).

[119] M. D. Smet, Y. Matsumoto, A.-M. J. Zwerver, L. Tryputen, S. L. de Snoo, S. V. Amitonov, A. Sammak, N. Samkharadze, Ö. Gül, R. N. M. Wasserman, M. Rimbach-Russ, G. Scappucci, and L. M. K. Vandersypen, *High-fidelity single-spin shuttling in silicon*, 2024.

[120] H. Jnane, B. Undseth, Z. Cai, S. C. Benjamin, and B. Koczor, "Multicore quantum computing", Phys. Rev. Appl. **18**, 044064 (2022).

[121] G. Pagano, P. W. Hess, H. B. Kaplan, W. L. Tan, P. Richerme, P. Becker, A. Kyprianidis, J. Zhang, E. Birckelbaw, M. R. Hernandez, Y. Wu, and C. Monroe, "Cryogenic trapped-ion system for large scale quantum simulation", Quantum Science and Technology **4**, 014004 (2018).

[122] M. Cetina, L. Egan, C. Noel, M. Goldman, D. Biswas, A. Risinger, D. Zhu, and C. Monroe, "Control of Transverse Motion for Quantum Gates on Individually Addressed Atomic Qubits", PRX Quantum **3**, 010334 (2022).

[123] K. A. Landsman, Y. Wu, P. H. Leung, D. Zhu, N. M. Linke, K. R. Brown, L. Duan, and C. Monroe, "Two-qubit entangling gates within arbitrarily long chains of trapped ions", Physical Review A **100**, 022332 (2019).

[124] A. Auffèves, "Quantum technologies need a quantum energy initiative", PRX Quantum **3**, 020101 (2022).

[125] M. Fellous-Asiani, J. H. Chai, Y. Thonnart, H. K. Ng, R. S. Whitney, and A. Auffèves, "Optimizing resource efficiencies for scalable full-stack quantum computers", PRX Quantum **4**, 040319 (2023).

[126] J. D. Cifuentes, T. Tanttu, P. Steinacker, S. Serrano, I. Hansen, J. P. Slack-Smith, W. Gilbert, J. Y. Huang, E. Vahapoglu, R. C. C. Leon, N. D. Stuyck, K. Itoh, N. Abrosimov, H.-J. Pohl, M. Thewalt, A. Laucht, C. H. Yang, C. C. Escott, F. E. Hudson, W. H. Lim, R. Rahman, A. S. Dzurak, and A. Saraiva, *Impact of electrostatic crosstalk on spin qubits in dense cmos quantum dot arrays*, 2023.

[127] J. Schuff, M. J. Carballido, M. Kotzagiannidis, J. C. Calvo, M. Caselli, J. Rawling, D. L. Craig, B. van Straaten, B. Severin, F. Fedele, S. Svab, P. C. Kwon, R. S. Eggli, T. Patlatiuk, N. Korda, D. Zumbühl, and N. Ares, *Fully autonomous tuning of a spin qubit*, 2024.

[128] A. Strikis, S. C. Benjamin, and B. J. Brown, "Quantum computing is scalable on a planar array of qubits with fabrication defects", Phys. Rev. Appl. **19**, 064081 (2023).

[129] *Ibm roadmap*, `https://research.ibm.com/blog/ibm-quantum-roadmap-2025`.

[130] Y. Wu, W.-S. Bao, S. Cao, F. Chen, M.-C. Chen, X. Chen, T.-H. Chung, H. Deng, Y. Du, D. Fan, M. Gong, C. Guo, C. Guo, S. Guo, L. Han, L. Hong, H.-L. Huang, Y.-H. Huo, L. Li, N. Li, S. Li, Y. Li, F. Liang, C. Lin, J. Lin, H. Qian, D. Qiao, H. Rong, H. Su, L. Sun, L. Wang, S. Wang, D. Wu, Y. Xu, K. Yan, W. Yang, Y. Yang, Y. Ye, J. Yin, C. Ying, J. Yu, C. Zha, C. Zhang, H. Zhang, K. Zhang, Y. Zhang, H. Zhao, Y. Zhao, L. Zhou, Q. Zhu, C.-Y. Lu, C.-Z. Peng, X. Zhu, and J.-W. Pan, "Strong Quantum Computational Advantage Using a Superconducting Quantum Processor", Physical Review Letters **127**, 180501 (2021).

[131] Y. Kim, A. Eddins, S. Anand, K. X. Wei, E. Van Den Berg, S. Rosenblatt, H. Nayfeh, Y. Wu, M. Zaletel, K. Temme, et al., "Evidence for the utility of quantum computing before fault tolerance", Nature **618**, 500 (2023).

[132] Z. Cai, "Resource Estimation for Quantum Variational Simulations of the Hubbard Model", Physical Review Applied **14**, 014059 (2020).

[133] 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).

[134] A. Ovide, S. Rodrigo, M. Bandic, H. V. Someren, S. Feld, S. Abadal, E. Alarcon, and C. G. Almudever, *Mapping quantum algorithms to multi-core quantum computing architectures*, 2023.

[135] A. Pastor, P. Escofet, S. B. Rached, E. Alarcón, P. Barlet-Ros, and S. Abadal, *Circuit partitioning for multi-core quantum architectures with deep reinforcement learning*, 2024.

[136] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, "Surface codes: Towards practical large-scale quantum computation", Physical Review A **86**, 032324 (2012).

[137] M. Li, D. Miller, M. Newman, Y. Wu, and K. R. Brown, "2D Compass Codes", Physical Review X **9**, 021041 (2019).

[138] C. Chamberland, G. Zhu, T. J. Yoder, J. B. Hertzberg, and A. W. Cross, "Topological and Subsystem Codes on Low-Degree Graphs with Flag Qubits", Physical Review X **10**, 011022 (2020).

[139] W. Dür and H.-J. Briegel, "Entanglement Purification for Quantum Computation", Physical Review Letters **90**, 067901 (2003).

[140] S. Krastanov, V. V. Albert, and L. Jiang, "Optimized Entanglement Purification", Quantum **3**, 123 (2019).

[141] N. H. Nickerson, J. F. Fitzsimons, and S. C. Benjamin, "Freely Scalable Quantum Technologies Using Cells of 5-to-50 Qubits with Very Lossy and Noisy Photonic Links", Physical Review X **4**, 041041 (2014).

[142] N. Lo Piparo, M. Hanks, C. Gravel, K. Nemoto, and W. J. Munro, "Resource Reduction for Distributed Quantum Information Processing Using Quantum Multiplexed Photons", Physical Review Letters **124**, 210503 (2020).

[143] R. Nigmatullin, C. J. Ballance, N. de Beaudrap, and S. C. Benjamin, "Minimally complex ion traps as modules for quantum communication and computing", New Journal of Physics **18**, 103028 (2016).

[144] N. Kalb, A. A. Reiserer, P. C. Humphreys, J. J. W. Bakermans, S. J. Kamerling, N. H. Nickerson, S. C. Benjamin, D. J. Twitchen, M. Markham, and R. Hanson, "Entanglement distillation between solid-state quantum network nodes", Science **356**, 928 (2017).

[145] 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).

[146] A. Blais, J. Gambetta, A. Wallraff, D. I. Schuster, S. M. Girvin, M. H. Devoret, and R. J. Schoelkopf, "Quantum-information processing with circuit quantum electrodynamics", Physical Review A **75**, 032329 (2007).

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

[148] M. A. Sillanpää, J. I. Park, and R. W. Simmonds, "Coherent quantum state storage and transfer between two phase qubits via a resonant cavity", Nature **449**, 438 (2007).

[149] P.-Q. Jin, M. Marthaler, A. Shnirman, and G. Schön, "Strong Coupling of Spin Qubits to a Transmission Line Resonator", Physical Review Letters **108**, 190506 (2012).

[150] G. Zheng, "Circuit Quantum Electrodynamics with Single Electron Spins in Silicon", PhD thesis (Delft University of Technology, 2021).

[151] X. Mi, M. Benito, S. Putz, D. M. Zajac, J. M. Taylor, G. Burkard, and J. R. Petta, "A coherent spin–photon interface in silicon", Nature **555**, 599 (2018).

[152] A. J. Landig, J. V. Koski, P. Scarlino, U. C. Mendes, A. Blais, C. Reichl, W. Wegscheider, A. Wallraff, K. Ensslin, and T. Ihn, "Coherent spin–photon coupling using a resonant exchange qubit", Nature **560**, 179 (2018).

[153] F. Borjans, X. G. Croot, X. Mi, M. J. Gullans, and J. R. Petta, "Resonant microwave-mediated interactions between distant electron spins", Nature **577**, 195 (2020).

[154] M. Benito, J. R. Petta, and G. Burkard, "Optimized cavity-mediated dispersive two-qubit gates between spin qubits", Physical Review B **100**, 10.1103/PhysRevB.100.081412 (2019).

[155] A. Warren, E. Barnes, and S. E. Economou, "Long-distance entangling gates between quantum dot spins mediated by a superconducting resonator", Physical Review B **100**, 161303 (2019).

[156] A. Warren, U. Güngördü, J. P. Kestner, E. Barnes, and S. E. Economou, "Robust photon-mediated entangling gates between quantum dot spin qubits", Physical Review B **104**, 115308 (2021).

[157] D. Culcer, X. Hu, and S. Das Sarma, "Interface roughness, valley-orbit coupling, and valley manipulation in quantum dots", Phys. Rev. B **82**, 205315 (2010).

[158] M. Benito, X. Mi, J. M. Taylor, J. R. Petta, and G. Burkard, "Input-output theory for spin-photon coupling in Si double quantum dots", Phys. Rev. B **96**, 235434 (2017).

[159] J. C. Abadillo-Uriel, C. King, S. N. Coppersmith, and M. Friesen, "Long-range two-hybrid-qubit gates mediated by a microwave cavity with red sidebands", Phys. Rev. A **104**, 032612 (2021).

[160] W. K. Wootters, "Entanglement of Formation of an Arbitrary State of Two Qubits", Phys. Rev. Lett. **80**, 2245 (1998).

[161] F. Martins, F. K. Malinowski, P. D. Nissen, E. Barnes, S. Fallahi, G. C. Gardner, M. J. Manfra, C. M. Marcus, and F. Kuemmeth, "Noise Suppression Using Symmetric Exchange Gates in Spin Qubits", Phys. Rev. Lett. **116**, 116801 (2016).

[162] M. Veldhorst, J. C. C. Hwang, C. H. Yang, A. W. Leenstra, B. de Ronde, J. P. Dehollain, J. T. Muhonen, F. E. Hudson, K. M. Itoh, A. Morello, and A. S. Dzurak, "An addressable quantum dot qubit with fault-tolerant control-fidelity", Nat. Nanotechnol. **9**, 981 (2014).

[163] N. Lo Piparo, M. Hanks, C. Gravel, K. Nemoto, and W. J. Munro, "Resource Reduction for Distributed Quantum Information Processing Using Quantum Multiplexed Photons", Phys. Rev. Lett. **124**, 210503 (2020).

[164] R. Nigmatullin, C. J. Ballance, N. de Beaudrap, and S. C. Benjamin, "Minimally complex ion traps as modules for quantum communication and computing", New J. Phys. **18**, 103028 (2016).

[165] X. Mi, J. V. Cady, D. M. Zajac, P. W. Deelman, and J. R. Petta, "Strong coupling of a single electron in silicon to a microwave photon", Science **355**, 156 (2017).

[166] B. Buonacorsi, B. Shaw, and J. Baugh, "Simulated coherent electron shuttling in silicon quantum dots", Phys. Rev. B **102**, 125406 (2020).

[167] F. Ginzel, A. R. Mills, J. R. Petta, and G. Burkard, "Spin shuttling in a silicon double quantum dot", Phys. Rev. B **102**, 195418 (2020).

[168] M. Künne, A. Willmes, M. Oberländer, C. Gorjaew, J. D. Teske, H. Bhardwaj, M. Beer, E. Kammerloher, R. Otten, I. Seidler, R. Xue, L. R. Schreiber, and H. Bluhm, *The spinbus architecture: scaling spin qubits with electron shuttling*, 2023.

[169] X. Zhao and X. Hu, *Coherent electron transport in silicon quantum dots*, 2019.

[170] X. Li, E. Barnes, J. P. Kestner, and S. Das Sarma, "Intrinsic errors in transporting a single-spin qubit through a double quantum dot", Phys. Rev. A **96**, 012309 (2017).

[171] B. Koczor, "Exponential Error Suppression for Near-Term Quantum Devices", Phys. Rev. X **11**, 031057 (2021).

[172] W. J. Huggins, S. McArdle, T. E. O'Brien, J. Lee, N. C. Rubin, S. Boixo, K. B. Whaley, R. Babbush, and J. R. McClean, "Virtual Distillation for Quantum Error Mitigation", Phys. Rev. X **11**, 041036 (2021).

[173] B. Koczor, "The dominant eigenvector of a noisy quantum state", New J. Phys. **23**, 123047 (2021).

[174] R. Nandkishore and D. A. Huse, "Many-Body Localization and Thermalization in Quantum Statistical Mechanics", Annual Review of Condensed Matter Physics **6**, 15 (2015).

[175] D. J. Luitz, N. Laflorencie, and F. Alet, "Many-body localization edge in the random-field Heisenberg chain", Physical Review B **91**, 081103 (2015).

[176] A. M. Childs, D. Maslov, Y. Nam, N. J. Ross, and Y. Su, "Toward the first quantum simulation with quantum speedup", Proceedings of the National Academy of Sciences **115**, 9456 (2018).

[177] E. Farhi, J. Goldstone, and S. Gutmann, "A Quantum Approximate Optimization Algorithm", arXiv:1411.4028 [quant-ph] (2014).

[178] R. Babbush, N. Wiebe, J. McClean, J. McClain, H. Neven, and G. K.-L. Chan, "Low-Depth Quantum Simulation of Materials", Physical Review X **8**, 011044 (2018).

[179] D. Wecker, M. B. Hastings, and M. Troyer, "Progress towards practical quantum variational algorithms", Physical Review A **92**, 042303 (2015).

[180] C. Cade, L. Mineh, A. Montanaro, and S. Stanisic, "Strategies for solving the Fermi-Hubbard model on near-term quantum computers", Physical Review B **102**, 235122 (2020).

[181] R. Wiersema, C. Zhou, Y. de Sereville, J. F. Carrasquilla, Y. B. Kim, and H. Yuen, "Exploring Entanglement and Optimization within the Hamiltonian Variational Ansatz", PRX Quantum **1**, 020319 (2020).

[182] O. Higgott, D. Wang, and S. Brierley, "Variational Quantum Computation of Excited States", Quantum **3**, 156 (2019).

[183] T. Jones, S. Endo, S. McArdle, X. Yuan, and S. C. Benjamin, "Variational quantum algorithms for discovering Hamiltonian spectra", Physical Review A **99**, 062304 (2019).

[184] X. Yuan, S. Endo, Q. Zhao, Y. Li, and S. C. Benjamin, "Theory of variational quantum simulation", Quantum **3**, 191 (2019).

[185] S. Endo, J. Sun, Y. Li, S. C. Benjamin, and X. Yuan, "Variational Quantum Simulation of General Processes", Physical Review Letters **125**, 010501 (2020).

[186] B. Koczor and S. C. Benjamin, "Quantum natural gradient generalized to noisy and nonunitary circuits", Phys. Rev. A **106**, 062416 (2022).

[187] B. Koczor and S. C. Benjamin, "Quantum analytic descent", Physical Review Research **4**, 023017 (2022).

[188] A. Mari, T. R. Bromley, and N. Killoran, "Estimating the gradient and higher-order derivatives on quantum hardware", Physical Review A **103**, 012405 (2021).

[189] X. Yuan, J. Sun, J. Liu, Q. Zhao, and Y. Zhou, "Quantum Simulation with Hybrid Tensor Networks", Physical Review Letters **127**, 040501 (2021).

[190] E. Garlatti, T. Guidi, S. Ansbro, P. Santini, G. Amoretti, J. Ollivier, H. Mutka, G. Timco, I. J. Vitorica-Yrezabal, G. F. S. Whitehead, R. E. P. Winpenny, and S. Carretta, "Portraying entanglement between molecular qubits with four-dimensional inelastic neutron scattering", Nature Communications **8**, 14543 (2017).

[191] G. A. Timco, S. Carretta, F. Troiani, F. Tuna, R. J. Pritchard, C. A. Muryn, E. J. L. McInnes, A. Ghirri, A. Candini, P. Santini, G. Amoretti, M. Affronte, and R. E. P. Winpenny, "Engineering the coupling between molecular spin qubits by coordination chemistry", Nature Nanotechnology **4**, 173 (2009).

[192] A. Y. Kitaev, "Unpaired Majorana fermions in quantum wires", Physics-Uspekhi **44**, 131 (2001).

[193] J. D. Sau, S. Tewari, R. M. Lutchyn, T. D. Stanescu, and S. Das Sarma, "Non-Abelian quantum order in spin-orbit-coupled semiconductors: Search for topological Majorana particles in solid-state systems", Physical Review B **82**, 214509 (2010).

[194] S. Akibue and M. Murao, "Network Coding for Distributed Quantum Computation Over Cluster and Butterfly Networks", IEEE Transactions on Information Theory **62**, 6620 (2016).

[195] P. Andrés-Martínez and C. Heunen, "Automated distribution of quantum circuits via hypergraph partitioning", Physical Review A **100**, 032308 (2019).

[196] D. Ferrari, A. S. Cacciapuoti, M. Amoretti, and M. Caleffi, "Compiler Design for Distributed Quantum Computing", IEEE Transactions on Quantum Engineering **2**, 1 (2021).

[197] S. Bravyi, G. Smith, and J. A. Smolin, "Trading Classical and Quantum Computational Resources", Physical Review X **6**, 021043 (2016).

[198] T. Peng, A. W. Harrow, M. Ozols, and X. Wu, "Simulating Large Quantum Circuits on a Small Quantum Computer", Physical Review Letters **125**, 150504 (2020).

[199] W. Tang, T. Tomesh, M. Suchara, J. Larson, and M. Martonosi, "CutQC: using small Quantum computers for large Quantum circuit evaluations", in Proceedings of the 26th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, ASPLOS 2021 (Apr. 2021), pp. 473–486.

[200] Y. Li and S. C. Benjamin, "Hierarchical surface code for network quantum computing with modules of arbitrary size", Physical Review A **94**, 042303 (2016).

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

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

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

[204] S. Endo, S. C. Benjamin, and Y. Li, "Practical quantum error mitigation for near-future applications", Phys. Rev. X **8**, 031027 (2018).

[205] A. Kandala, K. Temme, A. D. Córcoles, A. Mezzacapo, J. M. Chow, and J. M. Gambetta, "Error mitigation extends the computational reach of a noisy quantum processor", Nature **567**, 491 (2019).

[206] Y. Kim, C. J. Wood, T. J. Yoder, S. T. Merkel, J. M. Gambetta, K. Temme, and A. Kandala, "Scalable error mitigation for noisy quantum circuits produces competitive expectation values", Nature Physics, 1 (2023).

[207] T. E. O'Brien, G. Anselmetti, F. Gkritsis, V. Elfving, S. Polla, W. J. Huggins, O. Oumarou, K. Kechedzhi, D. Abanin, R. Acharya, et al., "Purification-based quantum error mitigation of pair-correlated electron simulations", arXiv preprint arXiv:2210.10799 (2022).

[208] S. Endo, S. C. Benjamin, and Y. Li, "Practical quantum error mitigation for near-future applications", Phys. Rev. X **8**, 031027 (2018).

[209] A. Strikis, D. Qin, Y. Chen, S. C. Benjamin, and Y. Li, "Learning-based quantum error mitigation", PRX Quantum **2**, 040330 (2021).

[210] A. Lowe, M. H. Gordon, P. Czarnik, A. Arrasmith, P. J. Coles, and L. Cincio, "Unified approach to data-driven quantum error mitigation", Phys. Rev. Res. **3**, 033098 (2021).

[211] E. Van Den Berg, Z. K. Minev, A. Kandala, and K. Temme, "Probabilistic error cancellation with sparse pauli–lindblad models on noisy quantum processors", Nature Physics, 1 (2023).

[212] A. Montanaro and S. Stanisic, "Error mitigation by training with fermionic linear optics", arXiv preprint arXiv:2102.02120 (2021).

[213] Y. Li and S. C. Benjamin, "Efficient variational quantum simulator incorporating active error minimization", Phys. Rev. X **7**, 021050 (2017).

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

[215] S. McArdle, X. Yuan, and S. Benjamin, "Error-Mitigated Digital Quantum Simulation", Physical Review Letters **122**, 180501 (2019).

[216] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, "Quantum computational chemistry", Reviews of Modern Physics **92**, 015003 (2020).

[217] L. Botelho, A. Glos, A. Kundu, J. A. Miszczak, Ö. Salehi, and Z. Zimborás, "Error mitigation for variational quantum algorithms through mid-circuit measurements", Phys. Rev. A **105**, 022441 (2022).

[218] S. Aaronson, "Shadow tomography of quantum states", in Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018 (June 2018), pp. 325–338.

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

[220] H. C. Nguyen, J. L. Bönsel, J. Steinberg, and O. Gühne, "Optimizing Shadow Tomography with Generalized Measurements", Phys. Rev. Lett. **129**, 220502 (2022).

[221] K. Wan, W. J. Huggins, J. Lee, and R. Babbush, "Matchgate Shadows for Fermionic Quantum Simulation", arXiv:2207.13723 (2022).

[222] C. Bertoni, J. Haferkamp, M. Hinsche, M. Ioannou, J. Eisert, and H. Pashayan, "Shallow shadows: expectation estimation using low-depth random clifford circuits", arXiv preprint arXiv:2209.12924 (2022).

[223] 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).

[224] 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).

[225] J. Preskill, "Quantum Computing in the NISQ era and beyond", Quantum **2**, 79 (2018).

[226] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, "Variational quantum algorithms", Nature Reviews Physics **3**, 625 (2021).

[227] S. Endo, Z. Cai, S. C. Benjamin, and X. Yuan, "Hybrid Quantum-Classical Algorithms and Quantum Error Mitigation", Journal of the Physical Society of Japan **90**, 032001 (2021).

[228] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, W.-K. Mok, S. Sim, L.-C. Kwek, and A. Aspuru-Guzik, "Noisy intermediate-scale quantum algorithms", Rev. Mod. Phys. **94**, 015004 (2022).

[229] B. van Straaten and B. Koczor, "Measurement cost of metric-aware variational quantum algorithms", PRX Quantum **2**, 030324 (2021).

[230] A. Seif, Z.-P. Cian, S. Zhou, S. Chen, and L. Jiang, "Shadow distillation: quantum error mitigation with classical shadows for near-term quantum processors", PRX Quantum **4**, 010303 (2023).

[231] Y. Zhou and Z. Liu, "A hybrid framework for estimating nonlinear functions of quantum states", arXiv (2023).

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

[233] D. E. Koh and S. Grewal, "Classical Shadows With Noise", Quantum **6**, 776 (2022).

[234] A. K. Ekert, C. M. Alves, D. K. L. Oi, M. Horodecki, P. Horodecki, and L. C. Kwek, "Direct estimations of linear and nonlinear functionals of a quantum state", Phys. Rev. Lett. **88**, 217901 (2002).

[235] W. Hoeffding, in *Breakthroughs in Statistics*, edited by S. Kotz and N. L. Johnson, Springer Series in Statistics (Springer, New York, 1992), pp. 308–334.

[236] M. C. Tran, C.-F. Chen, A. Ehrenberg, A. Y. Guo, A. Deshpande, Y. Hong, Z.-X. Gong, A. V. Gorshkov, and A. Lucas, "Hierarchy of linear light cones with long-range interactions", Phys. Rev. X **10**, 031009 (2020).

[237] M. C. Tran, K. Sharma, and K. Temme, "Locality and Error Mitigation of Quantum Circuits", arXiv:2303.06496 (2023).

[238] D. Grinko, J. Gacon, C. Zoufal, and S. Woerner, "Iterative quantum amplitude estimation", en, npj Quantum Information **7**, 1 (2021).

[239] G. Boyd and B. Koczor, "Training variational quantum circuits with covar: covariance root finding with classical shadows", Phys. Rev. X **12**, 041022 (2022).

[240] H. H. S. Chan, R. Meister, M. L. Goh, and B. Koczor, "Algorithmic shadow spectroscopy", arXiv:2212.11036 (2023).

[241] B. Koczor and S. C. Benjamin, "Quantum analytic descent", Physical Review Research **4**, 023017 (2022).

[242] R. Nandkishore and D. A. Huse, "Many-Body Localization and Thermalization in Quantum Statistical Mechanics", Annu. Rev. Condens. Matter Phys. **6**, 15 (2015).

[243] A. M. Childs, D. Maslov, Y. Nam, N. J. Ross, and Y. Su, "Toward the first quantum simulation with quantum speedup", Proc. Natl. Acad. Sci. U.S.A. **115**, 9456 (2018).

[244] D. J. Luitz, N. Laflorencie, and F. Alet, "Many-body localization edge in the random-field Heisenberg chain", Phys. Rev. B **91**, 081103 (2015).

[245] R. Meister, *PyQuEST - A Python interface for the Quantum Exact Simulation Toolkit*, Aug. 2022.

[246] I. Pogorelov, T. Feldker, C. D. Marciniak, L. Postler, G. Jacob, O. Krieglsteiner, V. Podlesnic, M. Meth, V. Negnevitsky, M. Stadler, B. Höfer, C. Wächter, K. Lakhmanskiy, R. Blatt, P. Schindler, and T. Monz, "Compact ion-trap quantum computing demonstrator", PRX Quantum **2**, 020343 (2021).

[247] S. Moses, C. Baldwin, M. Allman, R. Ancona, L. Ascarrunz, C. Barnes, J. Bartolotta, B. Bjork, P. Blanchard, M. Bohn, et al., "A race track trapped-ion quantum processor", arXiv preprint arXiv:2305.03828 (2023).

[248] O. Gühne, M. Reimpell, and R. F. Werner, "Estimating entanglement measures in experiments", Phys. Rev. Lett. **98**, 110502 (2007).

[249] N. Friis, O. Marty, C. Maier, C. Hempel, M. Holzäpfel, P. Jurcevic, M. B. Plenio, M. Huber, C. Roos, R. Blatt, and B. Lanyon, "Observation of Entangled States of a Fully Controlled 20-Qubit System", Phys. Rev. X **8**, 021012 (2018).

[250] Y. Dai, Y. Dong, Z. Xu, W. You, C. Zhang, and O. Gühne, "Experimentally Accessible Lower Bounds for Genuine Multipartite Entanglement and Coherence Measures", Phys. Rev. Applied **13**, 054022 (2019).

[251] H.-Y. Huang, R. Kueng, G. Torlai, V. V. Albert, and J. Preskill, "Provably efficient machine learning for quantum many-body problems", Science **377**, `10.1126/science.abk3333` (2022).

[252] T. E. Baker and D. Poulin, "Density functionals and kohn-sham potentials with minimal wavefunction preparations on a quantum computer", Physical Review Research **2**, 043238 (2020).

[253] C. Dasgupta and S.-k. Ma, "Low-temperature properties of the random heisenberg antiferromagnetic chain", Phys. Rev. B **22**, 1305 (1980).

[254] S.-k. Ma, C. Dasgupta, and C.-k. Hu, "Random antiferromagnetic chain", Phys. Rev. Lett. **43**, 1434 (1979).

[255] A. Rath, C. Branciard, A. Minguzzi, and B. Vermersch, "Quantum fisher information from randomized measurements", Phys. Rev. Lett. **127**, 260501 (2021).

[256] B. Koczor, S. Endo, T. Jones, Y. Matsuzaki, and S. C. Benjamin, "Variational-state quantum metrology", New Journal of Physics **22**, 083038 (2020).

[257] Z. Cai, A. Chapman, H. Jnane, and B. Koczor, *Biased estimator channels for classical shadows*, 2024.

[258] H.-Y. Huang, R. Kueng, and J. Preskill, "Efficient estimation of pauli observables by derandomization", Phys. Rev. Lett. **127**, 030503 (2021).

[259] A. Zhao, N. C. Rubin, and A. Miyake, "Fermionic partial tomography via classical shadows", Phys. Rev. Lett. **127**, 110504 (2021).

[260] C. Ferrie and R. Blume-Kohout, "Maximum likelihood quantum state tomography is inadmissible", arXiv e-prints, arXiv:1808.01072, arXiv:1808.01072 (2018).

[261] M. Schuld, V. Bergholm, C. Gogolin, J. Izaac, and N. Killoran, "Evaluating analytic gradients on quantum hardware", Physical Review A **99**, 032331 (2019).

[262] D. Wierichs, J. Izaac, C. Wang, and C. Y.-Y. Lin, "General parameter-shift rules for quantum gradients", Quantum **6**, 677 (2022).

[263] A. Zhao and A. Miyake, *Group-theoretic error mitigation enabled by classical shadows and symmetries*, 2023.

[264] A. Gresch and M. Kliesch, *Guaranteed efficient energy estimation of quantum many-body hamiltonians using shadowgrouping*, 2023.

[265] Z. Zhu, J. M. Lukens, and B. T. Kirby, *On the connection between least squares, regularization, and classical shadows*, 2024.

[266] D. J. Ibberson, T. Lundberg, J. A. Haigh, L. Hutin, B. Bertrand, S. Barraud, C.-M. Lee, N. A. Stelmashenko, G. A. Oakes, L. Cochrane, J. W. Robinson, M. Vinet, M. F. Gonzalez-Zalba, and L. A. Ibberson, "Large Dispersive Interaction between a CMOS Double Quantum Dot and Microwave Photons", PRX Quantum **2**, 020315 (2021).

[267] M. Am-Shallem, A. Levy, I. Schaefer, and R. Kosloff, "Three approaches for representing Lindblad dynamics by a matrix-vector notation", arXiv:1510.08634 [quant-ph] (2015).

[268] X. Xue, B. Patra, J. P. G. van Dijk, N. Samkharadze, S. Subramanian, A. Corna, B. Paquelet Wuetz, C. Jeon, F. Sheikh, E. Juarez-Hernandez, B. P. Esparza, H. Rampurawala, B. Carlton, S. Ravikumar, C. Nieva, S. Kim, H.-J. Lee, A. Sammak, G. Scappucci, M. Veldhorst, F. Sebastiano, M. Babaie, S. Pellerano, E. Charbon, and L. M. K. Vandersypen, "CMOS-based cryogenic control of silicon quantum circuits", Nature **593**, 205 (2021).

[269] E. J. Connors, ̇*Nelson*, and J. M. Nichol, "Rapid High-Fidelity Spin-State Readout in Si/Si-Ge Quantum Dots via rf Reflectometry", Physical Review Applied **13**, 024019 (2020).

[270] G. A. Oakes, V. N. Ciriano-Tejel, D. F. Wise, M. A. Fogarty, T. Lundberg, C. Lainé, S. Schaal, F. Martins, D. J. Ibberson, L. Hutin, B. Bertrand, N. Stelmashenko, J. W. A. Robinson, L. Ibberson, A. Hashim, I. Siddiqi, A. Lee, M. Vinet, C. G. Smith, J. J. L. Morton, and M. F. Gonzalez-Zalba, "Fast high-fidelity single-shot readout of spins in silicon using a single-electron box", Phys. Rev. X **13**, 011023 (2023).

[271] M. R. Jerrum, L. G. Valiant, and V. V. Vazirani, "Random generation of combinatorial structures from a uniform distribution", Theoret. Comput. Sci. **43**, 169 (1986).

[272] M. Lerasle, "Selected topics on robust statistical learning theory", arXiv:1908.10761 (2019).