Exotic Circuits for Enhanced Quantum Algorithms and Computation

by

Albie Chan

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Physics (Quantum Information)

Waterloo, Ontario, Canada, 2023

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

This thesis consists of material adapted from two manuscripts (arXiv:2305.19200 and arXiv:2304.08529) submitted for publication, in which I was lead and contributing author respectively.

Chapter 2: This chapter is based on arXiv:2305.19200 (lead author: A. Chan). I produced all VQE simulation results shown in Figs. 2.3, 2.4, 2.6, and 2.7, and all supplementary data presented in App. C. I designed and constructed all of Figs. 2.1–2.7. Fig. 2.1 was designed with assistance from Z. Shi, who also provided invaluable support with performing simulations involving the IBM Quantum platform. C. Muschik conceived design ideas for Figs. 2.1, 2.3, and 2.6, and provided iterative feedback during their construction.

Chapter 3: This chapter is an independent, abridged writeup based on arXiv:2304.08529 (lead author: J.M. Ramiro). The material in this chapter adapted ideas and concepts from the manuscript and supplementary notes by J.M. Ramiro, L. Dellantonio, and C. Muschik. I produced all numerical results involving the coherent gate-based (Fig. 3.3) and measurement-based (Fig. 3.5) protocols as an independent verification of the results presented in the manuscript. The measurement-based protocol described in Sec. 3.3.1 and App. H is a modified version of the original protocol, which I developed following discussions with Z. Shi. J.M. Ramiro, L. Dellantonio, and W. Dür formulated the theory behind both gate-based (Sec. 3.2.1) and measurement-based (Sec. 3.3.1) protocols. L. Dellantonio originally derived Eqs. (3.17) and (3.23). I designed and constructed Figs. 3.3 and 3.5. Figs. 3.1, 3.2, and 3.4 were designed by J.M. Ramiro and L. Dellantonio, and are reproduced in this thesis with kind permission from the aforementioned authors.

Contributions to other parts of the thesis are as follows:

C. Muschik and Z. Shi provided substantial input on the introduction (**Chapter 1**) and conclusions (**Chapter 4**), and contributed key ideas discussed therein.

I designed and constructed Figs. A.1, B.1, B.2, F.1, F.2, and G.1, and performed all equation derivations shown in the appendixes.

Appendix D: The procedure for error analysis was proposed by Z. Shi during one of many impromptu whiteboard discussions.

Appendix F: The `fast_wolf` algorithm was formulated by W. Dür and first employed in the MB-VQE scheme introduced in Phys. Rev. Lett. 126, 220501. The pseudocode shown (Alg. 1) is based on Z. Shi's implementation of the algorithm.

All chapters and appendixes in this thesis incorporated helpful suggestions and feedback from L. Dellantonio, Z. Shi, and C. Muschik.

Abstract

Quantum circuits play an essential role in many disciplines of quantum information science. They can not only be represented in the traditional gate-based (GB) paradigm, but also an alternative measurement-based (MB) paradigm. The latter begins from a resource state possessing many entangled ancillary qubits, and proceeds via mid-circuit measurements of these ancillas. More generally, ancillary qubits feature prominently in quantum error-correcting schemes, opening up the possibility for exploiting both paradigms on an equal footing. In this thesis, we explore how circuits designed with such features (referred to as "exotic") can be employed to enhance the quality of quantum algorithms and their general computations.

First, we examine variational quantum eigensolvers (VQEs) – a highly successful class of quantum algorithms for simulating physical models on quantum computers. Initially developed as a GB technique, they were recently extended to the MB model which enabled a first paradigm shift in the field. Our work seeks to push the integration and design frontiers of the VQE further by blending aspects of the GB and MB paradigms to form *hybrid* variational circuits. This can facilitate the design of a problem-informed variational ansatz, while enabling the efficient implementation of many-body Hamiltonians on noisy-intermediate-scale-quantum (NISQ) devices. To test the effectiveness of our approach, we demonstrate our hybrid circuits on IBM superconducting quantum computers by performing VQEs involving the perturbed planar code, $\mathbb{Z}_2$ lattice gauge theory, quantum chromodynamics, and the lithium hydride molecule.

Second, we propose a technique to mitigate the problem of noise and decoherence in quantum computation, which affect VQEs and many other quantum algorithms. Specifically, we consider circuits containing auxiliary registers and perform identical noisy computations on them in a superimposed fashion. We describe a concrete, scalable protocol for executing these circuits in both the GB and the MB paradigms, comparing and contrasting the mechanisms in which noise impacts them. Through analytical and numerical analyses, we illustrate how performing circuits in this manner leads to a noise interference process that yields improvements in computational fidelity.

The works discussed in this thesis constitute novel applications of exotic circuits that are both resource-efficient and versatile in nature. Together, these qualities can pave the way for more reliable quantum algorithms and computations on NISQ hardware.

Acknowledgements

"Complacency is the forerunner of mediocrity. You can never work too hard on attitudes, effort and technique." – Don Meyer

I would like to extend my sincere gratitude to those who have supported me throughout my MSc program at Waterloo. As an newcomer to quantum physics, I knew that embarking on such a journey would be rife with unexpected challenges and learning curves. However, these individuals helped bring a human side to a uncompromising discipline, and that has made it all the more bearable.

First, to my research collaborators from IQC and Innsbruck: Luca Dellantonio, Zheng Shi, Jorge Miguel-Ramiro, and Wolfgang Dür. Thank you for giving me the opportunity to work on two fantastic projects, as well as a forum to present and express my ideas. I thoroughly enjoyed our often lengthy, but friendly discourses on variational algorithms and quantum information theory. They were full of interesting twists and nothing short of enlightening. Many thanks for your thoughtful suggestions and feedback, and for working tirelessly to steer progress in the right direction. I would also like to acknowledge the other group members – Jinglei, Yasar, Jan, Victor, and Abhijit for being such a positive presence, and my committee members Chris Wilson and Thomas Jennewein for taking the time to hear about my work. It's been a privilege getting to know you all.

A warm, heartfelt acknowledgement goes to my supervisor, Christine Muschik, who accepted me under her wing despite my knowing little about high-energy physics or lattice gauge theories. I have learned much about these topics from her, and I could not be more thankful for the patience, guidance, and understanding she has shown me. Her attention to detail has also inspired me to strive for excellence in all my endeavours. I want to thank her for being a terrific group leader and for staying on top of things when it came to critical deadlines and paper submissions. I especially admire her knack for conceiving brilliant "rescue" plans whenever we seemed to hit brick walls.

Further acknowledgement goes to Emma Barlett and Erica Boland for being great admin supports at IQC. Their efforts to facilitate meetings and other logistics have helped me tremendously during hectic times. They are both exceptional individuals I could always count on.

Lee, Neil, and Russell – thank you for supporting my decision to pursue grad studies. To my parents and my extended family – thanks for all the love and support you've given me. Terence, without your timely advice and encouragement, I would not be where I am today. And finally, to everyone else I've gotten to know over the past eight years, thanks for being part of a budding stage in my career.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Quantum circuits are a core ingredient for describing quantum computations (QC) and their associated operations [1]. They have been instrumental to the development of fundamental concepts in quantum algorithms, cryptography, and communication [1]. Quantum circuits are conventionally expressed in the *gate-based* (GB)[1] paradigm, where a sequence of black-box unitaries act on a target state of interest. Additionally, they may incorporate non-unitary measurements during the qubit readout stage. These measurements are notable for having no analogue in classical circuits, and highlight the probabilistic nature of quantum-mechanical systems [1].

While GB circuits cast measurements in a supporting role, Raussendorf, Browne, and Briegel [2] proposed a computing approach that relies primarily on measurements. Known as the one-way scheme, it employs the *measurement-based* (MB) paradigm [3, 4] and performs computations using an entangled cluster or graph state (containing ancillary qubits) instead of unitary gates. Here, projective measurements of the ancillas drive the computation forward and are made possible by the correlations existing between them. Despite its irreversibility, MB-QC provides a convenient realization on platforms adept at generating large entangled clusters, particularly optical lattices and solid-state systems [2, 5]. These have proved useful in applications involving multipartite systems such as quantum communication [6], blind quantum computation [7], and quantum error correction with surface codes [8]. On the other hand, the theoretical description of MB-QC employs powerful concepts from graph theory and the stabilizer formalism, enabling one to express it in both graph and circuit languages. Interestingly, their circuit representation possesses elements

---

[1]Although the terms "gate-based" and "circuit-based" are often used synonymously, we use the former throughout since the measurement-based paradigm is also expressible in terms of circuits (albeit less frequently than clusters or graphs).

that are uncommon in standard computing applications thus far, including mid-circuit measurements, and conditional operations. Yet remarkably, they maintain a direct correspondence with GB circuits, with the existence of well-known techniques (e.g. phase-map decomposition [9], ZX-calculus [10]) for translating between the paradigms. Thus, the MB paradigm may be regarded as a alternative perspective for harnessing quantum computing power, with its own unique circuit elements and practical considerations.

Aside from MB-QC, the use of *auxiliary* registers (wires) in a circuit also carries significance in quantum information theory. A well-known example is that of quantum error-correcting codes [8, 11, 12], where multiple physical qubits are needed to construct a single fault-tolerant logical qubit. The auxiliary registers are entangled with the input state and then measured to obtain the syndrome of possible errors, such that they may be detected and corrected according to the measurement outcomes. On the other hand, quantum error *mitigation* in the NISQ era [13, 14] remains the more practical approach to leveraging current progress in quantum hardware development. While not all error mitigation techniques rely on auxiliary registers, entanglement purification offers us some inspiration [15, 16]: by employing one of the two low-fidelity Bell pairs as auxiliary systems, we can probabilistically increase the entanglement of the other pair. One may easily imagine the same idea being applied to the error mitigation problem.

It is apparent that the concepts of MB-QC, graph states, mid-circuit measurements, and auxiliary qubits suggest new, "exotic" possibilities for quantum circuits. Given this flexibility, it is natural to ask whether these features carry any benefits, for instance in enhancing quantum algorithms or general computations. This is precisely the question investigated in this thesis, as we examine two distinct applications of exotic circuits. In Chapter 2, we formulate these circuits as ansätze (trial states) for variational quantum eigensolvers (VQEs), a well-known quantum-classical algorithm used to study low-lying energy states. Specifically, we blend aspects of the GB and MB computational models within these ansätze in various ways, and demonstrate how such techniques can yield a more versatile approach to performing VQEs from both design and implementation-level standpoints. In Chapter 3, we present a more abstract formulation of exotic circuits in the context of noise analysis. Here, we show how coherently superimposing multiple instances of a noisy computation can lead to consistent reductions in the overall noise level, and hence improved performance. Akin to the first approach, we showcase how these superimposed circuits may be designed in both the GB and MB paradigms, while highlighting subtle differences in the way noise processes act on them. Finally, in Chapter 4, we draw conclusions based on the performance and viability of the exotic circuits, and look ahead to future areas of exploration where they may play a substantial role.

Chapter 2

Merging Computational Models for Hybrid Variational Quantum Eigensolvers

Many hard problems in physics require the simulation of complex Hamiltonians. Since their computational cost often renders classical methods intractable, quantum algorithms are required to obtain accurate solutions. However, these algorithms are curtailed by NISQ devices that are limited in both the number of qubits and the attainable circuit depths [17]. Since then, research efforts aimed at reducing requirements have focused on the class of *quantum-classical* algorithms. They employ quantum and classical processors working together in a feedback loop to approximate solutions numerically. One such algorithm is known as the *variational quantum eigensolver* (VQE), which determines a system's low-lying energy states via the classical optimization of a variational ansatz (trial state) [18, 19, 20]. This can provide useful information toward analyzing the system's quantum mechanical properties, including its finite-temperature behaviour and phase transitions [21, 22] Consequently, VQEs have been successful in areas spanning quantum chemistry [23], materials science [24, 25] and high-energy physics [21, 26, 27, 28, 29, 30]. Such developments have enabled opportunities for NISQ devices to tackle problems that would otherwise be out of reach.

Current demonstrations of VQEs are GB, in which the ansatz is expressed as a sequence of unitary black-box operations. However, as proposed in Ref. [22], VQEs can also operate using MB-QC (see App. A). The resulting algorithm, known as MB-VQE, expresses the ansatz as an entangled resource state [2, 31, 32, 33] which undergoes computation via a series of local rotations and measurements [22]. As an alternative, entirely gate-free

concept, MB-VQEs permit the use of VQEs on platforms for which the implementation of entangling gates is challenging. A unique feature of performing MB-QC and MB-VQE is the measurements that condition subsequent operations on their specific outcomes. These intermediate measurements are crucial to performing MB-QC deterministically, since they implement corrective operations that combat the random nature of measurement [2, 33].

In this chapter, we introduce *hybrid* VQEs that merge the strengths of the MB and GB paradigms for circuit implementations. In particular, we leverage current quantum hardware developments that have enabled fast, high-fidelity mid-circuit measurements on superconducting qubits, trapped ions, and ultracold neutral atoms [34, 35, 36, 37, 38, 39, 40]. This allows us to incorporate MB elements within these architectures, which predominately use the GB paradigm in their existing applications. We demonstrate the capabilities of our approach by performing a series of VQE simulations on IBM superconducting quantum computers. In doing so, we address two major challenges toward realizing effective VQEs – namely, the design of a good variational ansatz, and the physical realization of robust VQEs on NISQ devices.

To address the first challenge, we utilize MB principles to formulate an alternative approach for designing GB-VQE circuits. The versatility of this approach is seen when applied to problems with graph representations (MB patterns), as it permits one to access graph-inspired variational families directly in the GB formalism. It entails modifying each entangling gate with suitable operations [Fig. 2.1(a)], such that the equivalent graph is modified in a variationally tunable way. We apply this idea to the perturbed planar code (PC), as an example of how an effective GB-VQE ansatz can be designed based on an initial graph state. More broadly, this design tool can be useful when formulating a problem-informed VQE ansatz.

For the second challenge, we exploit MB elements that employ parameterized ancilla qubits and conditional measurements, and integrate them within VQE circuits. Here, the parameters dictate the choice of a measurement *basis* as opposed to rotations in parameter space, thus lending a unique variational aspect. We demonstrate this idea via MB patterns known as "Pauli gadgets", which realize n-body interaction terms as shown in Fig. 2.1(b). Such terms appear in many target Hamiltonians as a consequence of the underlying physics, particularly in models from condensed matter physics or in gauge theories. By realizing such interactions via measurements, these gadgets are implemented efficiently via the number of entangling operations, a current bottleneck on NISQ platforms. In conjunction with local rotation gates, the Pauli gadget gives rise to novel VQEs that merge GB and MB paradigms in their circuits. We perform VQE demonstrations of these "hybrid" circuits for 1D quantum chromodynamics, $\mathbb{Z}_2$ lattice gauge theory, and the lithium hydride molecule, showcasing their compact depth and effectiveness across a broad range of applications.
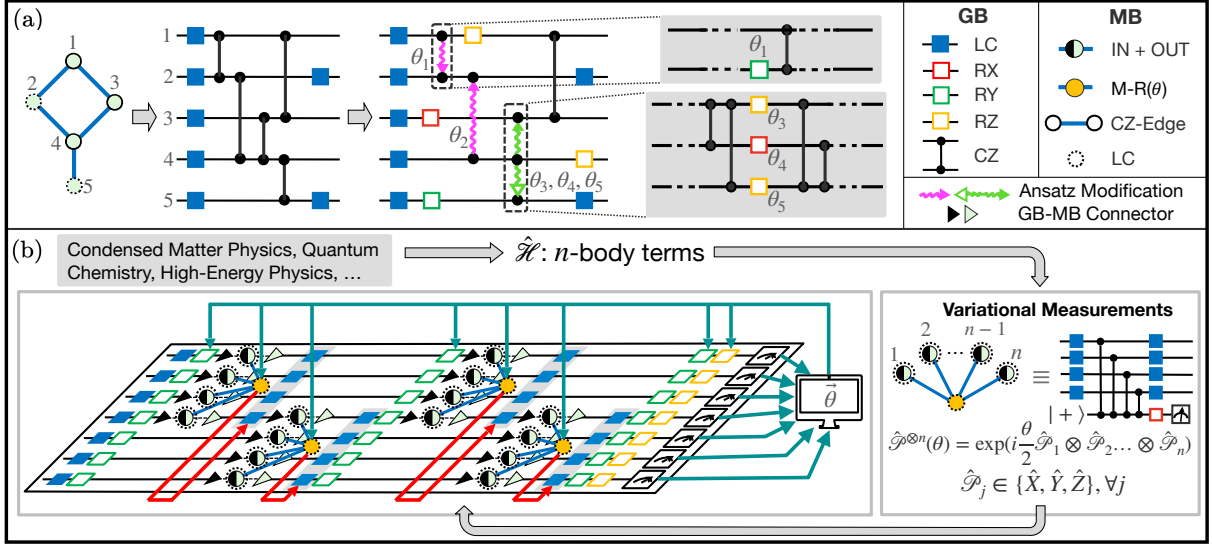
FIG. 2.1: **Enhancing VQEs by combining MB and GB principles to form hybrid ansatz circuits**. *(a)* Based on MB principles, gate modification (wavy lines with arrows) allows access to graph-inspired variational state families in the GB formalism. Starting from a base circuit derived from a graph state (MB pattern), we dress each entangling gate with unitary operations that tune the circuit's entanglement structure. *(b)* The Pauli gadgets $\hat{\mathcal{P}}^{\otimes n}(\theta)$ are MB patterns (graphs) that permit a VQE implementation of a hybrid ansatz on NISQ devices. Each gadget, when combined with local rotations, realizes an n-body interaction term in the target Hamiltonian. The parameterized ancilla qubits (in orange) provide a variational aspect by tuning the measurement bases used to perform MB-QC on them. These compact gadgets provide resource advantages in the number of entangling operations, and allows one to efficiently realize VQE circuits of arbitrary complexity. Grey, teal, and red arrows denote the overall workflow, classical–quantum VQE feedback loop, and corrective operations (i.e. byproduct Pauli operators $\hat{U}_{\Sigma}$) on the measured qubits respectively. On circuit platforms such as IBM Quantum, MB elements are implemented dynamically with mid-circuit measurements.

5

2.1 Background

In this section, we summarize the relevant concepts underlying hybrid circuits and VQEs.

In its simplest form, a VQE approximates the ground state (GS) energy E_0 of a target Hamiltonian $\hat{\mathcal{H}}$ with a parameterized ansatz state $|\psi_a(\vec{\theta})\rangle$. The traditional (GB) VQE represents the ansatz as a circuit whose gates are parameterized by $\vec{\theta}$. From the ansatz, the quantum processor computes the mean energy $E = \langle\psi_a|\hat{\mathcal{H}}|\psi_a\rangle$, which a classical optimizer minimizes by updating $\vec{\theta}$ [19]. The quantum processor then reevaluates the ansatz and the process repeats iteratively until a stopping condition for convergence (in E) is attained. A schematic of the VQE algorithm is shown in App. B.

The accuracy of a VQE in approximating E_0 relies heavily on the specific ansatz chosen. Despite the existence of well-established techniques (e.g. domain-specific and heuristics [41]) to formulate the ansatz, they are often beset by practical limitations. These include a susceptibility to barren plateaus and local minima in the energy landscape [42, 43]. It is also possible for the GS to be highly entangled, which can result in large circuit depth requirements. For example, heuristic designs often encapsulate many repeating layers with all-to-all connectivity [41]. This can pose challenges on NISQ devices due to restrictions in their quantum volumes and qubit connectivity. Such drawbacks motivate the use of our hybrid VQE approaches, which present a contrast through efficient ansatz design and implementation methods.

MB elements (patterns) process quantum information via the MB-QC protocol, which treats a resource state as a cluster or graph connecting a set of input, ancilla, and output qubits. The computation proceeds as a series of projective measurements on the resource in the eigenbases of the Pauli operators $\hat{X}, \hat{Y}, \hat{Z}$, or a rotated basis (e.g. $R(\theta) = (|0\rangle \pm e^{i\theta}|1\rangle)/\sqrt{2}$). Owing to the random nature of measurement, corrective operations (depending on the measurement outcomes) are applied to ensure the output result is always the same. These encompass byproduct operators $\hat{U}_\Sigma \in \{\hat{I}, \hat{X}, \hat{Z}, \hat{X}\hat{Z}, \hat{Z}\hat{X}\}$ acting on each output and angle adaptations $R(\theta) \to R(\pm\theta)$ for qubits measured in rotated bases. Together, the entangling, measurement, and corrective operators form an MB-QC realizing a specific unitary.

On IBM Quantum systems, MB-QC is performed on circuits using the *dynamic* mode of execution. It performs the projective measurements as mid-circuit measurements, and permits the application of corrective operations conditioned on their outcomes. By doing so, we avoid the necessity of having to postselect non-corrective outcomes (e.g. all '0'). This scales inefficiently since one would require on average $2^m N$ measurement samples (shots) to retain only N samples, where m is the number of mid-circuit measurements. We refer to Apps. A and B for further details.

6

2.2 Protocol design: ansatz modification

Here, we describe an MB-inspired tool for designing GB-VQE algorithms. It constructs VQE circuits by creating a problem-tailored ansatz and introducing variational modifications to its entangling operations, permitting one to access a physics-informed variational family. To understand the working principle, we first consider the tool on the level of graph states. These states are an equivalent representation of multi-qubit Pauli stabilizer states, being the $+1$ eigenvector of the correlation operator:

$$\hat{S}_n = \hat{X}_n \prod_k \hat{Z}_k, \tag{2.1}$$

where k are the nearest neighbour qubits (vertices) of qubit n. Eq. (2.1) is important as it describes the connectivity of graph states – each qubit n is stabilized by $\hat{X}$ and the qubits k it is connected to are each stabilized by $\hat{Z}$ (e.g. realizing a CZ interaction). It follows that when all qubits are initialized in $|+\rangle$, a CZ-edge between qubits n and k will realize the maximally entangled state:

$$CZ|+\rangle_n|+\rangle_k = \frac{|0\rangle_n|+\rangle_k + |1\rangle_n|-\rangle_k}{\sqrt{2}}. \tag{2.2}$$

From Eq. (2.2), we may view graphs as providing the "entanglement structure" (Fig. 2.2) necessary for MB-QC, which is consumed during qubit measurement. Additionally, their resource nature suggests that modifications can be made for instance, by attaching other graph state operations to their edges (see App. A). This novel idea was first explored in Ref. [22] to create custom ansätze that are highly adept at accessing finer corners of a system's Hilbert space (i.e. the states between two perturbative regimes). Provided that the modifications take place before any destructive measurements, they preserve the entanglement structure and thus provide an efficient way to generate arbitrarily complex graphs. These properties were found to be advantageous for problems whose unperturbed ground states possess a graph representation, such as the toric [44], planar [45, 46], and color [46] stabilizer codes. In these scenarios, one employs the ground state as the initial ansatz and performs the edge modifications when it is subjected to local perturbations.

If the edge modifications are unitary, GB circuit analogues of these graphs can be derived and modified in an equivalent fashion (see App. G). We focus on this version, known as "gate modification", where parameterized single- and two-qubit gates are used to dress each entangling gate (corresponding to a graph edge) to yield the same variational effect. As an added benefit, one can also enclose the ansatz with local rotation layers, which can further broaden the range of accessible states without incurring extra entangling gates.
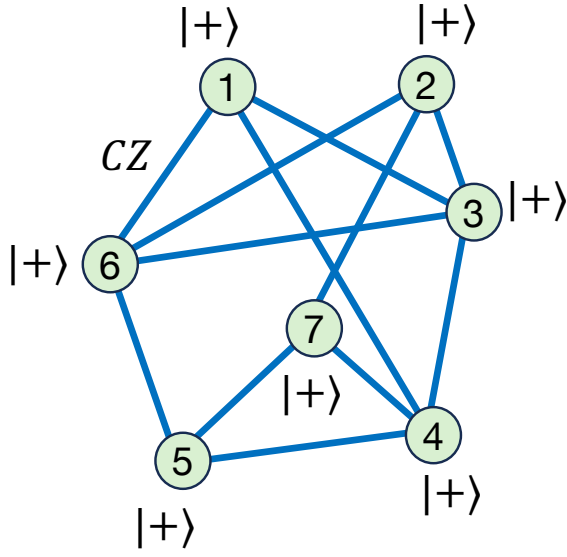
$$\hat{S}_1 = \hat{X}_1 \hat{Z}_3 \hat{Z}_4 \hat{Z}_6$$
$$\hat{S}_2 = \hat{X}_2 \hat{Z}_3 \hat{Z}_6 \hat{Z}_7$$
$$\hat{S}_3 = \hat{Z}_1 \hat{Z}_2 \hat{X}_3 \hat{Z}_4 \hat{Z}_6$$
$$\hat{S}_4 = \hat{Z}_1 \hat{Z}_3 \hat{X}_4 \hat{Z}_5 \hat{Z}_7$$
$$\hat{S}_5 = \hat{Z}_4 \hat{X}_5 \hat{Z}_6 \hat{Z}_7$$
$$\hat{S}_6 = \hat{Z}_1 \hat{Z}_2 \hat{Z}_3 \hat{Z}_5 \hat{X}_6$$
$$\hat{S}_7 = \hat{Z}_2 \hat{Z}_4 \hat{Z}_5 \hat{X}_7$$

FIG. 2.2: **Example of a graph state and its typical connectivity**. The vertices represent qubits (initialized in the $|+\rangle$ state) and the edges represent CZ operations between them. The graph's connectivity is fully captured by its n correlation operators $\hat{S}$, where the state of each qubit n is defined as a stabilizer state of $\hat{S}_n$.

2.3 2D planar code

We illustrate the ansatz-modification tool for the planar code (PC), a frequently employed surface code used for quantum error correction. Such codes are at the core of modern experimental efforts to realize fault-tolerant computation.

The PC is defined on a $M \times N$ lattice with open boundary conditions and qubits residing on the edges. Each qubit edge produces a gauge field that imposes local symmetries on each vertex. This results in a partitioning of the Hilbert space into non-overlapping charge sectors (via the Gauss' law), signifying the existence of electric field operators. When stray fields are present (i.e. in experimental settings), these act in the form of local perturbations on each qubit edge. In 2D, the gauge fields also self-interact over a square path of links (plaquette) which gives rise to a magnetic field. Consequently, the PC is stabilized by star ($\hat{S}_s$) and plaquette ($\hat{P}_p$) operators acting on the neighbouring qubits (edges) of each vertex s and plaquette p, respectively [44]. The overall PC Hamiltonian is therefore $\hat{\mathcal{H}}_{PC} = \hat{\mathcal{H}}_+ + \hat{\mathcal{H}}_\Box + \hat{\mathcal{H}}_\triangle$, where $\hat{\mathcal{H}}_+$ and $\hat{\mathcal{H}}_\Box$ are the star and plaquette terms respectively, and $\hat{\mathcal{H}}_\triangle$ is an added perturbation term. These terms are given by:

$$\hat{\mathcal{H}}_+ = -\sum_{s=1}^{n_s} \hat{S}_s; \;\; \hat{\mathcal{H}}_\Box = -\sum_{p=1}^{n_p} \hat{P}_p; \;\; \hat{\mathcal{H}}_\triangle = \xi \sum_{q=1}^{n_q} \hat{Z}_q;$$
$$\hat{P}_p = \prod_{i \in \Box_p} \hat{X}_i; \;\; \hat{S}_s = \prod_{i \in +_s} \hat{Z}_i.$$

(2.3)

Here, n_s and n_p are the numbers of vertices and plaquettes, respectively, and ξ is the strength of the perturbation acting on each of the n_q qubits. Furthermore, as shown in Fig. 2.3(a), $i \in \Box_p$ ($i \in +_s$) indicates the set of qubit edges belonging to a plaquette p (vertex s). The perturbation $\hat{\mathcal{H}}_\triangle$ imparts changes to the degree of entanglement in the ground state, yielding a continuum of states from maximally entangled (for $\xi = 0$), to the fully polarized $|1\rangle^{\otimes n_q}$ state (for $\xi \to \infty$). The intermediate regime, in particular, characterizes imperfect realizations of the error-correcting code [47].

The hybrid ansatz is constructed from a graph state representation of the GS of the PC Hamiltonian ($\xi = 0$) [Fig. 2.3(a; i–ii)]. It carries an equivalent [2] circuit representation consisting of CZ and H gates [Fig. 2.3(a)]. While this state starts the VQE in a suitable corner of the Hilbert space, the GS of the full Hamiltonian is found by adding modifications containing variational parameters following the strategy in Fig. 2.1(a). Here, we modify each CZ gate between qubits m and n so that it realizes $CZ_{m,n}RY(\theta)_n$ [Fig. 2.3(a; iii)]. The parameter θ effectively tunes the entanglement generated by a CZ gate and the resulting variable ansatz permits us to approximate the GS for different $\xi \in [0, \infty)$.

9
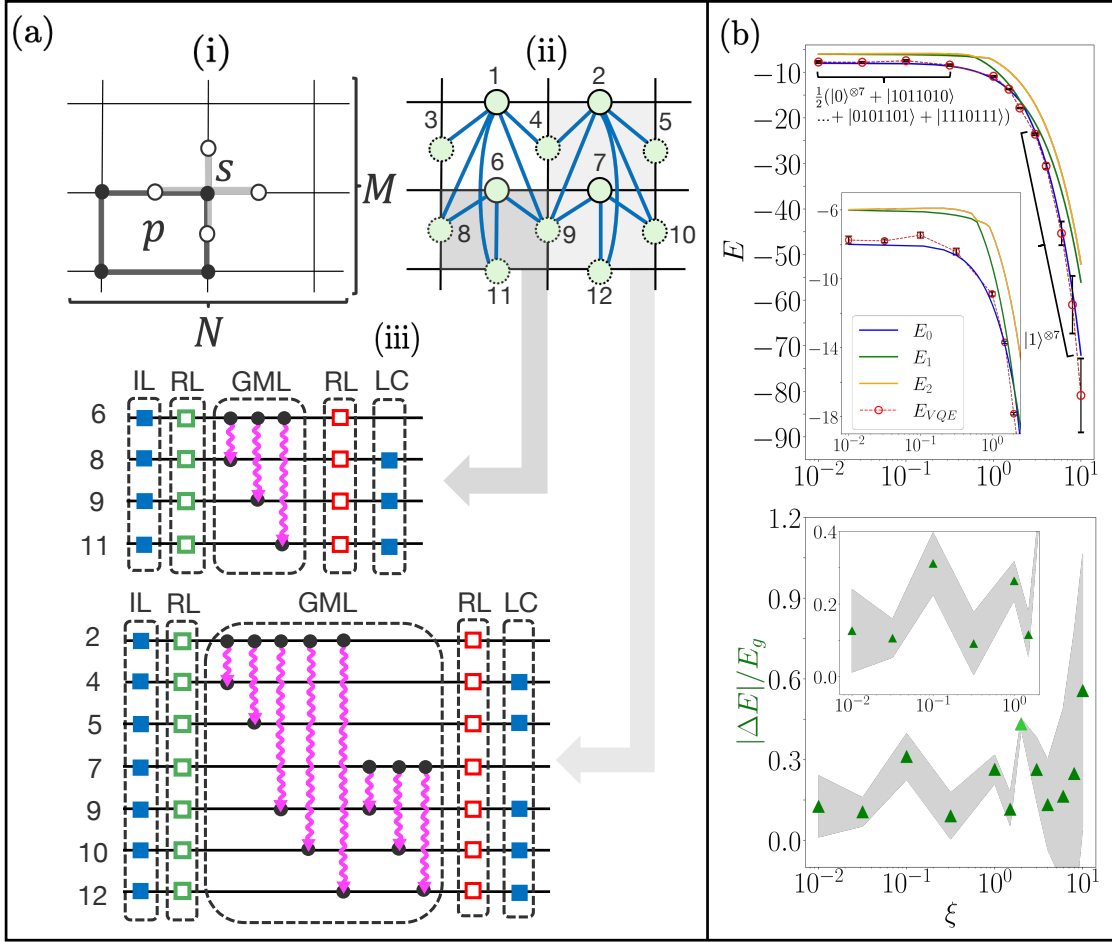
FIG. 2.3: **VQE demonstration of ansatz modification on IBM Quantum systems, applied to the perturbed planar code (PC)**. *(a) (i)* PC depicted as an $M \times N$ lattice in 2D, with star $(\hat{S}_s)$ and plaquette $(\hat{P}_p)$ stabilizer operators. *(ii)* GS for the $M \times N$ PC as a graph, showing qubits (circles) residing on the edges. *(iii)* VQE ansatz circuit for the 1×1 and 2×1 perturbed PC based on *(ii)*, where each entangling CZ (corresponding to a graph edge) is modified with an extra $RY(\theta)$ gate placed before (magenta line; see Fig. 2.1(a)). IL, RL, GML, and LC stand for initialization, rotation, gate modification, and local Clifford layers respectively. All LC gates are Hadamards. *(b)* VQE results for the modified 2×1 perturbed PC acquired on `ibm_lagos` and `ibm_perth`. The exact GS (E_0) and VQE (E_{VQE}) energies, relative error $(|\Delta E|/E_g)$, and the state fidelity $(\mathcal{F})$ are shown vs. perturbation strength ξ. The exact energies for the first two excited states $(E_1$ and $E_2)$ are also shown for comparison. The relative error is calculated from the energy difference $|\Delta E| = |E_{VQE} - E_0|$, expressed in units of the first[2] energy gap $E_g = E_1 - E_0$. Shaded grey regions indicate statistical uncertainty on the error. All energies are unitless and may be expressed in absolute units (e.g. eV, J, Ha) using a scaling factor with appropriate dimensions.

10

2.3.1 VQE demonstration

For the VQE demonstration[1] of the PC, we consider two vertical plaquettes $(M, N) = (2, 1)$ as a non-trivial example of a gate-modified ansatz (i.e. one that is not a simple repetition of rotation and entanglement layers). It possesses the Hamiltonian given in Eq. (2.4).

We ran the corresponding variational circuit in Fig. 2.3(a; iii) on `ibm_lagos` and `ibm_perth`. The circuit contains 7 qubits and $2 \times 7 + 8 = 22$ parameters from the gate modifications and added $RY + RX$ rotation layers. Since a uniform perturbation strength ξ is employed, the problem symmetry is preserved (App. E) in that qubits on *equivalent* vertices of the graph (i.e. with similarly positioned neighbours and connections; see Sec. 2.3.2) are described by the same θ. This enabled a reduction from 22 to 11 parameters.

The data [shown in Fig. 2.3(b)] exhibit close agreement with the GS energies from exact diagonalization (ED), as confirmed by the small energy differences attained at convergence; $0.091 \leq \Delta E/E_g \leq 0.556$, where $E_g = E_1 - E_0$ is the exact first energy gap.[2] This is also evidenced by the span of the error bars (see App. D) remaining consistently below E_1. Importantly, the expected system behaviour is recovered as we clearly identify a transition region at intermediate ξ (~ 1) where the total energy crosses over from approximating a randomly oriented to completely polarized set of spins. These limiting behaviours correspond to the entangled state $\frac{1}{2}(|0\rangle^{\otimes 7} + |1011010\rangle + |0101101\rangle + |1110111\rangle)$ and separable state $|1\rangle^{\otimes 7}$, respectively (see App. E for more details).

2.3.2 Further analysis and VQE demonstrations

We further analyze the ansatz-modification technique by considering other PC lattices containing multiple plaquettes in both spatial directions. In addition to the 7-qubit $(M, N) = (2, 1)$ case studied on a quantum computer (Sec. 2.3.1), we conduct fake backend simulations[3] for $(M, N) = (1, 2)$ and $(2, 2)$, which require lattices of 7 and 12 qubits

---

[1] Further information regarding the simulation methods and data are presented in Apps. B and C, respectively. Calibration data (error rates, qubit properties, etc.) corresponding to the IBM Quantum systems employed are reported in Ref. [48].

[2] At $\xi = 2$, the first excited state is nearly degenerate with the GS ($E_1 - E_0 = 0.120$). Therefore, for a more apt comparison, we employ the second energy gap $E_2 - E_1$ instead; see App. C. To indicate that a different calculation was employed for $\xi = 2$, we denote the corresponding point with a lighter colour in the plot of $|\Delta E|/E_g$ [Fig. 2.3(b)].

[3] These backends serve to mimic the behavior of IBM Quantum systems, albeit with more stable characteristics and simplified noise models [41]. Any simulations involving such backends are performed entirely on a classical computer.
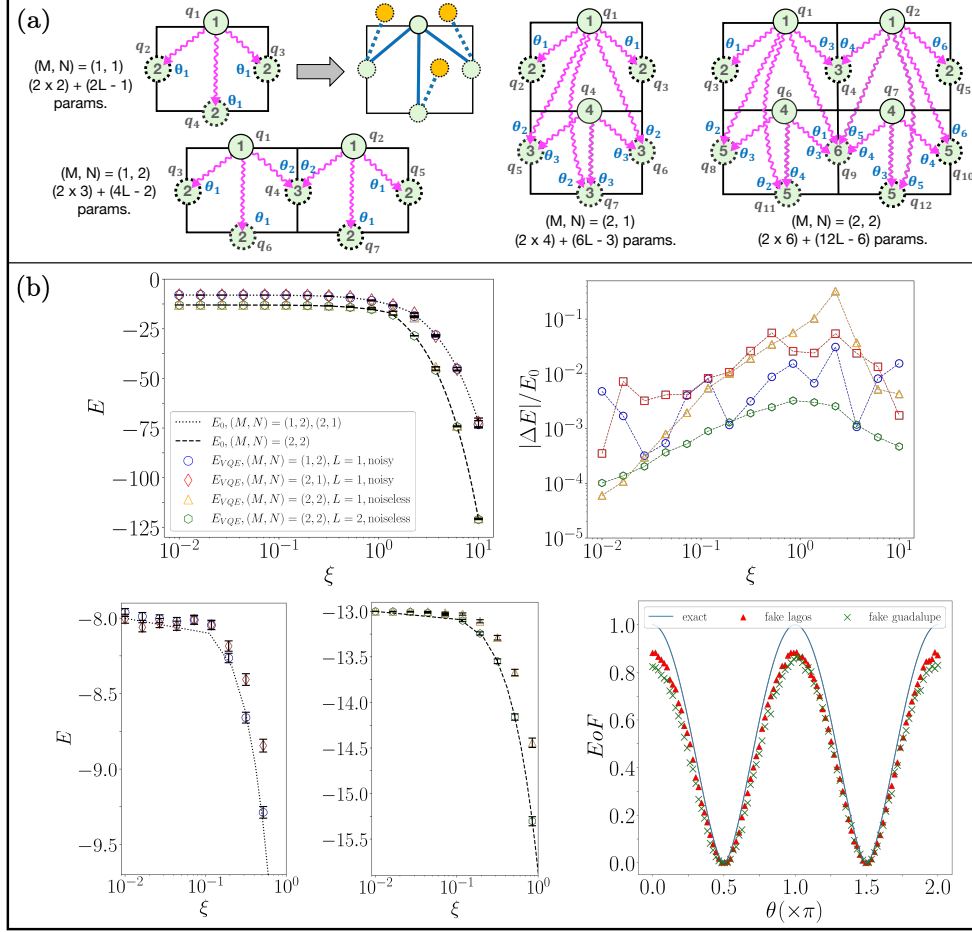
FIG. 2.4: **VQE demonstrations of ansatz modification on fake IBM backends, applied to the perturbed planar code (PC)**. *(a)* The edge-modified graphs used to construct the gate-modified circuits for the VQE, employing the notation in Fig. 2.1. For $(M, N) = (1, 1)$, the modified graph is also depicted as an LC-reduced graph (see Apps. A and F), with the effective edge modifications shown as dotted blue lines and parameterized qubits in orange. The applied symmetry is represented as follows: qubits with the same number inside their circles are parameterized with the same angle (in the rotation layers), while each modified CZ-edge is parameterized with the angle θ shown next to the arrow. The total number of variational parameters is also expressed in terms of the number of modification layers L. *(b)* VQE results for the perturbed PC, with $(M, N) = (1, 2), (2, 1), (2, 2)$, and the gate modification being $RY(\theta)$–CZ. Shown are plots of the EoF for $CZ(RY(\theta))|++\rangle$, E vs. perturbation strength ξ, and $\Delta E/E_0$ vs. ξ. The black dashed lines indicate ED GS energies, and all energies are unitless. The backends used are `fake_lagos` $[(M, N) = (1, 2)$ and $(2, 1)]$, and `fake_guadalupe` $[(M, N) = (2, 2)]$ with 5×10^4 measurement shots. The optimizer used is COBYLA with a maximum of 250 iterations.

12

respectively. Furthermore, applying the model symmetry yields 8 and 18 free parameters respectively. Both characteristics are reflected in the graph state for each scenario [Fig. 2.4(a)]. Their corresponding Hamiltonians are:

$$\hat{\mathcal{H}}_{PC}^{(2,1)} = - \hat{X}_1\hat{X}_2\hat{X}_3\hat{X}_4 - \hat{X}_4\hat{X}_5\hat{X}_6\hat{X}_7 - \hat{Z}_2\hat{Z}_4\hat{Z}_5 - \hat{Z}_3\hat{Z}_4\hat{Z}_6$$
$$... - \hat{Z}_1\hat{Z}_2 - \hat{Z}_1\hat{Z}_3 - \hat{Z}_5\hat{Z}_7 - \hat{Z}_6\hat{Z}_7 + \xi \sum_{i=1}^{7} \hat{Z}_i; \tag{2.4}$$

$$\hat{\mathcal{H}}_{PC}^{(1,2)} = - \hat{X}_1\hat{X}_3\hat{X}_4\hat{X}_6 - \hat{X}_2\hat{X}_4\hat{X}_5\hat{X}_7 - \hat{Z}_1\hat{Z}_2\hat{Z}_4 - \hat{Z}_4\hat{Z}_6\hat{Z}_7$$
$$... - \hat{Z}_1\hat{Z}_3 - \hat{Z}_3\hat{Z}_6 - \hat{Z}_2\hat{Z}_5 - \hat{Z}_5\hat{Z}_7 + \xi \sum_{i=1}^{7} \hat{Z}_i; \tag{2.5}$$

$$\hat{\mathcal{H}}_{PC}^{(2,2)} = - \hat{X}_1\hat{X}_3\hat{X}_4\hat{X}_6 - \hat{X}_2\hat{X}_4\hat{X}_5\hat{X}_7 - \hat{X}_6\hat{X}_8\hat{X}_9\hat{X}_{11} - \hat{X}_7\hat{X}_9\hat{X}_{10}\hat{X}_{12}$$
$$... - \hat{Z}_4\hat{Z}_6\hat{Z}_7\hat{Z}_9 - \hat{Z}_3\hat{Z}_6\hat{Z}_8 - \hat{Z}_5\hat{Z}_7\hat{Z}_{10} - \hat{Z}_1\hat{Z}_2\hat{Z}_4 - \hat{Z}_9\hat{Z}_{11}\hat{Z}_{12}$$
$$... - \hat{Z}_1\hat{Z}_3 - \hat{Z}_2\hat{Z}_5 - \hat{Z}_8\hat{Z}_{11} - \hat{Z}_{10}\hat{Z}_{12} + \xi \sum_{i=1}^{12} \hat{Z}_i, \tag{2.6}$$

which encompass all plaquette, star, and perturbation operators, with numbers corresponding to the qubit labels in Fig. 2.4(a).

For the choice of modification, we continue to employ the $RY(\theta)$–CZ operation used in the $(M, N) = (2, 1)$ case. To verify that it fully modifies the degree of entanglement between connected vertices (qubits), we plot its entanglement of formation (EoF)[4] vs. θ when acting on the input state $|++\rangle$. As shown in Fig. 2.4(b), we observe a periodic function varying between zero and one (near one for fake backends). These values correspond to fully separable and maximally entangled respectively, which confirms the tunable behavior.

We now present the VQE simulation results on `fake_lagos` (noisy) and `fake_guadalupe` (noiseless). Each scenario was optimized using COBYLA (see App. B), with a maximum of 250 iterations and 5×10^4 measurement shots. For $(M, N) = (1, 2)$ and $(2, 1)$, fake backend noise was introduced, while the $(M, N) = (2, 2)$ simulation was performed without noise and utilized $L = 1$ and 2 modification layers. The VQE circuits employed were derived from the connectivity of the graph states in Fig. 2.4(a) (see Fig. 2.3(a) for the $(M, N) = (2, 1)$ circuit). Here, the concept of multiple layers is best understood in the

---

[4]As a general measure of entanglement, the EoF for a bipartite mixed state ρ_{AB} is calculated as $\min[\sum_i p_i E(|\psi_i\rangle)]$, where the $|\psi_i\rangle$ are pure states derived from all possible spectral decompositions of ρ_{AB} (i.e. $\sum_i p_i|\psi_i\rangle\langle\psi_i|$), and $E(|\psi_{AB}\rangle) = -\mathrm{Tr}(\rho_A\log_2\rho_A) = -\mathrm{Tr}(\rho_B\log_2\rho_B)$ is the entanglement entropy for a bipartite pure state $|\psi_{AB}\rangle$ [49].

GB paradigm. Specifically, a circuit of L layers involves concatenating $2L-1$ modification operations, which is needed to avoid gate cancellation when θ is a multiple of 2π.

Our use of extra layers serves two purposes. It not only enables the optimization to access finer corners of the parameter space, but also permits each modification to influence a larger area of the graph's entanglement structure. This is particularly important when adjacent vertical plaquettes are present ($M > 1$), given the nonlocality of its connections [Fig. 2.4(a)]. However, these advantages can be greatly offset by hardware noise because of the increased number of CXs. Thus, it is best to study the effects of multiple layers in the *absence* of noise, as we do here. Furthermore, the model symmetry may be applied separately to each layer, which yields $3+2L$, $4+4L$, $5+6L$, and $6+12L$ total parameters for $(M,N) = (1,1), (1,2), (2,1)$, and $(2,2)$ respectively.

The energy plots in Fig. 2.4(b) for $(M,N) = (1,2)$ and $(2,1)$ exhibit good agreement with the exact GS energies in the presence of fake noise (order of relative errors between 10^{-4} and 10^{-2}), with slightly reduced accuracy in the latter case due to the presence of adjacent vertical plaquettes. However, we observe an improved approximation with $L = 2$ compared to $L = 1$ in the noiseless $(M,N) = (2,2)$ case, which becomes prominent at $\xi \approx 1$. This corresponds to the regime where the perturbation begins to strongly affect the ground state, and therefore requires more variational parameters (i.e. layers) for an accurate approximation. For a further discussion of the perturbative effects, refer to App. E.

We remark that any parameterized unitary operation which attains a continuum of EoF values between zero and one constitutes an effective modification. While this imparts flexibility toward the circuit design, we emphasize that our specific choice of $RY(\theta)$–CZ is motivated by the limiting performance factors on IBM Quantum systems, namely the measurement budget for VQE optimization and CX error rates. These factors are mitigated when the modification contains minimal free parameters and entangling operations, resulting in a less intensive optimization and reduced noise levels respectively. Since qubit connectivity restrictions on IBM systems often incur additional entangling (SWAP) gates, our modification also permits a more feasible extension to multiple layers.

2.4 Protocol implementation: Pauli gadgets

Next, we consider the implementation of hybrid VQEs, in which GB and MB elements are incorporated within the same circuit ansatz. To achieve this, we introduce MB patterns that employ variational measurements on ancillas. Here, the variational parameters $\vec{\theta}$ are encoded in the choice of rotated measurement bases. While there exists a variety of MB techniques [2, 31, 32, 33, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65] that can be adapted for this purpose, we demonstrate the concept with a class of MB patterns known as "Pauli gadgets" ($\hat{\mathcal{P}}^{\otimes n}$). They consist of n vertices (serving as input and output) connected to a single vertex measured in the rotated basis $R(\theta)$ [Fig. 2.1(b)]. Through a series of gadgets and rotation gates, one can readily design a hybrid ansatz circuit that generates multi-qubit Pauli interactions of the form:

$$\bigotimes_{i=1}^{n} \hat{\mathcal{P}}_i; \quad \hat{\mathcal{P}} \in \{\hat{I}, \hat{X}, \hat{Y}, \hat{Z}\}. \tag{2.7}$$

Such interactions are present in many target Hamiltonians encountered in quantum chemistry, high energy- or condensed matter physics. To better understand their realization, we remark that the n gadget vertices implement only Z-interactions (i.e. $\hat{Z}^{\otimes n}$) [66], and thus require additional operations. One straightforward choice is to apply local Clifford $\hat{H}$ ($\hat{S}\hat{H}$) operations on these vertices which enable transformations into $X(Y)$-interactions respectively. [67, 68, 69]. The action of the Pauli gadget then becomes:

$$\hat{\mathcal{P}}^{\otimes n}(\theta) = \left(\bigotimes_{\ell \neq k} \hat{H}_\ell\right)\left(\bigotimes_{k}(\hat{S}^\dagger\hat{H})_k\right)\exp\left[\frac{i\theta}{2}\bigotimes_{j}\hat{Z}_j\right]\left(\bigotimes_{k}(\hat{H}\hat{S})_k\right)\left(\bigotimes_{\ell \neq k}\hat{H}_\ell\right)$$
$$= \exp\left[\frac{i\theta}{2}\left(\bigotimes_{\ell \neq k}\hat{X}_\ell\right)\left(\bigotimes_{k}\hat{Y}_k\right)\left(\bigotimes_{j \neq k \neq \ell}\hat{Z}_j\right)\right], \tag{2.8}$$

where $j, k, \ell \in \{1, 2..., n\}$. In VQEs however, we may use traditional rotation gates to perform these operations. Doing so yields further advantages since their continuous parameterization provides enhanced tunability of the measurement bases (e.g. ability to generate linear combinations of Eq. (2.7)) while contributing extra variational parameters toward the classical optimization. These benefits highlight the synergistic role of rotation gates alongside Pauli gadgets, and showcase the hybrid paradigm nature of the ansatz.

The advantages of employing Pauli gadgets are their efficient resource requirement and versatility in engineering any desired interaction. Compared to the $2n - 2$ entangling operations required by their GB counterparts (deployed in circuit synthesis and T-gate

15

optimization), the Pauli gadget requires only n. This provides a linear depth reduction at the cost of a single ancilla qubit, which becomes increasingly relevant upon increasing the number of qubits, variational layers, or many-body interaction terms. As shown in Figs. 2.6 and 2.7, these hybrid circuits contain $n + g$ qubits, where g is the total number of Pauli gadgets (ancillas) measured mid-circuit. This can be reduced to $n + 1$ qubits if one recycles the first ancilla after its measurement, thus yielding a further resource advantage. We remark that these mid-circuit measurements are inherent to MB-QC, and contrast them with typical readout measurements performed at the end of a circuit (see App. B).

2.5 VQE demonstrations

We now showcase the broad applicability of the Pauli gadgets and their hybrid ansätze via three concrete demonstrations.[1] Here, we remark that our use of MB elements to reduce the number of entangling operations is relevant, as CXs are a bottleneck (i.e. incurring the largest cost and error rates) on the IBM Quantum platform.

2.5.1 Particle physics: QCD in 1D

Particle physics simulations are a promising candidate for achieving quantum advantage [27, 70, 71, 72], and have motivated the study of benchmark models on quantum computers [21, 28, 29, 30, 73, 74]. For our first gadget demonstration, we examine a phase transition phenomena in quantum chromodynamics (QCD). As a component of the Standard Model, QCD describes the strong force between quarks mediated by gluons and is governed by the $SU(3)$ gauge group. Here, we consider QCD on a one-dimensional lattice, where each site contains a quark or antiquark corresponding to a color component of the fermionic field. We remark that a lattice gauge theory formulation is of crucial importance – it provides a non-perturbative tool that enables quantum computers to simulate QCD phenomena that are intractable with standard perturbative techniques [75, 76].

To formulate the Hamiltonian, we follow the procedure of Ref. [77] and cast it into a form suitable for quantum computation. The resulting expression describes a unit cell () possessing three quarks and three antiquarks in the zero-baryon number ($B = 0$) sector, and is a sum of kinetic (hopping), mass, and electric field terms:

$$\hat{\mathcal{H}}_k^{(B=0)} = -\frac{1}{2}(\hat{X}_1\hat{Z}_2\hat{Z}_3 + \hat{Z}_1\hat{X}_2\hat{Z}_3 + \hat{Z}_1\hat{Z}_2\hat{X}_3);$$
$$\hat{\mathcal{H}}_m^{(B=0)} = \tilde{m}(3 - \hat{Z}_1 - \hat{Z}_2 - \hat{Z}_3); \tag{2.9}$$
$$\hat{\mathcal{H}}_e^{(B=0)} = \frac{1}{6x}(3 - \hat{Z}_1\hat{Z}_2 - \hat{Z}_1\hat{Z}_3 - \hat{Z}_2\hat{Z}_3),$$

where the dimensionless numbers $\tilde{m}$ and x are the bare quark mass and coupling constant, respectively. We note the specific form of Eq. (2.9) corresponds to a tetraquark state in the strong coupling limit (Fig. 2.5), with its specific arrangement of quarks and antiquarks corresponding to an elementary block of two adjacent unit cells [77].

We conducted a VQE demonstration on `ibm_perth` and `ibm_peekskill` to determine the ground state of the Hamiltonian in Eq. (2.9), with model parameters $x = 0.8$ and $\tilde{m} \in [-1, 1]$. The ansatz used consists of a single $\hat{\mathcal{P}}^{\otimes 3}$ gadget ($g = 1$) with RY and
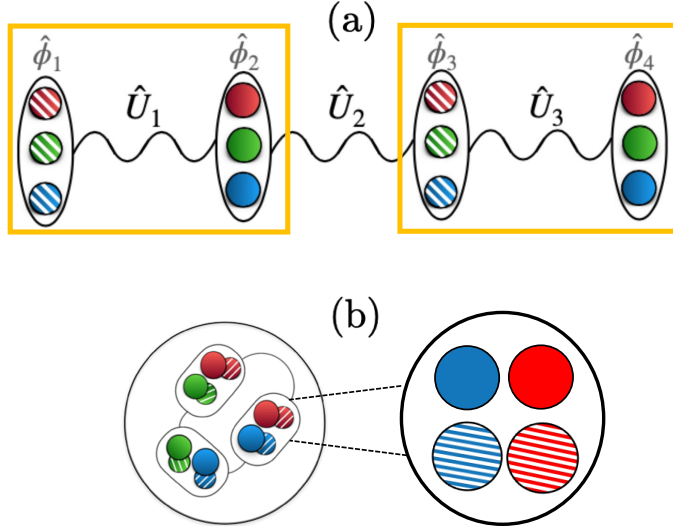
FIG. 2.5: **Discretizing quantum chromodynamics on a one-dimensional lattice [77]**. (a) Each unit cell of the lattice (highlighted in orange) contains up to three quarks (red, green, and blue; represented by solid circles) and three antiquarks (antired, antigreen, and antiblue; represented by stripped circles). The various terms in the Hamiltonian (Eq. 2.9) arise from the physics governing lattice gauge theories. The fermionic matter fields $\hat{\phi}$ (on the lattice sites) correspond to a mass term $\hat{\mathcal{H}}_m$, and are coupled to gauge fields via connecting links $\hat{U}$. The gauge fields employ electric field operators ($\hat{\mathcal{H}}_e$) to impose local charge symmetries on each site, while the links maintain the symmetries and permit matter to "hop" between sites, yielding a kinetic term $\hat{\mathcal{H}}_k$. (b) Depiction of the strongly coupled tetraquark state corresponding to two adjacent unit cells.

RY–RZ gates added before and after, respectively [Fig. 2.6(a)]. We exploit the model symmetry to reduce the number of free parameters from $3 \times 3 + 1 = 10$ to 4. We then calculate the quantity $\langle \hat{N} \rangle = \langle \hat{\mathcal{H}}_m \rangle / \tilde{m}$ for different values of $\tilde{m}$ to observe transitions in the mean occupation number [Fig. 2.6(a)]. For consistency checks, we also plot a second, related number operator $\langle \hat{N}' \rangle = (1/3) \sum_{i<j} (1 - \hat{Z}_i)(1 - \hat{Z}_j)$ for $i, j = 1, 2, 3$.

We observe in Fig. 2.6(a) that in the negative (positive) mass limits ($|x/\tilde{m}| < 1$), $\langle \hat{N} \rangle$ and $\langle \hat{N}' \rangle$ attain their maximal (minimal) occupation numbers, indicating the presence (absence) of quarks and antiquarks on all lattice sites. These correspond to the baryonium (vacuum) state. The state transitions (i.e. intermediate occupation numbers) between the two extremes occur in the region $\tilde{m} \in [-0.2, 0.2]$ [Fig. 2.6(a)], in accordance with Ref. [77]. Fig. 2.6(a) also exhibits good agreement with ED values in both large mass limits $|\tilde{m}| > 0.2$ ($\mathcal{F} \geq 83.4\%$; $|\Delta E|/E_g \leq 9.92 \times 10^{-2}$), and in the transition region $|\tilde{m}| \leq 0.2$ ($\mathcal{F} \geq 77\%$;

$|\Delta E|/E_g \leq 3.11 \times 10^{-1}$). The generally lower fidelity in the latter can be attributed to a smaller energy gap between the ground state and first excited states. However, we note that a single gadget with local rotations suffices to resolve the energy difference while maintaining state fidelities above 75%. These observations highlight the suitability of our Pauli gadgets in VQE problems involving many-body interactions.

2.5.2 Lattice gauge theory: pure $\mathbb{Z}_2$

Next, we analyze a pure (i.e. matterless) LGT possessing a $\mathbb{Z}_2$ symmetry, which describes invariance under spin flips as observed in the transverse-field Ising model [78]. For context, we note that the PC model discussed in Sec. 2.3 corresponds to a $\mathbb{Z}_2$ theory, where the charge configurations and signs of the Hamiltonians are fixed by $\hat{\mathcal{H}}_+$. The $\mathbb{Z}_2$ LGT achieves the same effect by enforcing Gauss' law within its Hilbert space. This allows us to switch to a more general formulation by omitting $\hat{\mathcal{H}}_+$ and the negative signs in Eq. (2.5). We may then express the $\mathbb{Z}_2$ Hamiltonian as a sum of electric and magnetic contributions:

$$\hat{\mathcal{H}}_{\mathbb{Z}_2} = \lambda \hat{\mathcal{H}}_\square + \frac{1}{\lambda} \hat{\mathcal{H}}_\triangle, \tag{2.10}$$

where $\hat{\mathcal{H}}_\square$ and $\hat{\mathcal{H}}_\triangle$ are defined in Eq. (2.5), and λ is the coupling strength [79].

We performed a VQE demonstration on `ibm_peekskill` for a single plaquette [Fig. 2.6(b)], and plot the expectation of the normalized plaquette $\langle \hat{\mathcal{H}}_\square \rangle$[5] and E-field operators $\langle \hat{\mathcal{H}}_\triangle/4 \rangle$ for $\lambda \in [0.5, 3.3]$. Since the only multi-qubit interaction is the plaquette term $X^{\otimes 4}$ in Eq. (2.10), the ansatz contains a $\hat{\mathcal{P}}^{\otimes 4}$ gadget ($g = 1$), with RY and RY–RZ gates added before and after, respectively [Fig. 2.6(b)]. Furthermore, the elimination of redundant parameters in the rotation layers allows us to reduce the number of parameters from $3 \times 4 + 1 = 13$ to 4.

The target plot in Fig. 2.6(b) exhibits a competition between the two operators $\hat{\mathcal{H}}_\triangle$ and $\hat{\mathcal{H}}_\square$. The former (latter) dominates in the small (large) λ regime and is indicative of decoupled spins (ferromagnetic ordering). As expected, there is a crossover around $\lambda = 2$ where the spins begin coupling due to the formation of a circulating B-field around the plaquette. This crossover is generally well captured by the VQE, as evidenced by $\mathcal{F}$ ranging from 76.8% to 86.2% and consistent resolution of the energy difference ($|\Delta E|/E_g \leq 0.573$). While the large-$\lambda$ regime (≥ 2.88) yielded GS energies closer to and beyond the first excited

---

[5]$\langle \hat{\mathcal{H}}_\square \rangle$ and $\langle \hat{\mathcal{H}}_\triangle \rangle$ provide strong indicators of VQE performance since they contribute little to the overall energy in the small and large-λ regimes respectively. This implies that a high accuracy and precision is needed to ensure their agreement with ED.

state ($|\Delta E|/E_g \geq 0.9$), we remark that its near-degeneracy with E_0 ($E_g \leq 0.2349$) and use of a limited resource budget can pose optimization challenges with resolving the energy gap (see App. B for more details).

2.5.3 Quantum chemistry: lithium hydride

Our final demonstration showcases how Pauli gadgets may be used to simulate molecular systems. We determine the ground state of the lithium hydride (LiH) molecule, and follow Ref. [80] to obtain a reduced 4-qubit Hamiltonian consisting of 100 many-body Pauli terms (App. E).

We performed a VQE simulation on `ibm_perth` at an interatomic distance of 1.6 Å. The hybrid ansatz used is similar to that of the $\mathbb{Z}_2$ demonstration, consisting of a $\hat{\mathcal{P}}^{\otimes 4}$ gadget surrounded by RY and RX gates [Fig. 2.6(c)]. Here, the prior elimination of redundant qubits in Ref. [80] implies that we must retain all 9 parameters in the optimization. To highlight the gadget's ability in this more complex scenario, we choose initial parameters far from the GS (corresponding to $E > -7$ Ha), as opposed to the near-optimal settings ($E < -7.7$ Ha) attainable from UCCSD and Hartree-Fock methods [41].

The optimization plots in Fig. 2.6(c) verify the effective performance of the hybrid ansatz, as it converges within 1.1% of $E_0 \approx -7.88$ Ha after only 99 iterations, with $\mathcal{F}$ improving from 6.7 to $\sim 78\%$. We remark that the final error [68] and fidelity not only contains contributions from hardware noise, but also from the limited ability of a single gadget to capture specific two-, three-, and four-body interactions terms in the Hamiltonian. For the latter, we may improve the result by appending extra Pauli gadgets (with local rotations) to the ansatz, as discussed in the next section.

2.5.4 Extension to multiple gadgets

We analyze the efficacy of Pauli gadgets by examining how multiple Pauli gadgets can improve the approximation of the LiH GS energy. Similar to the PC model, we achieve this by layering multiple gadgets in series, dispersing local rotations in between them. To minimize the number of added parameters, we apply a single RY rotation layer between each gadget, with the usual RX layer only after the last gadget [Fig. 2.7(a)]. This yields a total of $5g + 4$ variational parameters. We simulated[3] $g = 1, 2, 3, 4$ $\hat{\mathcal{P}}^{\otimes 4}$ gadgets in a noiseless fashion on `fake_guadalupe` with 512 or 1024 shots and 250 iterations. This normally requires $4 + g$ qubits, however, as mentioned in Sec. 2.4, one may reduce the required qubits to 5 if reset operations are available.

From the progress plots [Fig. 2.7(b)], we observe that a larger g requires more iterations to converge but attains a lower relative error in the GS energy and higher fidelity. This is expected since the greater number of variational parameters enables the ansatz to thoroughly explore greater regions of the Hilbert space. and thus improve the GS approximation. We also note the significant increase in $\mathcal{F}$ from 58.6% ($g = 1$) to 89.3% ($g = 2$). This reflects the sensitivity of the quantity since ρ is constructed by fitting an exponential number of matrix elements (see App. B), and thus incurs larger inaccuracies with relatively few variational parameters and/or shots. With respect to capturing the many-body interactions in Table C.1, we remark that appending identical 4-body gadgets suffices since each is capable of generating k-body interactions for $k < 4$. This can offer a practical advantage as it reduces the variety of gadgets that must be experimentally realized.
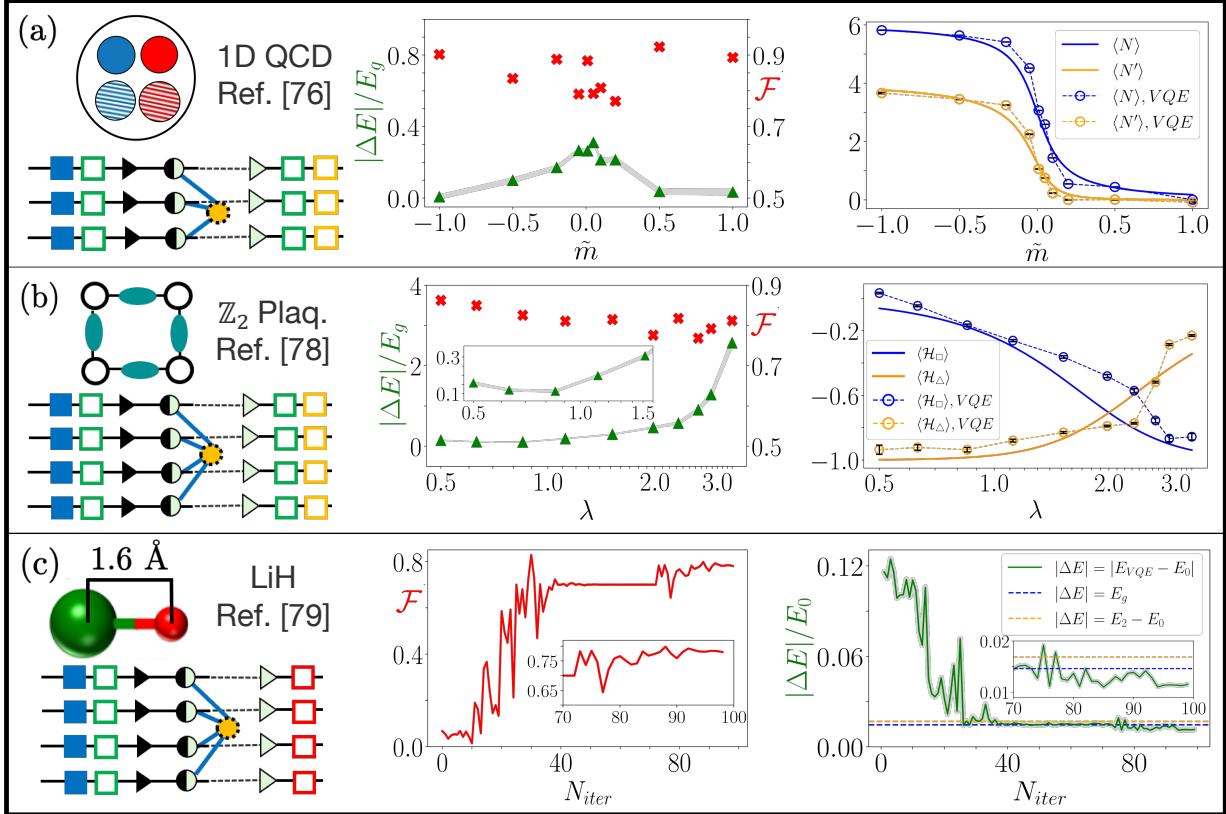
FIG. 2.6: **VQE demonstrations of hybrid ansatz circuits on IBM Quantum systems, involving a single Pauli gadget**. All VQE runs were performed dynamically, i.e. with mid-circuit measurements. All LC gates are Hadamards. The relative errors ($\Delta E/E_g$) are calculated from the energy difference $|\Delta E| = |E_{VQE} - E_0|$, expressed in units of the energy gap $E_g = E_1 - E_0$ or E_0. Shaded grey regions indicate statistical uncertainty on the error. *(a)* 1D QCD model (on `ibm_lagos` and `ibm_peekskill`): $\langle \hat{N} \rangle$ and $\langle \hat{N}' \rangle$ vs. reduced bare quark mass $\tilde{m}$ ($x = 0.8$, $-1 \leq \tilde{m} \leq 1$), with plots of $|\Delta E|/E_g$ and $\mathcal{F}$. *(b)* $\mathbb{Z}_2$ lattice gauge theory model (on `ibm_peekskill`): $\langle \hat{\mathcal{H}}_\square \rangle$ and $\langle \hat{\mathcal{H}}_\triangle \rangle$ vs. coupling strength λ ($0.5 \leq \lambda \leq 3.3$), with plots of $|\Delta E|/E_g$ and $\mathcal{F}$. *(c)* LiH molecule (at interatomic distance = 1.6 Å, on `ibm_perth`): Convergence plots of $|\Delta E|/E_0$ and $\mathcal{F}$ with respect to the VQE iteration number N_{iter}. For comparison, the errors corresponding to excited state energies ($|\Delta E| = E_g$ and $|\Delta E| = E_2 - E_0$) are plotted as horizontal lines.

22
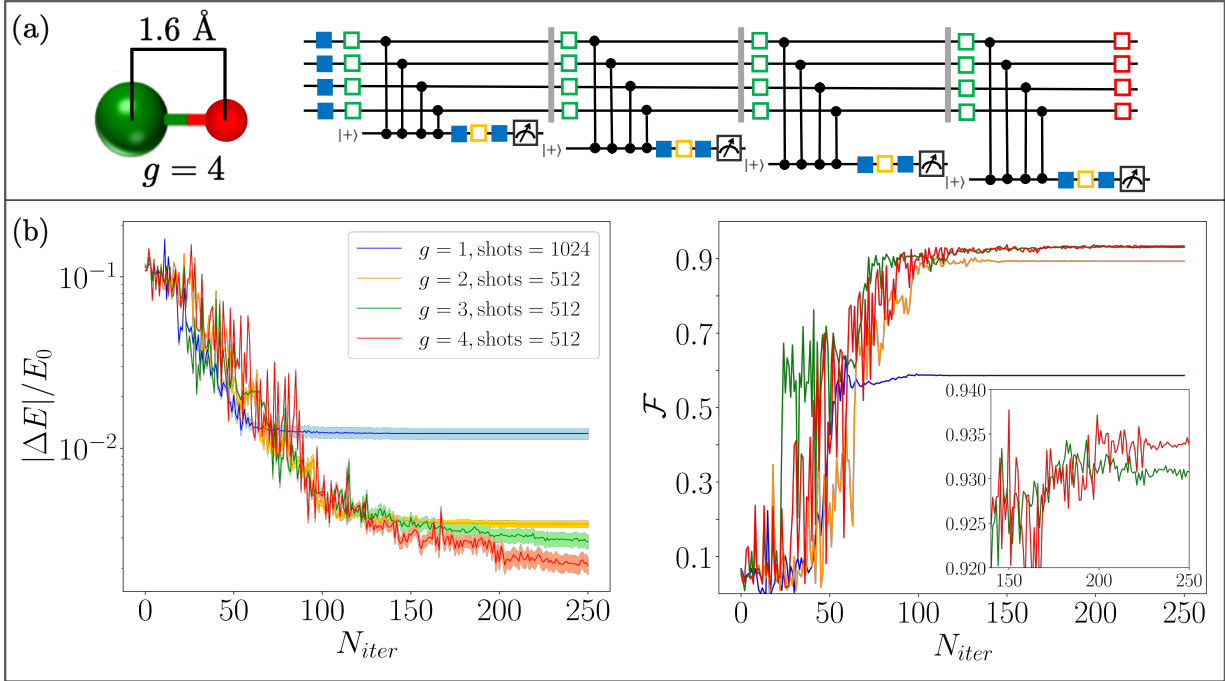
FIG. 2.7: **Applying multiple Pauli gadgets to improve VQE accuracy.** *(a)* Depiction of $g = 4$ $\hat{\mathcal{P}}^{\otimes 4}$ gadgets in a hybrid circuit, with CZ, RZ and H (LC) operations in the MB pattern for each gadget explicitly shown. *(b)* Noiseless VQE simulation results (relative error $\Delta E/E$ and state fidelity $\mathcal{F}$ vs. iteration number) for determining the GS energy of the LiH molecule (at 1.6 Å) with varying numbers of $\hat{\mathcal{P}}^{\otimes 4}$ gadgets ($g = 1, 2, 3, 4$). Lighter color shades around the solid lines indicate the uncertainty in $\Delta E/E_0$. The backend used is `fake_guadalupe` with 512 or 1024 measurement shots. The optimizer employed is COBYLA with a maximum of 250 iterations.

23

2.6   Conclusions

In this chapter, we have detailed a new hybrid approach to performing VQEs which merges the strengths of the GB and MB computing paradigms within a variational circuit (ansatz). While MB approaches are known for having high error thresholds [81, 82, 83, 84] and permit more efficient implementations of certain operations [22, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 85], GB approaches offer greater flexibility, and high fidelity implementations are available on a multitude of physical platforms.

We have introduced our hybrid approach on two levels showcasing how MB aspects can enhance the efficiency of standard GB-VQEs. On the design level, we adapted the MB principle of modifying the entanglement structure of variational graph states to circuits, thus enabling the creation of a problem-informed ansatz in the GB paradigm. On the implementation level, we utilized MB elements known as Pauli gadgets, which employ mid-circuit measurements to efficiently implement multi-qubit interactions. We then expanded the range of possible interactions via local rotation gates, and integrated them alongside the gadgets to form a "hybrid paradigm" circuit ansatz. Bringing the benefits of MB elements to NISQ devices is especially timely in view of mid-circuit measurements becoming available on various platforms [34, 35, 36, 37, 38, 39, 40]. These measurements implement the corrective operations necessary for deterministic MB-QC and avoids the need for inefficient postselection.

Our VQE demonstrations on IBM Quantum systems demonstrated the wide applicability of our hybrid approach in testbed scenarios, and their effectiveness in capturing the underlying physics. These constitute a major step toward practical VQEs, capable of solving hard problems in quantum chemistry, quantum error correction, and lattice gauge theory. We also performed VQE simulations that extrapolated these models toward larger system sizes and circuit depths, and found that they maintained a comparable performance (with improvements present in noiseless runs). These results indicate the potential of our techniques given the availability of more scalable and fault-tolerant devices.

The tools we described in this chapter can be directly applied outside the realm of VQEs towards the broader class of variational quantum algorithms [86], e.g. in quantum neural networks and quantum machine learning [87, 88, 89, 90, 91, 92, 93, 94, 95]. In addition, the Pauli gadget may also be utilized to efficiently simulate real-time, dynamic evolutions of many-body Hamiltonians [73, 77, 96, 97, 98]. Finally, it will be interesting to explore novel readout schemes which use MB elements such as the Pauli gadget to improve the measurement efficiency for many-body observables [19, 68, 99]. Such endeavours may yield new, exciting opportunities to further elucidate the behaviours of fundamental physical systems.

Chapter 3

Superimposing Quantum Operations for Effective Noise Mitigation

In recent years, quantum computing has become a frontrunner in the pursuit of solving classically intractable problems. Despite their burgeoning potential in a variety of mathematical and scientific areas, their present experimental realizations are highly susceptible to noise processes. These stem from apparatus imperfections, interactions with external fields, and system defects [17, 100, 101], all of which result in leakage of the quantum information into the environment. To obtain reliable results, protocols such as error correction [11, 12] and fault-tolerant computation (by virtue of the threshold theorem) [102, 103] are often deployed. However, they all require large ratios between the number of physical qubits and logical qubits they seek to protect. Such an overhead in resources poses challenges for near-term computation, given their limited scalability and ability to maintain control of the qubit interactions.

In this chapter, we present an alternative protocol to the above approaches. Rather than subjecting a target input to a noisy computation, we perform these computations *multiple* times in a coherently controlled superposition. With the help of added resources (i.e. control and auxiliary circuit registers), our approach generates correlations between the auxiliary and input states in such a way that the noise level is effectively reduced. While these involve probabilistic aspects such as measurements and postselection, it is remarkably possible to achieve a deterministic, on-average improvement when combined with unitary optimization. Furthermore, our protocol is capable of reaching full correction in the infinite limit of superpositions or branches. This provides one with the freedom to design circuits with varying degrees of noise mitigation based on the available hardware. More broadly, the protocol is conceptually simple to analyze on the theory and implementation levels, thanks

to how superposition is generated in our approach. This contrasts with previous works, which investigated superposition using purely communication-based approaches involving the quantum switch and space-time trajectories [104, 105].

First, we will discuss the basic mathematical theory underlying the "fidelity enhancement" stemming from our approach. Then, we present numerical demonstrations that showcase different scenarios where effective noise reduction is achieved, including those where the added resources themselves incur noise. In a similar spirit to Chapter 2, we also formulate and perform our protocol in both the GB and MB paradigms. While they yield similar advantages in noise reduction, these advantages stem from different physical processes inherent to their computational procedures, which we highlight throughout our analysis. We emphasize here that all our techniques are geared toward *mitigation* (i.e. partial correction) of the noise as opposed to its complete elimination. This is what enables the required computational resources to be lessened, thereby providing a suitable approach for current NISQ devices.
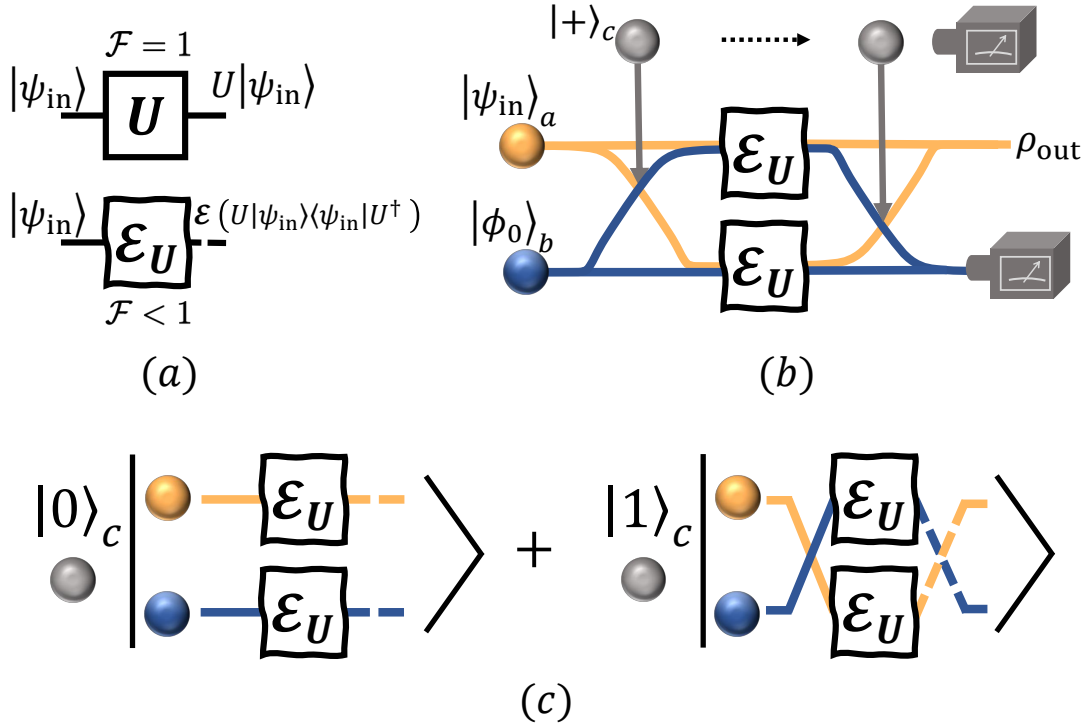
FIG. 3.1: **Conceptual illustration of the fidelity enhancement protocol, applied to an arbitrary computation U acting on m input qubits**. *(a)* The fidelity $\mathcal{F}$ is calculated with respect to the output of the ideal (noiseless) computation. In the ideal case, the absence of error guarantees unit fidelity, i.e. $\mathcal{F} = 1$. When considering realistic (i.e experimental) settings, we must consider noise processes acting on the system. In this *incoherent* case, the fidelity is generally reduced, i.e. $\mathcal{F} < 1$. *(b)–(c)* Schematic of the underlying idea behind the protocol. By employing a control register c initialized in the srtate $|+\rangle$, one can generate a coherent superposition of two or more identical noisy computations (between a target input and auxiliary registers). In this *coherent* case, correlations are generated between the registers such that it is possible for the noises to interfere destructively, thereby improving $\mathcal{F}$.

3.1 Background

Before introducing the protocol, we describe relevant concepts for understanding noise processes and their influence on a quantum computation. These include the mathematical formalism used to describe noise models, and metrics to quantity the performance of a noisy computation.

3.1.1 Steinspring, Kraus and process matrix representation

Noisy processes such as decoherence are captured by the Steinspring representation [1, 106], which characterizes the dynamics of *open* quantum systems. These systems comprise a target system interacting with an external *environment*, which corrupts the desired result of a computation. This interaction may be regarded as a time-evolution of the target (ρ_{in}^t) and environment ($\rho_{in}^e = |e_0\rangle\langle e_0|$) states, where the latter is often unknown. Together, they form the initial input state $\rho_{in} = \rho_{in}^t \otimes \rho_{in}^e$, which is assumed to be pure and separable. The evolution then proceeds as a joint unitary transformation U_{te} on both systems, yielding the output $\rho_{out} = U_{te}\rho_{in}U_{te}^\dagger$. To obtain the target output, we trace out the environmental system to obtain:

$$
\begin{aligned}
\rho_{out}^t &= \mathrm{Tr}_e\left[U_{te}(\rho_{in}^t \otimes |e_0\rangle\langle e_0|)U_{te}^\dagger\right] \\
&= \sum_i \langle e_i|U_{te}(\rho_{in}^t \otimes |e_0\rangle\langle e_0|)U_{te}^\dagger|e_i\rangle \\
&= \sum_i K_i(U_t\rho_{in}^t U_t^\dagger)K_i^\dagger,
\end{aligned}
\tag{3.1}
$$

where Tr_e is the partial trace over the environment, $|e_i\rangle$ are the environmental basis states, and K_i are known as Kraus operators. The latter describes the influence of the environment and obeys the completeness relation $\sum_i K_i^\dagger K_i = I$. Eq. (3.1) is important as it arises from the action of a quantum channel $\mathcal{E}_{U_t}$ (a completely positive and trace-preserving map) on ρ_{in}^t, so that $\mathcal{E}_{U_t}(\rho_{in}) = \rho_{out}^t$. This exemplifies the notion that quantum channels are a unitary evolution in a larger, purified Hilbert space, where their Kraus operators are associated with the environmental basis states,

$$
|\psi\rangle_t|e_0\rangle_e \mapsto \sum_j U_{te}K_j\,|\psi\rangle_t \otimes |j\rangle_e.
\tag{3.2}
$$

Here, the leakage of information into the environment (e) enables a characterization of the noise, and the partial trace over the environment allows one to obtain a description solely

in terms of Kraus operators.

We may also decompose any Kraus operator into a weighted sum of Pauli operators, as they themselves obey the completeness relation:

$$K_i = \sum_j c_{i,j} \mathcal{P}_j, \tag{3.3}$$

where $\mathcal{P} \in \{I, X, Y, Z\}$ and $\sum_j |c_{i,j}|^2 = 1$. This allows us to re-express Eq. (3.1) as:

$$\rho_{\text{out}}^{\text{t}} = \sum_{a,b} \chi_{ab} \mathcal{P}_a \left(U_{\text{t}} \rho_{\text{in}}^{\text{t}} U_{\text{t}}^{\dagger}\right) \mathcal{P}_b^{\dagger}, \tag{3.4}$$

where the coefficients $\chi_{ab} = \sum_i c_{ia} c_{ib}^*$ and $\mathcal{P} \in \{I, X, Y, Z\}^{\otimes m}$ is an m-qubit Pauli string (for an m-qubit input). This is known as the canonical or *process matrix representation* [1], which facilitates the description of noise types naturally described in the Pauli basis. It is particularly relevant as we focus on depolarizing and dephasing channels in our analysis. Their corresponding Kraus operators are given by:

$$\text{Dephasing: } K_0 = \sqrt{p_0} I, \ K_1 = \sqrt{1 - p_0} Z, \tag{3.5a}$$

$$\text{Depolarizing: } K_0 = \sqrt{p_0} I, \ K_i = \sqrt{\frac{1 - p_0}{3}} \mathcal{P}_i, \tag{3.5b}$$

where $\mathcal{P}_i = X, Y, Z$ for $i = 1, 2, 3$ and p_0 denotes the probability of incurring no error. In this representation, we may conveniently characterize such channels by examining their process matrices χ (with normalized elements χ_{ab}). For example, the $m = 2$ matrices for dephasing and depolarizing channels[1] respectively are:

$$\chi = diag(p_0^2, 0, 0, p_0(1 - p_0), 0, 0, 0, 0, 0, 0, 0, 0, (1 - p_0)p_0, 0, 0, (1 - p_0)^2); \tag{3.6a}$$

$$\begin{aligned} \chi = diag(&p_0^2, p_0(1 - p_0)/3, p_0(1 - p_0)/3, p_0(1 - p_0)/3, (1 - p_0)p_0/3, (1 - p_0)^2/9, \\ &(1 - p_0)^2/9, (1 - p_0)^2/9, (1 - p_0)p_0/3, (1 - p_0)^2/9, (1 - p_0)^2/9, (1 - p_0)^2/9, \\ &(1 - p_0)p_0/3, (1 - p_0)^2/9, (1 - p_0)^2/9, (1 - p_0)^2/9), \end{aligned} \tag{3.6b}$$

with no-error probability $p_{\text{ne}} = \chi_{00} = p_0^2$.

---

[1] In cases where $m \geq 2$, we compose these channels by considering tensor products of all m-operator permutations of K_i. For instance, the $m = 2$ dephasing channel has Kraus operators: $K_0 = \sqrt{p_0^2}(I \otimes I)$, $K_1 = \sqrt{p_0(1 - p_0)}(I \otimes Z)$, $K_2 = \sqrt{(1 - p_0)p_0}(Z \otimes I)$, and $K_3 = \sqrt{(1 - p_0)^2}(Z \otimes Z)$.

3.1.2 Principle for fidelity enhancement

Given a noisy computation, it is useful to quantify its detrimental effects by computing the state fidelity – a measure of how "close" two states are. We consider as our reference the noiseless situation, where only the desired computation acts on the target system $\rho_{\text{in}}^{\text{t}}$. Denoting the subscript "0" to mean the absence of noise, we express the output as $\rho_0^{\text{t}} = U_{\text{t}}\rho_{\text{in}}^{\text{t}}U_{\text{t}}^{\dagger}$. In the noisy case, the output $\rho_{\text{out}}^{\text{t}}$ has the form of Eq. (3.1). We then calculate the state fidelity $\mathcal{F}$ between these two states as:

$$\mathcal{F} = \text{Tr}\left\{ (\sqrt{\rho_0^{\text{t}}}\rho_{\text{out}}^{\text{t}}\sqrt{\rho_0^{\text{t}}})^{\frac{1}{2}} \right\}^2, \tag{3.7}$$

which equals unity when the computation is unaffected by the noise (i.e. $\rho_{\text{out}}^{\text{t}} = \rho_0^{\text{t}}$). If the input is pure, then $\rho_{\text{in}}^{\text{t}} = |\psi_{\text{in}}^{\text{t}}\rangle\langle\psi_{\text{in}}^{\text{t}}|$ and $\rho_0^{\text{t}} = U_{\text{t}}|\psi_{\text{in}}^{\text{t}}\rangle\langle\psi_{\text{in}}^{\text{t}}|U_{\text{t}}^{\dagger}$, allowing us to simplify Eq. (3.7) to:

$$\mathcal{F} = \langle\psi_{\text{out}}^{\text{t}}|\rho_{\text{out}}^{\text{t}}|\psi_{\text{out}}^{\text{t}}\rangle. \tag{3.8}$$

However, for a pure $|\psi_{\text{in}}\rangle$, a noisy computation will generally transform the state into a mixed one, i.e. ρ_{out}. Thus, it is more apt to define the fidelity as:

$$\mathcal{F}_{\text{ic}} = \text{Tr}(\rho_{\text{out}}U|\psi_{\text{in}}\rangle\langle\psi_{\text{in}}|U^{\dagger}). \tag{3.9}$$

We refer to $\mathcal{F}_{\text{ic}}$ as the *incoherent* fidelity, which is the quantity we aim to improve on with our coherent superposition protocol.

We now describe the fidelity enhancement principle underlying our protocols, using the simplest example of two superposition branches in Fig. 3.1. The superposition is generated by an auxiliary control register c, which we set to $|+\rangle$ to maximize its effect. We then apply $\mathcal{E}_U$ – a noisy version of the desired computation U – identically to both branches. Due to the coherent superposition, the noises become correlated and interfere with each other. Since this interference can be destructive in nature, one may achieve partial cancellation (reduction) of the noise[2]. We then collapse the superposition by measuring the control in the X-basis, which leads to two outcomes (assuming the control state remains fixed) and their associated output states. Here, we let their state fidelities be $\mathcal{F}_1$ and $\mathcal{F}_2$, with probabilities p_1 and p_2. Our protocol is then capable of yielding probabilistic

---

[2]It is interesting to note the similarities between coherently superimposing noise, and the use of coherent states in quantum metrology to reduce shot noise toward the Heisenberg limit. In particular, both draw benefits from the assumption of *uncorrelated* noise. In our protocol, it allows fidelity improvements to be more readily obtained and understood, while in quantum metrology, it enables a desired sampling variance (precision) to be achieved with a fewer number of measurement shots [107, 108].

(i.e. $\max(\mathcal{F}_1, \mathcal{F}_2) > \mathcal{F}_{\text{ic}}$) and/or deterministic (i.e. $p_1\mathcal{F}_1 + p_2\mathcal{F}_2 > \mathcal{F}_{\text{ic}}$) fidelity improvements. We remark that a generalization to d branches ($d > 2$) is achievable by employing a multi-level (qudit) control system where $|+\rangle_c = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |i\rangle$.

3.1.3 Figures of merit

Despite the usefulness of calculating state fidelity [Eq. (3.7)], it can vary depending on the input state chosen. This makes it difficult to describe as an indicator of performance, and as such, we turn to more specific figures of merit to quantify our advantages. Here, we employ the Choi-Jamiołkowski (CJ) isomorphism, which directly relates quantum states with channels [109]. At the center of the description is the $2m$-qubit Bell state:

$$|\Phi_m^+\rangle_{\text{a,b}} = \left(\frac{|0\rangle_{\text{a}}|0\rangle_{\text{b}} + |1\rangle_{\text{a}}|1\rangle_{\text{b}}}{\sqrt{2}} \right)^{\otimes m}, \tag{3.10}$$

where a and b refer to the two halves of the state, corresponding to two m-qubit systems commonly named "Alice" and "Bob". We subject Bob's half of the state to a noisy computation, while Alice's half remains unaffected. If we set $|\psi_{\text{in}}\rangle = |\Phi_m^+\rangle$[3], one may define a fidelity-type quantity known as the Choi-Jamiołkowski (CJ) fidelity:

$$\mathcal{F}_{\text{CJ}} = \langle\Phi_m^+|_{\text{a,b}}(I \otimes U^\dagger)\rho_{\text{out}}^{\text{a,b}}(I \otimes U)|\Phi_m^+\rangle_{\text{a,b}}, \tag{3.11}$$

where $\rho_{\text{out}}^{\text{a,b}}$ is the joint output state after the noisy computation, and U acts on Bob's (the target) system. The CJ fidelity is notable for imposing a lower bound on the achievable fidelity $\mathcal{F}$ over all arbitrary input states. This remarkable property arises from the maximally entangled nature of Eq. (3.10), and its relation to teleportation. The presence of noise serves to decohere and destroy the entanglement, which lowers $\mathcal{F}_{\text{CJ}}$ and the degree the transmitted state resembles $(I \otimes U)|\Phi_m^+\rangle_{\text{a,b}}$.

Throughout our numerical analyses (Secs. 3.2.5 and 3.3.2), we quantify the degree of advantage as the CJ-infidelity ratio $\mathcal{R}$ between the incoherent and coherent cases:

$$\mathcal{R} = \frac{1 - \mathcal{F}_{\text{CJ, ic}}}{1 - \mathcal{F}_{\text{CJ}}}, \tag{3.12}$$

such that the greater $\mathcal{R}$ is above 1, the greater the noise reduction, with $\mathcal{R} = 1$ indicating no reduction (i.e. same as the incoherent case).

---

[3]Note that $|\Phi_m^+\rangle$ requires double the number of input registers ($2m$), and conventionally, Alice's m registers are placed above Bob's m registers. After state initialization, Alice's registers remain idle throughout the computation.
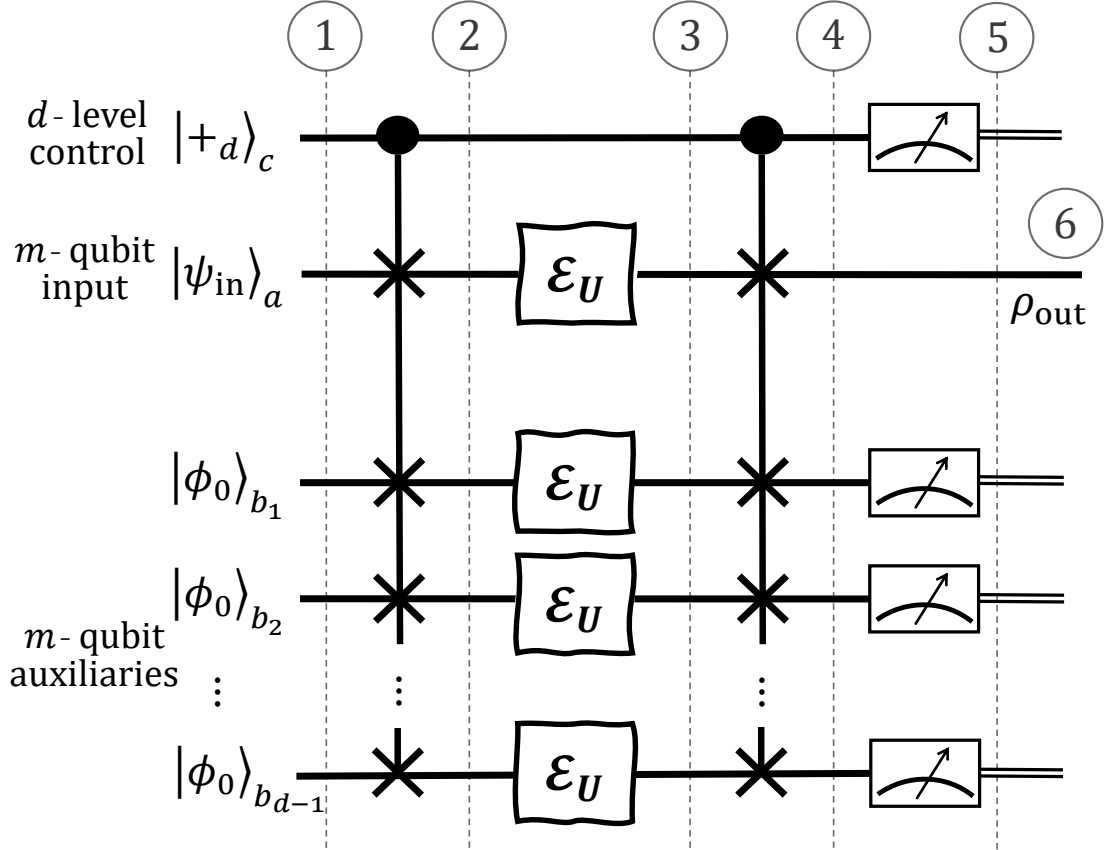
FIG. 3.2: **Schematic representation of the GB protocol for fidelity enhancement**. Vertical dashed lines, marked with numbers 1–6, identify the system's state after the corresponding steps of the protocol (refer to Ref. [65]). The initial $|\psi_{\mathrm{in}}\rangle_a$ and auxiliary $|\phi_0\rangle_{b_i}$ $(i = 1, \ldots, d-1)$ states, as well as the noisy computations $\mathcal{E}_U$ are explicity shown. The multi-qubit gates between lines 1 and 2, as well as 3 and 4, are cSWAP operations defined by Eq. (3.14).

3.2 Enhanced gate-based computation

With the background in place, we now present the protocol for performing coherent computations, starting with the original gate-based (GB) version of the protocol. We explain the steps from a high-level perspective, and refer to Ref. [65] for more details regarding the derivation and notations used.

3.2.1 Working protocol

The GB protocol entails performing the circuit shown in Fig. 3.2. It consists of a control system $|+_d\rangle_c$, an m-qubit register for the input state $|\psi_{\text{in}}\rangle_t$, and $d-1$ identical registers of m-qubit auxiliary states $|\phi_0\rangle_i$, where $i = 1, ..., d-1$.

Step 1: First, we prepare all systems mentioned above. The control register is initialized to $|+\rangle_c = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |i\rangle$, with d indicating the number of branches (superpositions) used to perform the desired computation U. If only qubits are available, one may implement the control register using $n > \log_2 d$ qubits with each in $|+\rangle$. The auxiliary states $|\phi_0\rangle_i$ may be chosen freely (i.e. as separable, entangled, or product of entangled states) and plays a direct role in the degree of noise reduction obtained. The initial state is then:

$$|+_d\rangle_c|\psi_{\text{in}}\rangle_t \bigotimes_{i=1}^{d-1} |\phi_0\rangle_i, \tag{3.13}$$

Step 2: Next, we apply a generalized d-level version of the cSWAP (Fredkin) gate, which acts according to:

$$\text{cSWAP} = |0\rangle\langle 0|_c \otimes I + \sum_{i=1}^{d-1} |i\rangle\langle i|_c \otimes \text{SWAP}_{t,i}, \tag{3.14}$$

where,

$$\text{SWAP}_{t,i} = \bigotimes_{j=1}^{m} \left(|\psi_{\text{in}}^{(j)}\rangle\langle\phi_0^{(j)}|_t \right) \left(|\phi_0^{(j)}\rangle\langle\psi_{\text{in}}^{(j)}|_i \right). \tag{3.15}$$

The cSWAP gate swaps the input $|\psi_{\text{in}}\rangle$ (i.e. Bob's part of the Bell state) with every auxiliary register, effectively dispersing it across all branches. In particular, the choice of which states to swap is dictated by the control state $|i\rangle$ for $i = 1, 2, ..., d$. For $i = 0$, no swap occurs and corresponds to the incoherent case. We note that in multi-qubit cases

33

$(m > 1)$, a SWAP exchanges the jth qubit of states $|\psi_{\text{in}}\rangle_{\text{t}}$ and $|\phi_0\rangle_i$ for $j = \{1, 2, .., m\}$, as described by Eq. (3.15).

Step 3: The computation of interest U is now applied to the input and all auxiliary registers. In the GB paradigm, we model noise as arising from the imperfect application of U given by Eq. (3.1), which corresponds to the map $\mathcal{E}_U$. Here, we assume that the noise affecting each register is uncorrelated (independent), which ensures they may be cancelled out, and that they are identical.

Step 4: We then apply a second cSWAP operation, which serves to undo all the register swaps in Step 2. In the ideal (noiseless) scenario, it completely cancels with the first cSWAP, leaving the overall state as U applied to the inputs and auxiliary systems in Eq. (3.13). However, in a noisy scenario, the overall state remains entangled among its registers. As described in the next step, it is this correlation that allows the different noise types to interfere, resulting in their partial or complete cancellation.

Step 5: The control and auxiliary systems are measured to collapse the overall super-position, yielding the output state ρ_{out}. This step is crucial since we use it to showcase its higher fidelity over the incoherent case [Eq. (3.9)]. We measure the control system in the generalized X- or Fourier basis, which is spanned by the set of states:

$$\left\{ \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} e^{\frac{2\pi ikl}{d}} |k\rangle_c \right\}_{l=0}^{d-1}. \tag{3.16}$$

The rationale is that any ideal result will involve the control in $|+\rangle$, and consequently, our choice of basis allows us to distinguish this state and extract the best possible advantage. Choosing a basis for the auxiliary states is less straightforward, and we discuss a systematic approach in Sec. 3.2.2. Assuming the control and auxiliaries are measured in the states $|+\rangle$ and $|\phi_f\rangle$ respectively, the output state is [65]:

$$\rho_{\text{out}} = \frac{\mathcal{A}_d}{d} \left[\sum_i K_i U |\psi_{\text{in}}\rangle \langle \psi_{\text{in}}| U^\dagger K_i^\dagger \right.$$

$$\left. ... + (d-1) \sum_{i,j} \left(\frac{\langle \phi_f | K_j U |\phi_0\rangle \langle \phi_0| U^\dagger K_i^\dagger |\phi_f\rangle}{\mathcal{A}_2} \right) K_i U |\psi_{\text{in}}\rangle \langle \psi_{\text{in}}| U^\dagger K_j^\dagger \right]; \tag{3.17}$$

$$\mathcal{A}_d = \left(\sum_i |\langle \phi_f | K_i U |\phi_0\rangle|^2 \right)^{d-1},$$

which is a sum of an incoherent term [$\propto$ Eq. (3.1)] and a coherent term that corresponds

34

to the input distributed over the auxiliary registers (i.e. from Step 2). We focus on the latter term, which contains the product:

$$\langle \phi_f | K_j U | \phi_0 \rangle \langle \phi_0 | U^\dagger K_i^\dagger | \phi_f \rangle. \tag{3.18}$$

Eq. (3.18) represents the interference between noise types, as evidenced by the cross-terms involving the Kraus operators K_i and K_j. Here, $|\phi_f\rangle$ refers to the chosen state that the auxiliary register is projected onto. We observe that the greater $U|\phi_0\rangle$ is susceptible to the noise, such that it more closely resembles $|\phi_f\rangle$, the greater the interference and extent of noise reduction. Although counter-intuitive, we may relate this property to our act of measuring the input and auxiliary systems, which "reveals" the nature of the noise acting on them. This knowledge is required for interference to occur, and is strengthened the more noise impacts the systems responsible for the interference (i.e. the auxiliaries).

Naturally, we wish to maximize Eq. (3.18) – which we achieve by defining two performance quantities ω_1 and ω_2:

$$\omega_1 = 1 - \frac{\sum_{j \geq 1} \left| \langle \phi_0 | U^\dagger K_j U | \phi_0 \rangle \right|^2}{1 - p_{\text{ne}}}; \tag{3.19a}$$

$$\omega_2 = |\langle \phi_f | U | \phi_0 \rangle|^2, \tag{3.19b}$$

where $p_{\text{ne}} = p_0^m$ is the probability of incurring no error. Both ω_1 and ω_2 range between 0 and 1, with the former due to $\sum_i K_i^\dagger K_i = I$. Considering $(\omega_1, \omega_2) = (1, 1)$, we find that Eq. (3.18) is indeed maximized by imposing that $|\phi_f\rangle = U|\phi_0\rangle$ and the only surviving Kraus term be $i = j = 0$, which corresponds to an ideal computation (since $K_0 = I$). The coherent term becomes $p_{\text{ne}}^d U |\psi\rangle_{\text{in}} \langle \psi | U^\dagger$ while the incoherent term (containing all K_i with $i \geq 1$) becomes proportional to p_{ne}^{d-1}/d (see Eq. (27) of Ref. [65]). We observe that in the asymptotic limit of branches ($d \to \infty$), one can arbitrarily minimize the incoherent term, with $\mathcal{F}_{\text{CJ}}$ tending to one. We also see that the extremes $(\omega_1, \omega_2) = (1, 1)$ and $(0, 0)$ correspond to maximum and no noise reduction respectively, with intermediate values indicating partial reduction.

Step 6: Finally, we perform a postprocessing step on the results in one of two ways. We either postselect a *single* outcome corresponding to the highest fidelity, or retain all or part of the outcomes and perform a unitary optimization to improve the on-average fidelity. We discuss these approaches further in the next sections.

3.2.2 Probabilistic approach

The idea behind postselection revolves around the desire to attain maximal fidelity. It involves retaining a single outcome over the control and auxiliary registers corresponding to the state:

$$|+_d\rangle_c \bigotimes_{i=1}^{d-1} |\phi_f\rangle_i. \tag{3.20}$$

which corresponds to the output in Eq. (3.17). From Sec. 3.2.1, it is ideal to set $|\phi_f\rangle = U|\phi_0\rangle$ which allows $\omega_2 = 1$. However, from an experimental standpoint, our initialization of each auxiliary in $|\phi_0\rangle$ requires an application of $U^\dagger$ to undo the computation. This can introduce further noise that lowers the fidelity improvement. We can circumvent this with a more generic choice for $|\phi_f\rangle$, one that is completely sensitive to *arbitrary* noises. More specifically, setting $|\phi_0\rangle$ to the m-qubit Bell state $|\Phi_m^+\rangle$[4] [Eq. (3.10)] fulfils this condition since:

$$\langle\Phi_m^+|U^\dagger K_i U|\Phi_m^+\rangle = \sqrt{p_0^m}\delta_{i,0} \tag{3.21}$$

for $i = 0, 1, ..., 4^m - 1$ (i.e. all m-qubit Paulis), where $K_0 = \sqrt{p_0^m}I^{\otimes m}$. Since this implies $\langle\phi_f|U|\Phi_m^+\rangle = 1$, we are ensured that $(\omega_1, \omega_2) = (1, 1)$.[5] Furthermore, by expressing arbitrary noise in the form of Eq. (3.4) with process matrix χ, one can prove (see Ref. [65]) that:

$$\mathcal{F}_{\text{CJ}} = \chi'_{00} > \chi_{00} = \mathcal{F}_{\text{CJ, ic}}, \tag{3.22}$$

where $\chi'_{00} = \chi_{00}^d/p_s$, and p_s is the success probability of obtaining Eq. (3.20):

$$p_s = \frac{\chi_{00}^{d-1}}{d} + \frac{d-1}{d}\chi_{00}^{d-2}\left(\sum_i |\chi_{i0}|^2\right). \tag{3.23}$$

Eqs. (3.22) and (3.23) highlight two important features of the protocol: firstly, $\mathcal{F}_{\text{CJ}}$ increases asymptotically toward 1 (its maximum) as $d \to \infty$, and secondly, such behaviour is independent of the specific noise acting on the computation.

Despite these general advantages, Bell states are inherently challenging to prepare ex-

---

[4]Here, we may treat the auxiliary registers similarly to the input (i.e. each is augmented with an idle register). However, we note if $m = 2$ and $d \geq 2$, one may encode $|\Phi_1^+\rangle$ directly in each auxiliary to avoid incurring extra resources. Alternatively, one may resort to maximally entangled GHZ states $\frac{|0\rangle^{\otimes m} + |1\rangle^{\otimes m}}{\sqrt{2}}$ both separately and in conjunction with Bell states to achieve comparable effects. For simplicity, we do not consider such scenarios here and refer to Ref. [65] for more details.

[5]Under this condition, the state in Eq. (3.20) reduces to the ideal (noiseless) output. One can then show that projection onto this state yields the *maximum* possible $\mathcal{F}_{CJ}$ over all possible outputs; see Ref. [65].

perimentally and require extra resources ($2m$ qubits for each $|\phi_0\rangle_i$). However, it turns out for noises described by rank-2 matrices such as dephasing, we may attain $w_1 = 1$ with non-entangled $|\phi_0\rangle_i$. In this example, since Z is the only non-identity Kraus, setting each $|\phi_0\rangle_i$ to either eigenstate of X or Y guarantees maximal sensitivity. As for $w_2 = 1$, we must possess individual knowledge of the unitary U and its action on $|\phi_0\rangle$, to avoid acting $U^\dagger$. From there, we may select the auxiliary such that it is an eigenstate of U (i.e. $U|\phi_0\rangle = |\phi_0\rangle$). For example, if U is an RZ-gate, then it is appropriate to set $|\phi_0\rangle = |0\rangle$ or $|1\rangle$. Note that if U is Clifford (i.e. composed of H, S and CX gates), then we may initialize $|\phi_0\rangle$ and determine $U|\phi_0\rangle$ classically efficiently, and such knowledge is not required. In the numerical analysis (Sec. 3.2.5), we take into account such considerations. We emphasize, however, that the benefit of this approach lies in its plug-and-play nature – one need not know anything about the unitary or noise type, and fidelity improvement can still be obtained provide both ω_1 and ω_2 are effectively > 0.

3.2.3 ϵ–deterministic approach

While the probabilistic approach is capable of mitigating a noisy computation, it relies on a high p_s (i.e. probability of measuring the postselected outcome) to be effective. However, there exist scenarios where p_s is low, such as when d is relatively large or when ω_1 and ω_2 are closer to unity. In these scenarios, additional outcomes must be retained in order to increase p_s. Although this often lowers the (now) on-average fidelity $\bar{\mathcal{F}}_{\mathrm{CJ}}$, we may boost this quantity by performing a prior optimization step in which correcting unitaries are applied to each output state. If we label the possible outcomes as $q = 1, 2, ..., 2^{m(d-1)}d$, each corresponding to an output $\rho_{\mathrm{out}}^{(q)}$ with probability $p_s^{(q)}$, then the optimization seeks to find a unitary $U_c^{(q)}$ that improves $\mathcal{F}_{\mathrm{CJ}}^{(q)}$ as much as possible to $\mathcal{F}_{\mathrm{CJ, opt}}^{(q)}$ (corresponding to $U_c^{(q)}\rho_{\mathrm{out}}^{(q)}U_c^{(q)\dagger}$). Moreover, a user may specify a desired failure threshold ϵ which dictates the number of outcomes $r \leq 2^{m(d-1)}d$ retained such that $p_s = \sum_{q=1}^{r} p_s^{(q)} \geq 1-\epsilon$. This forms the basis for the ϵ-*deterministic* approach, which yields the improved on-average fidelity:

$$\bar{\mathcal{F}}_{\mathrm{CJ, opt}} = \frac{1}{p_s} \sum_{q=1}^{r} p_s^{(q)} \mathcal{F}_{\mathrm{CJ, opt}}^{(q)} . \tag{3.24}$$

Here, we perform the optimization as a minimization problem for each outcome, where the cost function is the infidelity $1 - \mathcal{F}_{\mathrm{CJ}}^{(q)}$. From this, we show in Sec. 3.2.5 that it is possible to attain advantages even in situations where all outcomes are kept (i.e. $\epsilon = 0$).

In principle, the optimization may incorporate free parameters to vary in pursuit of finding the best corrections and fidelities (akin to a VQE). These may be used to optimize

other relevant aspects of the protocol, such as the initial auxiliary states $|\phi_0\rangle_i$ and measurement bases. However, we note the optimization challenges encountered when the number of branches (d) or inputs (m) is large (e.g. high problem dimensionality), and the need for a prior extensive characterization of the noise (e.g. via process tomography) in experimental settings. To avoid additional (hidden) resources, we employ local, single-qubit correcting unitaries in our optimizations, and restrict them to discrete gates in the Clifford group, which incur minimal noise due to their efficient classical implementation.

3.2.4 Protocol example

To illuminate the ideas behind the GB protocol, we apply it to a basic scenario involving three branches $(d = 3)$ and dephasing noise [Eq. (3.5a)] with noise parameter $p_0 = 0.6$ and noiseless cSWAPs. We employ a single-qubit input $(m = 1)$ and an S gate as the main computation $(U|\Phi_1^+\rangle = (I \otimes S)|\Phi_1^+\rangle;\ U|\phi_0\rangle = S|\phi_0\rangle)$. As discussed in Sec. 3.2.2, we may set $|\phi_0\rangle = |+\rangle$ to ensure $(\omega_1, \omega_2) = (1, 1)$. In the incoherent case $(d = 1)$, we have:

$$
\begin{aligned}
\rho_{\text{out, ic}} &= 0.6(U|\Phi_1^+\rangle\langle\Phi_1^+|U^\dagger) + 0.4(ZU|\Phi_1^+\rangle\langle\Phi_1^+|U^\dagger Z) \\
&= 0.5\,|00\rangle\langle00| - 0.1i\,|00\rangle\langle11| + 0.1i\,|11\rangle\langle00| + 0.5|11\rangle\langle11|,
\end{aligned}
\tag{3.25}
$$

and $\mathcal{F}_{\text{CJ, ic}} = 3/5 = 0.6$, where the ideal output is:

$$
U|\Phi_1^+\rangle\langle\Phi_1^+|U^\dagger = 0.5\,|00\rangle\langle00| - 0.5i\,|00\rangle\langle11| + 0.5i\,|11\rangle\langle00| + 0.5|11\rangle\langle11|.
\tag{3.26}
$$

Now, let us contrast this with the output of the coherent protocol. Plugging into Eq. (3.17), one finds that $\mathcal{A}_d = 0.6^{d-1}$, and the output state is therefore:

$$
\begin{aligned}
\rho_{\text{out}} &= \frac{0.6^2}{3}\left[\left(0.6(U|\Phi_1^+\rangle\langle\Phi_1^+|U^\dagger) + 0.4(ZU|\Phi_1^+\rangle\langle\Phi_1^+|U^\dagger Z)\right)\right. \\
&\quad \left. \ldots + 2\left(0.6(U|\Phi_1^+\rangle\langle\Phi_1^+|U^\dagger) + 0(\ldots)\right)\right] \\
&= 0.818(U|\Phi_1^+\rangle\langle\Phi_1^+|U^\dagger) + 0.182(ZU|\Phi_1^+\rangle\langle\Phi_1^+|U^\dagger Z) \quad \text{(normalized)} \\
&= 0.5\,|00\rangle\langle00| - 0.318i\,|00\rangle\langle11| + 0.318i\,|11\rangle\langle00| + 0.5|11\rangle\langle11|.
\end{aligned}
\tag{3.27}
$$

where $(\ldots)$ contains the remainder of the summation over Kraus operators with $(i, j) = (I, Z), (Z, I), (Z, Z)$. From Eq. (3.27), we can observe how the coherent term suppresses the undesired terms involving Z and boosts the amplitude of the ideal computation $(\propto U|\Phi_1^+\rangle\langle\Phi_1^+|U^\dagger)$, thereby resulting in an improved $\mathcal{F}_{\text{CJ}}$. Recall that in doing so, we have measured the control and auxiliaries in orthonormal bases spanned by $|+_3\rangle$ and $U|\phi_0\rangle$

respectively. The results of these measurements are shown in in Table 3.1.

Table 3.1: Sample results after performing the gate-based protocol for fidelity enhancement, before unitary optimization. The postselected outcome is highlighted in bold.

q	Outcomes (ctrl, aux)	$\rho_{\text{out}}^{(q)}(\lvert 00\rangle\langle 00\rvert, \lvert 00\rangle\langle 11\rvert, \lvert 11\rangle\langle 00\rvert, \lvert 11\rangle\langle 11\rvert)$	$\mathcal{F}_{\text{CJ}}^{(q)}$	$p_s^{(q)}$
1	**0, 00**	**0.5, -0.318i, 0.318i, 0.5**	**0.818**	**0.264**
2	0, 01	$0.662, -0.0004i, 0.0004i, 0.338$	0.5	0.123
3	0, 10	$0.469, -0.096i, 0.096i, 0.531$	0.596	0.125
4	0, 11	$0.36, 0.0116i, -0.0116i, 0.64$	0.488	0.115
5	1, 00	$0.5, 0.5i, -0.5i, 0.5$	0	0.048
6	1, 01	$0.247, 0.0789 - 0.256i, 0.0789 + 0.256i, 0.753$	0.756	0.0393
7	1, 10	$0.534, 0.0268 - 0.105i, 0.0268 + 0.105i, 0.466$	0.605	0.0567
8	1, 11	$0.6875, -0.108 - 0.25i, -0.108 + 0.25i, 0.3125$	0.75	0.0427
9	2, 00	$0.5, 0.5i, -0.5i, 0.5$	0	0.048
10	2, 01	$0.247, -0.0789 - 0.256i, -0.0789 + 0.256i, 0.753$	0.756	0.0393
11	2, 10	$0.534, -0.0268 - 0.105i, -0.0268 + 0.105i, 0.466$	0.605	0.0567
12	2, 11	$0.6875, 0.108 - 0.25i, 0.108 + 0.25i, 0.3125$	0.75	0.0427

Here, the measured outcomes '0', '1', and '2' correspond to the computational basis states $\lvert 0\rangle$, $\lvert 1\rangle$, and $\lvert 2\rangle$ respectively. We observe that the "all 0" outcome ($q = 1$) corresponds to ρ_{out} and is therefore postselected (with $p_s = 0.264$) when using the probabilistic approach (the other outcomes are associated with states that are orthogonal to the control and/or auxiliaries). Note that this outcome corresponds to the highest $\mathcal{F}_{\text{CJ}}$ (see Ref. [65] for details) and that:

$$\mathcal{F}_{\text{CJ}} = 0.818 > 0.6 = \mathcal{F}_{\text{CJ, ic}}. \tag{3.28}$$

From this, we find the infidelity ratio to be:

$$\mathcal{R} = (1 - 0.6)/(1 - 0.818) = 2.2. \tag{3.29}$$

One can enhance the relatively low p_s by employing the ϵ-deterministic approach. During optimization, we determine the unitary correction (from the single qubit Clifford group) for each output state (outcome q) that results in the greatest fidelity increase. The outcomes are then accepted one-by-one from the highest to lowest optimized fidelity until the desired $\epsilon = 1 - p_s$ is attained. Suppose one wishes to achieve $\epsilon = 0.25 \implies p_s = 0.75$. Then, the approach yields the optimized results in Table 3.2.

Table 3.2: Sample results after performing the gate-based protocol for fidelity enhancement, after unitary optimization. Outcomes are arranged from the highest optimized fidelity to the lowest.

q	Outcomes (ctrl, aux)	$U_c^{(q)}$	$U_c^{(q)}\rho_{\text{out}}^{(q)}U_c^{(q)\dagger}(\lvert 00\rangle\langle 00\rvert,\lvert 00\rangle\langle 11\rvert,\lvert 11\rangle\langle 00\rvert,\lvert 11\rangle\langle 11\rvert)$	$\mathcal{F}_{\text{CJ, opt}}^{(q)}$	$p_s^{(q)}$	Accept? $(\epsilon = 0.25)$
5	$1,00$	$I \otimes SS$	$0.5, -0.5i, 0.5i, 0.5$	1	0.048	✓
9	$2,00$	$I \otimes SS$	$0.5, -0.5i, 0.5i, 0.5$	1	0.048	✓
1	$0,00$	$I \otimes I$	Same as in Table 3.1	0.818	0.264	✓
6	$1,01$	$I \otimes I$	Same as in Table 3.1	0.756	0.0393	✓
10	$2,01$	$I \otimes I$	Same as in Table 3.1	0.756	0.0393	✓
8	$1,11$	$I \otimes I$	Same as in Table 3.1	0.75	0.0427	✓
12	$2,11$	$I \otimes I$	Same as in Table 3.1	0.75	0.0427	✓
7	$1,10$	$I \otimes I$	Same as in Table 3.1	0.605	0.0567	✓
11	$2,10$	$I \otimes I$	Same as in Table 3.1	0.605	0.0567	✓
3	$0,10$	$I \otimes I$	Same as in Table 3.1	0.596	0.125	✓
4	$0,11$	$I \otimes SS$	$0.36, -0.0116i, 0.0116i, 0.64$	0.512	0.115	✗
2	$0,01$	$I \otimes I$	Same as in Table 3.1	0.5	0.123	✗

We see that by accepting further outcomes (excluding $q = 2$ and 4), we are able to improve p_s from 0.264 to

$$\sum_{q\in\{1,\dots,12\}\setminus\{2,4\}} p_s^{(q)} = 0.762 \tag{3.30}$$

at the expense of slightly lower $\mathcal{F}_{\text{CJ}}$ and $\mathcal{R}$ values (note that exceeding the desired p_s is common given the finite number of outcomes). Nevertheless, we still obtain a degree of enhancement since:

$$\bar{\mathcal{F}}_{\text{CJ, opt}} = \sum_{q\in\{1,\dots,12\}\setminus\{2,4\}} p_s^{(q)}\mathcal{F}_{\text{CJ,opt}}^{(q)} = 0.7588 > 0.6 = \mathcal{F}_{\text{CJ, ic}}, \tag{3.31}$$

and,

$$\mathcal{R} = (1 - 0.6)/(1 - 0.7588) = 1.658 > 1. \tag{3.32}$$

3.2.5 Performance and analysis

We now proceed to analyze the numerical performance of the coherent GB protocol in a variety of settings involving the probabilistic and ϵ-deterministic approaches. Here, we examine two kinds of unitaries U, namely, T and CX gates for one- and two-qubit input scenarios respectively. These operations are essential to many quantum algorithms, with the former used in magic state distillation [110], and the latter essential for stabilizer measurements [111]. To assess the protocol directly, we first present results where the control register and generalized cSWAPs are assumed to be noiseless. Then, we analyze a separate scenario where noise affects them both, which we use to establish an upper bound on the achievable fidelity. For the noise, we apply either dephasing and depolarizing noises to T and CX. These choices are motivated by the ideas of Sec. 3.2.2, and ensures that a choice of $|\phi_0\rangle$ such that $\omega_2 = 1$ exists.

The GB protocol results are shown in Fig. 3.3 for $d = 2, 3, 4$ superposition branches and various noise parameters p_{ne}. In (a)–(e), we employ the parameterized auxiliary states:

$$|\phi_0\rangle = \cos(\theta)|0\rangle + \sin(\theta)|1\rangle \quad \text{(1 input)}; \tag{3.33}$$

$$|\phi_0\rangle = \sin^2(\theta)|00\rangle + \cos(\theta)\sin(\theta)(|01\rangle + |10\rangle) + \cos^2(\theta)|11\rangle \quad \text{(2 inputs)}, \tag{3.34}$$

which by varying θ, permits us to examine scenarios with differing $\omega_1 \in [0,1]$. As general observations, we find that (1) $\mathcal{R}$ is consistently higher when $\omega_1 = 1$ vs. < 1, and (2) $\mathcal{R}$ tends towards larger values as d increases. These agree with the asymptotic behavior of $\mathcal{F}_{\mathrm{CJ}}$ and theoretical predictions in the noiseless limit (see Ref. [65]). We further confirm observation (1) by varying ω_1 at a fixed p_{ne}, which we perform for the probabilistic CX case [Fig. 3.3(e)]. As expected, there is a steady increase in $\mathcal{R}$ upon increasing ω_1 toward 1. In the probabilistic scenarios [Figs. 3.3(a), (c), (e)], we confirm that the success probability p_s decays with increasing d and ω_1. These behaviours are a consequence of Eqs. (3.22) and (3.23)[6], and suggest an appropriate use for the ϵ-deterministic approach [Figs. 3.3(b), (d)]. In these scenarios, we study the limiting case where $\epsilon = 0$ and all outcomes are retained (i.e. fully deterministic). While they exhibit lower $\mathcal{R}$ values than in the probabilistic approach, we remarkably find that $\mathcal{R} > 1$ for a significant range of p_{ne} parameters, suggesting an advantage can be obtained by performing the protocol as-is. Moreover, while $\mathcal{R}$ decreases in the probabilistic approach with greater noise (smaller p_{ne}), we witness increases in the deterministic approach thanks to the optimization employed. Such unitary corrections can further reduce the noise if they are predominantly weighted toward single Pauli operators.

---

[6]As ω_1 increases, $\mathcal{R}$ and hence $\mathcal{F}_{\mathrm{CJ}} = \chi'_{00}$ undergo increases. Since χ_{00} (the incoherent fidelity) is fixed for some d, Eq. (3.22) implies p_s must decrease.
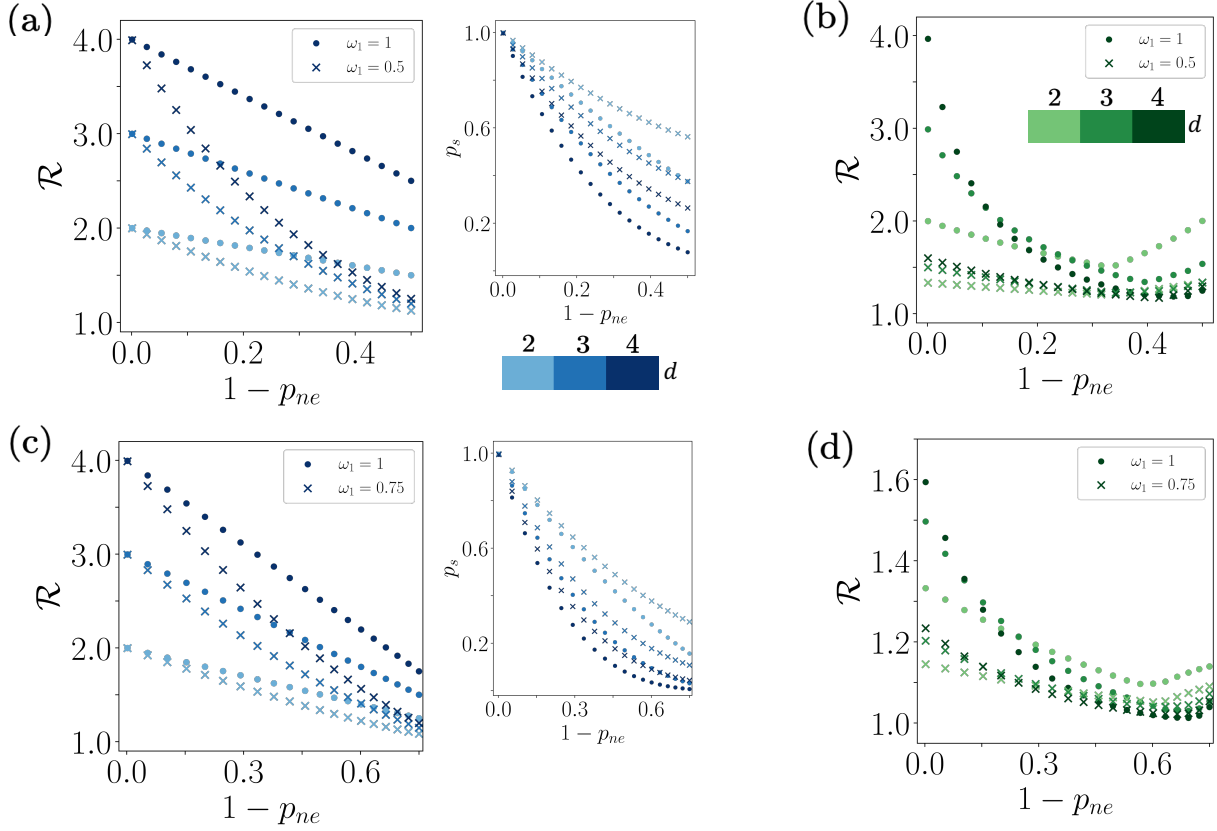
FIG. 3.3: **Numerical performance of the coherent gate-based (GB) protocol in mitigating various noisy scenarios, with variable number of superposition branches**. First row: Dephasing noise applied to the T gate ($m = 1$) using the *(a)* probabilistic and *(b)* fully deterministic ($\epsilon = 0$) approaches for $\omega_1 = 0.5, 1$. Second row: Dephasing noise applied to the CX gate ($m = 2$) using the *(c)* probabilistic and *(d)* fully deterministic approaches for $\omega_1 = 0.75, 1$. For all plots, the incoherent to coherent fidelity ratio $\mathcal{R}$ and postselection success probability p_s are plotted vs. the no-error probability p_{ne} [defined by Eq. (3.5a)].

42
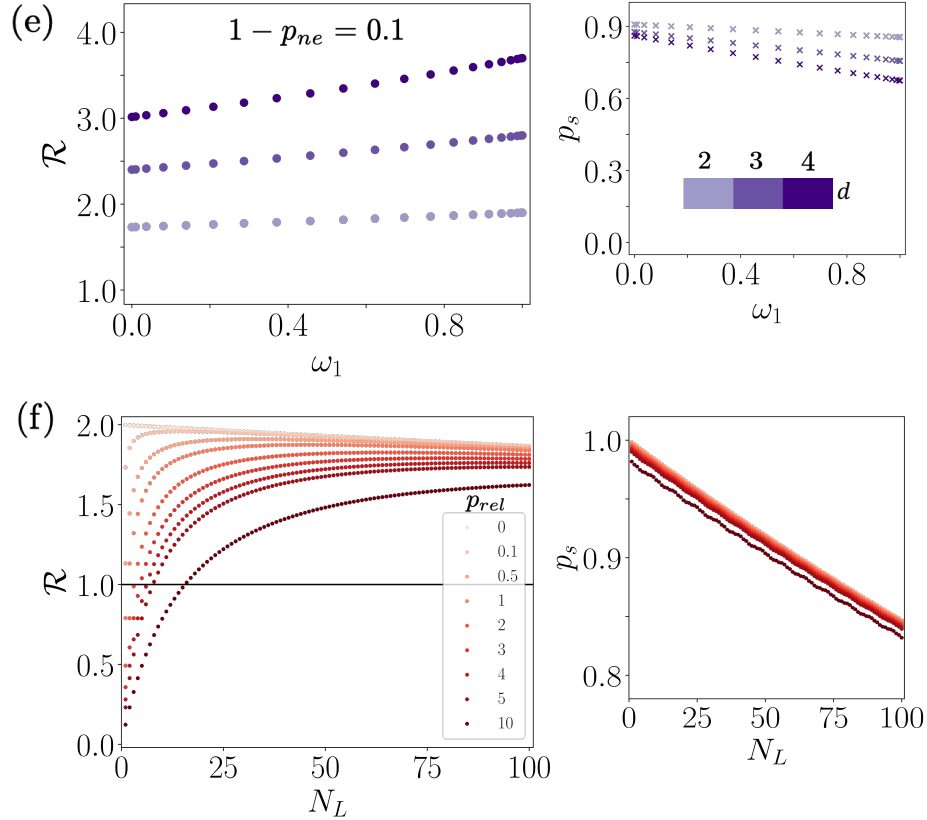
FIG. 3.3 (cont.): *(e)* Further analysis of *(c)* where ω_1 is varied [via the choice of $|\phi_0\rangle$ in Eq. (3.33)] between $[0, 1]$ at a fixed p_{ne} of 0.9. Each cSWAP is also assumed to be noiseless, with noise stemming from the main computation only (applied after each unitary gate). Both $\mathcal{R}$ and p_s are plotted as functions of ω_1. *(f)* Performance of the probabilistic approach ($d = 2$) when depolarizing noise acts on a 2-qubit operation ($m = 2$) with variable number of layers N_L. Each layer consists of T gates on each of the two qubits, followed by a CX. Here, the auxiliary state is chosen such that $(\omega_1, \omega_2) = (1, 1)$ [see Eq. (3.34)]. In addition to the main computation (acting with fixed noise parameter $1 - p_0 = 10^{-4}$), noise also acts after each cSWAP operation with parameter $1 - p$ [defined by Eq. (3.5b)]. The target plot shows $\mathcal{R}$ vs. N_L for various noise level ratios p_{rel} between cSWAPs and operations in U. For reference, the threshold for noise reduction ($\mathcal{R} = 1$) is plotted as a horizontal line. The corresponding success probabilities p_s are also plotted vs. N_L.

In Fig. 3.3(f), we analyze a more realistic scenario where the cSWAPs and control registers are noisy. In contrast to the noiseless cases above, we choose U to be a multi-gate operation $CX(T \otimes T)$ of N_L layers. The rationale for a variable computation depth stems from the cSWAP always being applied *twice*, which allows us to analyze a direct competition between noises affecting the cSWAPs and U. We apply depolarizing noise after applying each gate in U (only on registers acted by the particular gate), and to all registers after applying a cSWAP. We then analyze $\mathcal{R}$ vs. N_L for different ratios $p_{rel} = (1 - p)/(1 - p_0)$ between the cSWAP noise $(1 - p)$ and gate noise (fixed at $1 - p_0 = 10^{-4}$). Here, we set $|\phi_0\rangle = |\Phi_1^+\rangle$[7], $d = 2$, and employ the probabilistic approach. We observe that when $p_{rel} = 0$, the plot agrees with Fig. 3.3(d) for $d = 2$ in the noiseless limit. However, as noise is added to the cSWAP $(p_{rel} > 0)$, we see that the effect is detrimental for small N_L, with $\mathcal{R} < 1$ indicating a worse performance than the corresponding incoherent case. Interestingly, we observe that advantages $(\mathcal{R} > 1)$ begin manifesting as N_L increases, with more computational layers required for noisier cSWAPs. This directly relates to the impact of noise on the auxiliary states (see Sec. 3.2.1), and implies that the noise on U must *exceed* the noise on both cSWAPs for our protocol to yield a level of noise reduction. While the large circuit depths required by U pose challenges with experimental realization, we remark that our choice to break U up into separate operations lends a more practical approach, in light of recent proposals for realizing high-fidelity multi-qubit gates [112, 113, 114].

From the results, we can make further remarks regarding use of the two postprocessing approaches. The increasing $\mathcal{R}$ values and high p_s in Fig. 3.3(f) suggest the probabilistic approach is well-suited when applied to a noisy computation of large depth. In contrast, the ϵ-deterministic approach is best employed when applied to a single noisy gate [Fig. 3.3(b), (d)], especially when one wishes to both maximize p_s and achieve a non-negligible advantage. This can be useful even within larger computations – if several constituent gates are of low fidelity, one may apply the approach to each upon their execution, in a multi-stage application of the protocol. The suitability toward small computations is particularly relevant, owing to the extra resources needed to perform the optimization step (Sec. 3.2.3).

---

[7]See note 4.
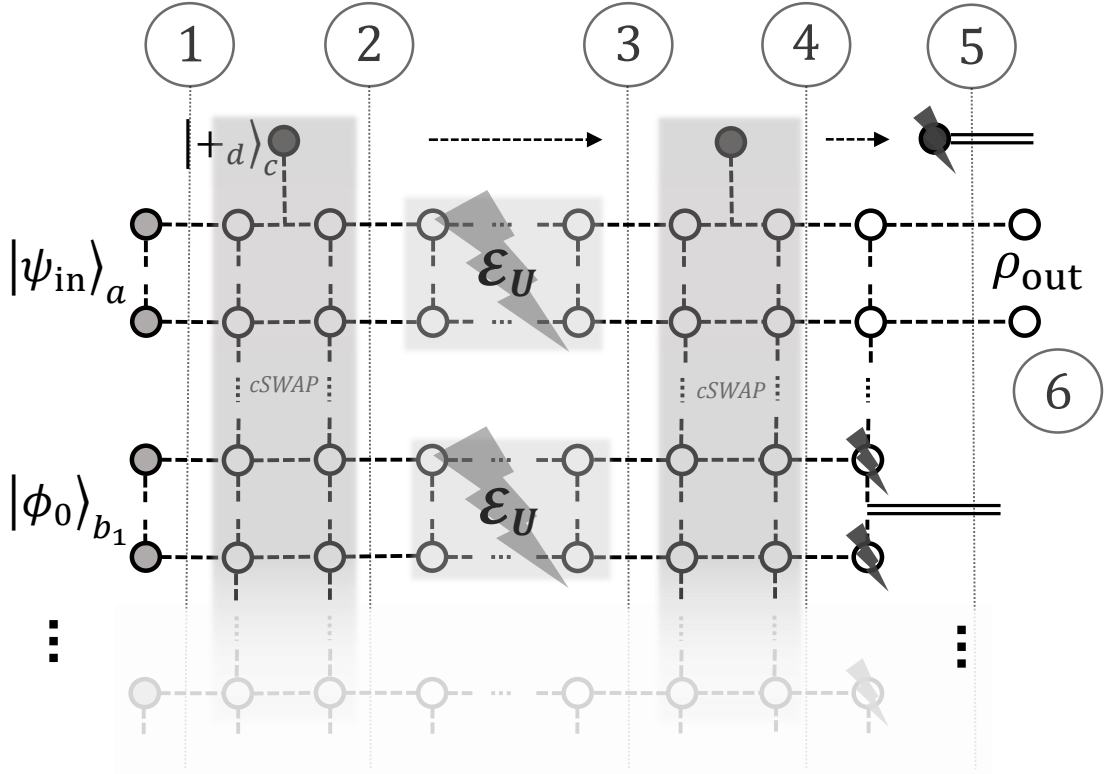
FIG. 3.4: **Schematic representation of the MB protocol for fidelity enhancement**. A d-dimensional control register in the state $|+\rangle_d$ generates the superposition by swapping the m-qubit input state with d m-qubit auxiliaries in a controlled manner. All operations (cSWAP, U) are performed in the MB-paradigm, with each represented by a cluster (resource) state (see App. A and Ref. [2]). To mirror the GB protocol (Fig. 3.2), all cluster states (including inputs and outputs) are concatenated together to form steps 1–5. Here, the noise $\mathcal{E}_U$ arises from the imperfect preparation of each cluster realizing U. To facilitate analysis, the control register is treated externally and not integrated as part of the overall cluster (see App. H).

3.3 Enhanced measurement-based computation

We now present the MB-QC version of our coherent protocol. Although the underlying manner in which noise acts is different, it also serves as a limiting factor towards scaling the size and complexity of computations. We first describe the protocol, then proceed to compare and contrast its performance in the same scenarios of Sec. 3.2.5.

3.3.1 Working protocol

The MB protocol is conceptually similar to the GB protocol, with the primary difference being the use of cluster states and *ancillary* qubit measurements (see App. A) to perform the cSWAPs and computations U. Our choice to perform MB-QC on clusters contrasts with the graphs employed in Chapter 2, and is well-suited since we regard the computations of interest on the level of individual elements (i.e. gates). More specifically, we replace each gate in Fig. 3.2 with its equivalent cluster state $|G\rangle$, and concatenate them to form the resource state depicted in Fig. 3.4. Each $|G\rangle$ is expressed in the form of Eq. (A.2), and may be derived from well-known MB patterns in Ref. [2].

Another difference concerns the noise model governing MB-QC. Due to the larger number of qubits present, errors are incurred *multiple* times within an operation, which reflect potential imperfections in preparing the cluster and the quality of the projective measurements. As such, we act noise locally on each qubit j in a given cluster prior to its measurement, i.e.:

$$\rho_G = \prod_{j=1}^{N} \mathcal{E}_j(|G\rangle\langle G|),$$

(3.35)

where N is the number of qubits in the cluster. Here, we may employ the same noise representation in terms of Kraus operators ($\mathcal{E} = \sum_i K_i \rho K_i^\dagger$; see Sec. 3.1.1). However, it is crucial to note that even if identical noise is applied to each qubit j, the additional factors inherent to MB-QC (e.g. chosen measurement bases, entanglement structure) suggests that the overall effect of noise becomes nontrivial. This can greatly influence the degree of noise reduction when compared to the GB version, and places emphasis on the specific design of the cluster used to implement a particular operation.

As Fig. 3.4 shows, the overall setup consists of the same d-level control, m-qubit input, and $d - 1$ auxiliary states. In Step 1, the input and auxiliary systems are prepared as components of the overall cluster state, with their input qubits initialized to Eq. (3.13). However, the control register is less straightforward to prepare in this manner, as there are no well-known MB patterns for $d > 2$. As such, we treat the control externally of the cluster

and prepare it according to the GB protocol. This also impacts the way we perform the cSWAP in Steps 2, as it requires entire clusters to be swapped at a time. To achieve this, we resort to a mathematical procedure involving block-diagonal matrices, which is detailed in App. H. We then perform MB-QC on the parts of the cluster realizing U and the final cSWAP (Steps 3 and 4), with noise applied prior to measuring each qubit. Depending on their outcomes, we apply correcting byproduct operators (App. A) to ensure the resulting state is always the same. As before, we measure the control and auxiliary systems at Step 5 to obtain ρ_{out}, with the same probabilistic (Sec. 3.2.2) and ϵ-deterministic (Sec. 3.2.3) methods to post-process the result.

3.3.2 Performance and analysis

We study the numerical performance of the coherent MB protocol in scenarios similar to that of the GB protocol (Sec. 3.2.5). For each simulation, the entire resource state is performed in a piecemeal fashion, such that the output of one cluster operation is used as input for the next. This keeps the size of the processed cluster relatively small, particularly as density matrices are used and manipulated throughout.

In Figs. 3.5(a) and (c) (i.e. probabilistic approach with noiseless cSWAPs and control), we observe that the overall behaviour and trends resemble that of the GB results; however, their $\mathcal{R}$ values tend to decrease more rapidly with increasing noise parameters. We attribute this to our multiple applications of noise (i.e. on every cluster qubit), which on average, leads to a noisier quantum channel overall. Consequently, we require nonzero ϵ (i.e. 0.5 and 0.75 for the T and CX operations respectively) in the ϵ-deterministic cases [Figs. 3.5(b), (d)] in order to attain comparable $\mathcal{R}$ values to the GB results. Nevertheless, these scenarios highlight situations in which noise reduction over the incoherent case is obtained to some degree. We also remark, that due to the more complex nature of the noise model, we heuristically choose each $|\phi_0\rangle$ to be $|+\rangle^{\otimes m}, m = 1, 2$, which generally corresponds to $(\omega_1, \omega_2) < (1, 1)$. As evidenced by Figs. 3.5(b) and (d), this is responsible for the $\mathcal{R}$ values being slightly above or consistently equal to 1 when all outcomes are retained ($\epsilon = 0$), indicating little to no advantage.

Similar to the GB protocol, we analyze the scenario where noise acts on the control and every cluster qubit involved in the cSWAP operation, and study $U = [CX(T \otimes T)]^{N_L}$ for $d = 2$. Here, N_L is linearly proportional to the cluster size. Again, we choose $|\phi_0\rangle = |\Phi_1^+\rangle$ and use the probabilistic approach. However, we set cSWAP and U noises to the same level p_0, and keep all measurement outcomes constant to minimize variability stemming from the byproduct operators (App. A) acting on noisy states. As Fig. 3.5(e) confirms, we recover the same observation that our protocol is beneficial ($\mathcal{R} > 1$) when N_L is sufficiently

47
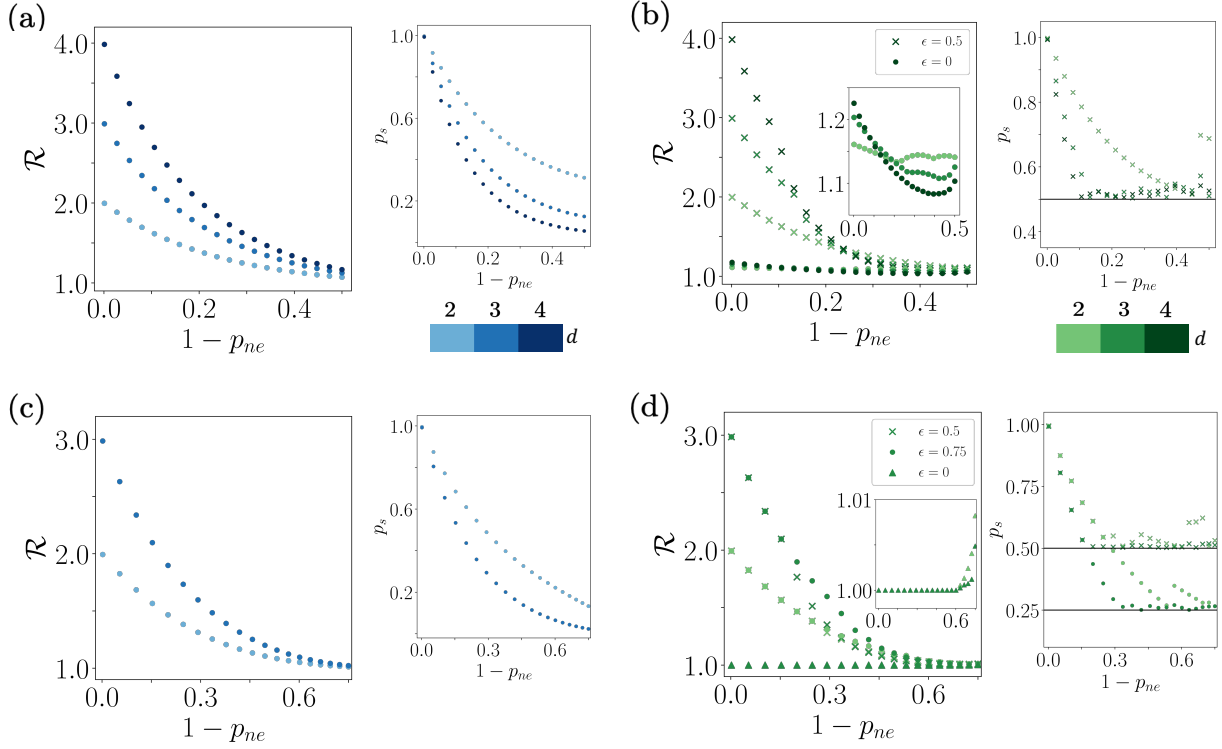
FIG. 3.5: **Numerical performance of the coherent measurement-based (MB) protocol in mitigating various noisy scenarios, with variable number of superposition branches**. All operations are implemented by preparing and concatenating cluster states, then measuring their qubits in prescribed bases. First row: Dephasing noise applied to the T operation ($m = 1$) using the *(a)* probabilistic and *(b)* ϵ-deterministic approaches for $\epsilon = 0, 0.5$. Second row: Dephasing noise applied to the CX operation ($m = 2$) using the *(c)* probabilistic and *(d)* ϵ-deterministic approaches for $\epsilon = 0, 0.5, 0.75$. For *(a)–(d)*, the incoherent to coherent fidelity ratio $\mathcal{R}$ and postselection success probability p_s are plotted vs. the no-error probability p_{ne} [defined by Eq. (3.5a)]. For the latter, all nonzero ϵ are indicated as horizontal lines (at $1 - \epsilon$) to indicate the minimum thresholds for accepting outcomes. Each cSWAP is also assumed to be noiseless, with noise stemming from the main computation only (applied to each cluster qubit comprising the unitary).

48

FIG. 3.5 (cont.): *(e)* Performance of the probabilistic approach ($d = 2$) when depolarizing noise acts on a 2-qubit operation ($m = 2$) with variable number of layers N_L. Each layer consists of a cluster state implementing T operations on each of the two qubits, followed by a CX. In addition to the main computation, noise also acts on each cluster qubit comprising the cSWAPs. Both sources of noise are applied with the same noise parameter p_0 [defined by Eq. (3.5b)], and the target plot shows $\mathcal{R}$ vs. N_L for various $1 - p_0$. For reference, the threshold for noise reduction ($\mathcal{R} = 1$) is plotted as a horizontal line. The corresponding success probabilities p_s are also plotted vs. N_L. In all scenarios, the auxiliary state is chosen heuristically as $|+\rangle$ ($m = 1$) or $|++\rangle$ ($m = 2$) which corresponds to $(\omega_1, \omega_2) < (1, 1)$.

large such that the noise on U exceeds that on the cSWAPs. This can be achieved with a wide range of p_0 and N_L values, albeit with generally lower p_s.

3.4   Conclusions

In this chapter, we have introduced and analyzed a protocol that is capable of mitigating noisy quantum computations. It involves performing an identical number of them in a coherent superposition, and leads to a interference process between noises in different branches. We demonstrated via numerical simulations that this interference can lead to varying degrees of noise mitigation, which can be attained in a range of different scenarios involving basic unitary operations and Pauli noises. We also provided a concrete way of performing the protocol in both the GB and MB paradigms, which employs exotic circuit features such as the inclusion of auxiliary registers, generalized cSWAP operations, and ancillary cluster qubits from MB-QC. In particular, we have seen that the amount of mitigation is dictated not only by how noise affects the inputs, but also the *auxiliary* systems, and that an appropriate choice for their initial states can maximize the advantage we obtain.

Ref. [65] presents further extensions of the protocol to a circuit-free version that does not require additional cSWAPs and auxiliary systems, and is tailored to an physical implementation using photonic interferometers and beamsplitters. A so-called nested scenario is also described, which enables one to readily scale the number of branches without increasing the dimensionality of the control system. It also enables the degree of noise mitigation to be optimized using *non-identical* auxiliary states.

Our coherent superposition protocol serves as a NISQ-friendly approach to overcoming noise and decoherence in quantum computations, standing in contrast to resource-heavy fault-tolerant computing and error correction methods. We remark however, that the tools and techniques of our protocol are not limited to this area, and have the potential to impact other consequential fields such as quantum communication, metrology, and sensing [115].

Chapter 4

Conclusions and Outlook

In this thesis, we examined two different ways in which exotic circuits may be employed to enhance quantum algorithms and computation. We employ the umbrella term "exotic" to mean elements of ancillary registers, graphs (or clusters), mid-circuit measurements, and conditional operations which are commonly found in the MB paradigm but less common for traditional GB-QC.

In Chapter 2, we employed exotic circuits in VQE algorithms, where they serve as the ansatz for determining ground state properties. We achieved this by blending elements of the GB and MB computing paradigms to produce a *hybrid* VQE ansatz. Using corresponding ansatz modification and Pauli gadget techniques, we obtain compact circuits that are both versatile and resource-efficient. The Pauli gadget, in particular, minimizes the overhead and offers a practical means for performing MB-QCs on platforms that do not implement highly entangled resource states.

To test our hybrid approaches, we performed these circuits on IBM superconducting qubit systems, where they succeeded in simulating key ground-state phenomena (i.e. phase transitions and perturbative effects) for a broad range of physical models. These models span different areas of physics, encompassing lattice gauge theory ($\mathbb{Z}_2$), 1D quantum chromodynamics, error-correcting surface codes (PC), and molecular systems (LiH). Although our results are similar in quality to those of standard GB-VQEs, our inclusion of both paradigms expands the range of compatible quantum platforms for VQEs, while retaining the resource and performance benefits of each. Our approach also displayed promise in scaling up to larger systems and circuit depths, maintaining a strong performance in their corresponding VQE simulations. These innovations constitute a new, enhanced way of thinking about VQE algorithms from both the design and realization perspectives.

In Chapter 3, we used exotic circuits to mitigate noise impacting general quantum computations. Specifically, we augment a circuit with multiple *auxiliary* registers and perform the same noisy computation on each in a controlled coherent superposition. This creates an interference process between registers that, under certain conditions, allows for a partial elimination of the noise. We identified such conditions in scenarios involving Pauli noises and basic unitary operations, and gave a concrete, scalable protocol to perform these circuits in the GB and MB paradigms. We performed our protocol as numerical simulations that confirmed and quantified the advantage via the infidelity ratio as a figure of merit. Most notably, they indicated that our advantages can be maximized by selecting an appropriate initial auxiliary state, and increasing the number of superposition branches and/or computational depth. Despite the MB and GB paradigms being equivalent in the noiseless limit, their differing noise models led to significant differences in their noise reduction mechanisms. Nevertheless, we find both paradigms to be capable of mitigating noise, with similar results and steps in their corresponding protocols.

We emphasize that both GB- and MB-computing paradigms were treated equally in our discussions. MB-QC has so far proven to be more challenging to implement experimentally, particularly with high-fidelity cluster states [116]. In contrast, our exotic circuits took a different approach by exploiting MB-QC in traditionally GB applications. As mentioned above, we have showcased tools (i.e. Pauli gadgets) that enable a resource-efficient implementation of MB-QC. These aspects serve to highlight MB-QC as an equally powerful approach not only in the subset of applications involving exotic circuits, but in other quantum physics and information science areas that rely on general quantum computations.

Looking ahead, our exotic circuits may be used to continue the lines of investigation begun in our works. For instance, one can identify other physical models and design possibilities suited for our hybrid VQE techniques, and attempt to perform simulations on hardware involving the scaled-up scenarios mentioned. Meanwhile, the coherent computation protocols may be applied within more concrete settings (e.g. using interferometric elements and photonic degrees of freedom), allowing one to ascertain whether our noise advantages have a consequential impact. These investigations may be facilitated by leveraging numerical packages (e.g. Qiskit [41], QuTiP [117], Cirq [118]) to validate our circuits, as well as tailoring the circuits' techniques toward the requirements of specific problems, which would enable more direct analyses.

In a broader sense, this thesis demonstrated how exotic circuits can be beneficial not only for near-term QC, but also for solving problems that are classically intractable and/or incur significant resources on quantum hardware. It is hoped that the work expounded upon here will provide new insights toward achieving reliable, high-fidelity results in the NISQ era and beyond.

References

[1] Michael A. Nielsen and Isaac L. Chuang. Chapter 4: Quantum Circuits, page 171–215. Cambridge University Press, 2010.

[2] Robert Raussendorf, Daniel E. Browne, and Hans J. Briegel. Measurement-based quantum computation on cluster states. Phys. Rev. A, 68:022312, Aug 2003.

[3] Richard Jozsa. An introduction to measurement based quantum computation. Quantum Information Processing: From Theory to Experiment, 199:137, 2006.

[4] Tzu-Chieh Wei. Measurement-based quantum computation. Oxford Research Encyclopedia of Physics, 2021.

[5] Hezi Zhang, Anbang Wu, Yuke Wang, Gushu Li, Hassan Shapourian, Alireza Shabani, and Yufei Ding. A compilation framework for photonic one-way quantum computation. arXiv:2209.01545, 2022.

[6] Tomoyuki Morimae. Acausal measurement-based quantum computing. Physical Review A, 90(1), 2014.

[7] Xiaoqian Zhang. Measurement-based universal blind quantum computation with minor resources. Quantum Information Processing, 21(1), 2021.

[8] Janna Hinchlif. Measurement-based quantum error correction. University of Brisftol, 2015.

[9] N. de Beaudrap, V. Danos, and E Kashefi. Phase map decomposition for unitaries. arXiv:quant-ph/0603266, 2006.

[10] R Duncan, A Kissinger, S Perdrix, and van der Wetering. Graph-theoretic simplification of quantum circuits with the ZX-calculus. Quantum, 4:279, 2020.

[11] Daniel Eric Gottesman. Stabilizer Codes and Quantum Error Correction. California Institute of Technology, 1997.

[12] Daniel A. Lidar and Todd A. Brun, editors. Quantum Error Correction. Cambridge University Press, September 2013.

[13] Sarmed A Rahman, Randy Lewis, Emanuele Mendicelli, and Sarah Powell. Self-mitigating trotter circuits for SU(2) lattice gauge theory on a quantum computer. Phys. Rev. D, 106:074502, Oct 2022.

[14] Zhenyu Cai, Ryan Babbush, Simon C. Benjamin, Suguru Endo, William J. Huggins, Ying Li, Jarrod R. McClean, and Thomas E. O'Brien. Quantum error mitigation. arXiv:2210.00921, 2022.

[15] Charles H. Bennett, Gilles Brassard, Sandu Popescu, Benjamin Schumacher, John A. Smolin, and William K. Wootters. Purification of noisy entanglement and faithful teleportation via noisy channels. Physical Review Letters, 76(5):722–725, 1996.

[16] Jian-Wei Pan, Christoph Simon, Časlav Brukner, and Anton Zeilinger. Entanglement purification for quantum communication. Nature, 410(6832):1067–1070, 2001.

[17] John Preskill. Quantum Computing in the NISQ era and beyond. Quantum, 2:79, August 2018.

[18] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J Love, Alán Aspuru-Guzik, and Jeremy L O'Brien. A variational eigenvalue solver on a photonic quantum processor. Nat. Commun., 5:4213, 2014.

[19] Jarrod R McClean, Jonathan Romero, Ryan Babbush, and Alán Aspuru-Guzik. The theory of variational hybrid quantum-classical algorithms. New J. Phys., 18(2):023023, 2016.

[20] Jules Tilly, Hongxiang Chen, Shuxiang Cao, Dario Picozzi, Kanav Setia, Ying Li, Edward Grant, Leonard Wossnig, Ivan Rungger, George H. Booth, and Jonathan Tennyson. The variational quantum eigensolver: A review of methods and best practices. Phys. Rep., 986:1–128, 2022.

[21] Danny Paulson, Luca Dellantonio, Jan F. Haase, Alessio Celi, Angus Kan, Andrew Jena, Christian Kokail, Rick van Bijnen, Karl Jansen, Peter Zoller, and Christine A. Muschik. Simulating 2d effects in lattice gauge theories on a quantum computer. PRX Quantum, 2:030334, Aug 2021.

[22] R. R. Ferguson, L. Dellantonio, A. Al Balushi, K. Jansen, W. Dür, and C. A. Muschik. Measurement-based variational quantum eigensolver. Phys. Rev. Lett., 126:220501, Jun 2021.

[23] Mario Motta and Julia E. Rice. Emerging quantum computing algorithms for quantum chemistry. Wiley Interdiscip. Rev. Comput. Mol. Sci., 12(3):e1580, 2022.

[24] Bela Bauer, Dave Wecker, Andrew J. Millis, Matthew B. Hastings, and Matthias Troyer. Hybrid quantum-classical approach to correlated materials. Phys. Rev. X, 6:031045, Sep 2016.

[25] Bela Bauer, Sergey Bravyi, Mario Motta, and Garnet Kin-Lic Chan. Quantum algorithms for quantum chemistry and quantum materials science. Chem. Rev., 120(22):12685–12717, 2020.

[26] Mari Carmen Bañuls, Rainer Blatt, Jacopo Catani, Alessio Celi, Juan Ignacio Cirac, Marcello Dalmonte, Leonardo Fallani, Karl Jansen, Maciej Lewenstein, Simone Montangero, et al. Simulating lattice gauge theories within quantum technologies. Eur. Phys. J. D, 74(8):1–42, 2020.

[27] Andrew J Daley, Immanuel Bloch, Christian Kokail, Stuart Flannigan, Natalie Pearson, Matthias Troyer, and Peter Zoller. Practical quantum advantage in quantum simulation. Nature, 607(7920):667–676, 2022.

[28] Christian Kokail, Christine Maier, Rick van Bijnen, Tiff Brydges, Manoj K Joshi, Petar Jurcevic, Christine A Muschik, Pietro Silvi, Rainer Blatt, Christian F Roos, et al. Self-verifying variational quantum simulation of lattice models. Nature, 569(7756):355–360, 2019.

[29] Jan F. Haase, Luca Dellantonio, Alessio Celi, Danny Paulson, Angus Kan, Karl Jansen, and Christine A. Muschik. A resource efficient approach for quantum and classical simulations of gauge theories in particle physics. Quantum, 5:393, February 2021.

[30] Yasar Y Atas, Jinglei Zhang, Randy Lewis, Amin Jahanpour, Jan F Haase, and Christine A Muschik. SU(2) hadrons on a quantum computer via a variational approach. Nat. Commun., 12(1):6499, 2021.

[31] Robert Raussendorf and Hans J. Briegel. A one-way quantum computer. Phys. Rev. Lett., 86:5188–5191, May 2001.

[32] Dan Browne and Hans Briegel. One-way quantum computation. Quantum information: From foundations to quantum technology applications, pages 449–473, 2016.

[33] Hans J Briegel, David E Browne, Wolfgang Dür, Robert Raussendorf, and Maarten Van den Nest. Measurement-based quantum computation. Nat. Phys., 5(1):19–26, 2009.

[34] A. D. Córcoles, Maika Takita, Ken Inoue, Scott Lekuch, Zlatko K. Minev, Jerry M. Chow, and Jay M. Gambetta. Exploiting dynamic quantum circuits in a quantum algorithm with superconducting qubits. Phys. Rev. Lett., 127:100501, Aug 2021.

[35] MS Allman, CH Baldwin, M Foss-Feig, D Hayes, K Mayer, Ryan-Anderson C, and Neyenhuis B. Demonstration of the trapped-ion quantum ccd computer architecture. Nature, 592(7853):209, 2021.

[36] J. P. Gaebler, C. H. Baldwin, S. A. Moses, J. M. Dreiling, C. Figgatt, M. Foss-Feig, D. Hayes, and J. M. Pino. Suppression of midcircuit measurement crosstalk errors with micromotion. Phys. Rev. A, 104:062440, Dec 2021.

[37] Emma Deist, Yue-Hui Lu, Jacquelyn Ho, Mary Kate Pasha, Johannes Zeiher, Zhenjie Yan, and Dan M. Stamper-Kurn. Mid-circuit cavity measurement in a neutral atom array. Phys. Rev. Lett., 129:203602, Nov 2022.

[38] Sainath Motlakunta, Nikhil Kotibhaskar, Chung-You Shih, Anthony Vogliano, Darian Mclaren, Lewis Hahn, Jingwen Zhu, Roland Hablützel, and Rajibul Islam. Preserving a qubit during adjacent measurements at a few micrometers distance. arXiv:2306.03075, 2023.

[39] Roman Stricker, Davide Vodola, Alexander Erhard, Lukas Postler, Michael Meth, Martin Ringbauer, Philipp Schindler, Thomas Monz, Markus Müller, and Rainer Blatt. Experimental deterministic correction of qubit loss. Nature, 585(7824):207–210, 2020.

[40] TM Graham, L Phuttitarn, R Chinnarasu, Y Song, C Poole, K Jooya, J Scott, A Scott, P Eichler, and M Saffman. Mid-circuit measurements on a neutral atom quantum processor. arXiv:2303.10051, 2023.

[41] Qiskit contributors. Qiskit: An open-source framework for quantum computing, 2023.

[42] Harper R. Grimsley, George S. Barron, Edwin Barnes, Sophia E. Economou, and Nicholas J. Mayhall. Adaptive, problem-tailored variational quantum eigensolver mitigates rough parameter landscapes and barren plateaus. npj Quantum Information, 9(1), 2023.

[43] Anton Nykänen, Matteo A. C. Rossi, Elsi-Mari Borrelli, Sabrina Maniscalco, and Guillermo García-Pérez. Mitigating the measurement overhead of adapt-vqe with optimised informationally complete generalised measurements. arXiv:2212.09719, 2022.

[44] A.Yu. Kitaev. Fault-tolerant quantum computation by anyons. Ann. Phys., 303(1):2–30, 2003.

[45] Paweł Mazurek, Andrzej Grudka, Michał Horodecki, Paweł Horodecki, Justyna Łodyga, Łukasz Pankowski, and Anna Przysiężna. Long-distance quantum communication over noisy networks without long-time quantum memory. Physical Review A, 90(6), 2014.

[46] Kao-Yueh Kuo and Ching-Yi Lai. Comparison of 2d topological codes and their decoding performances. 2022 IEEE International Symposium on Information Theory (ISIT), 2022.

[47] R. Hübener, C. Kruszynska, L. Hartmann, W. Dür, M. B. Plenio, and J. Eisert. Tensor network methods with graph enhancement. Phys. Rev. B, 84:125103, Sep 2011.

[48] Albie Chan, Zheng Shi, Luca Dellantonio, Wolfgang Dür, and Christine A. Muschik. Hybrid variational quantum eigensolvers: merging computational models. arXiv:2305.19200, 2023.

[49] William K. Wootters. Entanglement of formation of an arbitrary state of two qubits. Physical Review Letters, 80(10):2245–2248, 1998.

[50] Hans J. Briegel and Robert Raussendorf. Persistent entanglement in arrays of interacting particles. Phys. Rev. Lett., 86:910–913, Jan 2001.

[51] Daniel Gottesman and Isaac L Chuang. Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations. Nature, 402(6760):390–393, 1999.

[52] Barbara M. Terhal and David P. DiVincenzo. Adaptive quantum computation, constant depth quantum circuits and Arthur-Merlin games. Quantum Inf. Comput., 4(2):134–145, 2004.

[53] Peter Høyer and Robert Špalek. Quantum fan-out is powerful. Theory Comput., 1(5):81–103, 2005.

[54] Sergey Bravyi and Alexei Kitaev. Universal quantum computation with ideal Clifford gates and noisy ancillas. Phys. Rev. A, 71:022316, Feb 2005.

[55] Dan Browne, Elham Kashefi, and Simon Perdrix. Computational depth complexity of measurement-based quantum computation. In Wim van Dam, Vivien M. Kendon, and Simone Severini, editors, Theory of Quantum Computation, Communication, and Cryptography, pages 35–46, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.

[56] Richard Jozsa and Marrten Van Den Nest. Classical simulation complexity of extended Clifford circuits. Quantum Info. Comput., 14(7, 8):633–648, May 2014.

[57] Adam Paetznick and Krysta M Svore. Repeat-until-success: non-deterministic decomposition of single-qubit unitaries. Quantum Inf. Comput., 14(15-16):1277–1301, 2014.

[58] Yasuhiro Takahashi and Seiichiro Tani. Collapse of the hierarchy of constant-depth exact quantum circuits. Comput. Complex., 25:849–881, 2016.

[59] Guanyu Zhu, Ali Lavasani, and Maissam Barkeshli. Universal logical gates on topologically encoded qubits via constant-depth unitary circuits. Phys. Rev. Lett., 125:050502, Jul 2020.

[60] Lorenzo Piroli, Georgios Styliaris, and J. Ignacio Cirac. Quantum circuits assisted by local operations and classical communication: Transformations and phases of matter. Phys. Rev. Lett., 127:220503, Nov 2021.

[61] Ruben Verresen, Nathanan Tantivasadakarn, and Ashvin Vishwanath. Efficiently preparing Schrödinger's cat, fractons and non-Abelian topological order in quantum devices. arXiv:2112.03061, 2021.

[62] Nathanan Tantivasadakarn, Ryan Thorngren, Ashvin Vishwanath, and Ruben Verresen. Long-range entanglement from measuring symmetry-protected topological phases. arXiv:2112.01519, 2021.

[63] Zhenning Liu and Alexandru Gheorghiu. Depth-efficient proofs of quantumness. Quantum, 6:807, September 2022.

[64] Sergey Bravyi, Isaac Kim, Alexander Kliesch, and Robert Koenig. Adaptive constant-depth circuits for manipulating non-abelian anyons. arXiv:2205.01933, 2022.

[65] Jorge Miguel-Ramiro, Zheng Shi, Luca Dellantonio, Albie Chan, Christine A. Muschik, and Wolfgang Dür. Enhancing quantum computation via superposition of quantum gates. arXiv:2304.08529, 2023.

[66] D.E. Browne and H.J. Briegel. One-way quantum computation - a tutorial introduction. arXiv:quant-ph/0603226v2, 2006.

[67] A. Jena, S. Genin, and M. Mosca. Pauli partitioning with respect to gate sets. arXiv:1907.07859, 2019.

[68] Ariel Shlosberg, Andrew J. Jena, Priyanka Mukhopadhyay, Jan F. Haase, Felix Leditzky, and Luca Dellantonio. Adaptive estimation of quantum observables. Quantum, 7:906, January 2023.

[69] D. Gottesman. Fault-tolerant quantum computation. Chaos Solit. Fractals, 10:1749–1758, 1999.

[70] L. Funcke, T. Hartung, K. Jansen, S. Kühn, M. Schneider, P. Stornati, and X. Wang. Towards quantum simulations in particle physics and beyond on noisy intermediate-scale quantum devices. Philos. Trans. R. Soc. A, 380(2216):20210062, 2022.

[71] Natalie Klco, Alessandro Roggero, and Martin J Savage. Standard model physics and the digital quantum revolution: thoughts about the interface. Rep. Prog. Phys., 85(6):064301, May 2022.

[72] Erez Zohar. Quantum simulation of lattice gauge theories in more than one space dimension—requirements, challenges and methods. Philos. Trans. R. Soc. A, 380(2216):20210069, 2022.

[73] Esteban A Martinez, Christine A Muschik, Philipp Schindler, Daniel Nigg, Alexander Erhard, Markus Heyl, Philipp Hauke, Marcello Dalmonte, Thomas Monz, Peter Zoller, et al. Real-time dynamics of lattice gauge theories with a few-qubit quantum computer. Nature, 534(7608):516–519, 2016.

[74] Christine Muschik, Markus Heyl, Esteban Martinez, Thomas Monz, Philipp Schindler, Berit Vogell, Marcello Dalmonte, Philipp Hauke, Rainer Blatt, and Peter Zoller. U(1) Wilson lattice gauge theories in digital quantum simulators. New J. Phys., 19(10):103020, 2017.

[75] Erez Zohar, J Ignacio Cirac, and Benni Reznik. Quantum simulations of lattice gauge theories using ultracold atoms in optical lattices. Reports on Progress in Physics, 79(1):014401, 2015.

[76] D. Marcos, P. Widmer, E. Rico, M. Hafezi, P. Rabl, U. J. Wiese, and P. Zoller. Two-dimensional lattice gauge theories with superconducting quantum circuits. arXiv:1407.606, 2014.

[77] Y. Y. Atas, J. F. Haase, J. Zhang, V. Wei, S. M.-L. Pfaendler, R. Lewis, and C.A. Muschik. Simulating one-dimensional quantum chromodynamics on a quantum computer: Real-time evolutions of tetra- and pentaquarks. arXiv:2207.03473, 2022.

[78] Arata Yamamoto. Real-time simulation of (2+1)-dimensional lattice gauge theory on qubits. Progress of Theoretical and Experimental Physics, 2021(1), 2020.

[79] D. Horn, M. Weinstein, and S. Yankielowicz. Hamiltonian approach to $Z(N)$ lattice gauge theories. Phys. Rev. D, 19:3715–3731, Jun 1979.

[80] Abhinav Kandala, Antonio Mezzacapo, Kristan Temme, Maika Takita, Markus Brink, Jerry M Chow, and Jay M Gambetta. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. Nature, 549(7671):242–246, 2017.

[81] Robert Raussendorf and Jim Harrington. Fault-tolerant quantum computation with high threshold in two dimensions. Phys. Rev. Lett., 98:190504, May 2007.

[82] M. Zwerger, H. J. Briegel, and W. Dür. Universal and optimal error thresholds for measurement-based entanglement purification. Phys. Rev. Lett., 110:260503, Jun 2013.

[83] M Zwerger, HJ Briegel, and W Dür. Hybrid architecture for encoded measurement-based quantum computation. Sci. Rep., 4(1):5364, 2014.

[84] M. Zwerger, H. Briegel, and W. Dür. Measurement-based quantum communication. App. Phys. B, 122:50, 06 2015.

[85] Xiao Yuan, Jinzhao Sun, Junyu Liu, Qi Zhao, and You Zhou. Quantum simulation with hybrid tensor networks. Phys. Rev. Lett., 127:040501, Jul 2021.

[86] Marco Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, et al. Variational quantum algorithms. Nat. Rev. Phys., 3(9):625–644, 2021.

[87] Yunseok Kwak, Won Joon Yun, Soyi Jung, and Joongheon Kim. Quantum neural networks: Concepts, applications, and challenges. In 2021 Twelfth International Conference on Ubiquitous and Future Networks (ICUFN), pages 413–416. IEEE, 2021.

[88] Amira Abbas, David Sutter, Christa Zoufal, Aurélien Lucchi, Alessio Figalli, and Stefan Woerner. The power of quantum neural networks. Nat. Comput. Sci., 1(6):403–409, 2021.

[89] Oleksandr Kyriienko and Vincent E. Elfving. Generalized quantum circuit differentiation rules. Phys. Rev. A, 104:052417, Nov 2021.

[90] David Peral García, Juan Cruz-Benito, and Francisco José García-Peñalvo. Systematic literature review: Quantum machine learning and its applications. arXiv:2201.04093, 2022.

[91] Oleksandr Kyriienko, Annie E. Paine, and Vincent E. Elfving. Protocols for trainable and differentiable quantum generative modelling. arXiv:2202.08253, 2022.

[92] Junyu Liu, Frederik Wilde, Antonio Anna Mele, Liang Jiang, and Jens Eisert. Noise can be helpful for variational quantum algorithms. arXiv:2210.06723, 2022.

[93] Junyu Liu, Francesco Tacchino, Jennifer R. Glick, Liang Jiang, and Antonio Mezzacapo. Representation learning via quantum neural tangent kernels. PRX Quantum, 3:030323, Aug 2022.

[94] Han Zheng, Gokul Subramanian Ravi, Hanrui Wang, Kanav Setia, Frederic T. Chong, and Junyu Liu. Benchmarking variational quantum circuits with permutation symmetry. arXiv:2211.12711, 2022.

[95] Annie E. Paine, Vincent E. Elfving, and Oleksandr Kyriienko. Quantum kernel methods for solving regression problems and differential equations. Phys. Rev. A, 107:032428, Mar 2023.

[96] Seth Lloyd. Universal quantum simulators. Science, 273(5278):1073–1078, 1996.

[97] Ben P Lanyon, Cornelius Hempel, Daniel Nigg, Markus Müller, Rene Gerritsma, F Zähringer, Philipp Schindler, Julio T Barreiro, Markus Rambach, Gerhard Kirchmair, et al. Universal digital quantum simulation with trapped ions. Science, 334(6052):57–61, 2011.

[98] Y. Salathé, M. Mondal, M. Oppliger, J. Heinsoo, P. Kurpiers, A. Potočnik, A. Mezzacapo, U. Las Heras, L. Lamata, E. Solano, S. Filipp, and A. Wallraff. Digital quantum simulation of spin models with circuit quantum electrodynamics. Phys. Rev. X, 5:021027, Jun 2015.

[99] Hsin-Yuan Huang, Michael Broughton, Jordan Cotler, Sitan Chen, Jerry Li, Masoud Mohseni, Hartmut Neven, Ryan Babbush, Richard Kueng, John Preskill, et al. Quantum advantage in learning from experiments. Science, 376(6598):1182–1186, 2022.

[100] Chih-Chiao Hung, Tim Kohler, and Kevin D. Osborn. Quantum defects from single surface exhibit strong mutual interactions. arXiv:2302.00318, 2023.

[101] Kishor Bharti, Alba Cervera-Lierta, Thi Ha Kyaw, Tobias Haug, Sumner Alperin-Lea, Abhinav Anand, Matthias Degroote, Hermanni Heimonen, Jakob S. Kottmann, Tim Menke, Wai-Keong Mok, Sukin Sim, Leong-Chuan Kwek, and Alán Aspuru-Guzik. Noisy intermediate-scale quantum algorithms. Rev. Mod. Phys., 94:015004, Feb 2022.

[102] P.W. Shor. Fault-tolerant quantum computation. IEEE Comput. Soc. Press, 1996.

[103] Daniel Gottesman. Theory of fault-tolerant quantum computation. Physical Review A, 57(1):127–137, 1998.

[104] Philip Walther, Kevin J Resch, Terry Rudolph, Emmanuel Schenck, Harald Weinfurter, Vlatko Vedral, Markus Aspelmeyer, and Anton Zeilinger. Experimental one-way quantum computing. Nature, 434(7030):169–176, 2005.

[105] Giulio Chiribella and Hlér Kristjánsson. Quantum shannon theory with superpositions of trajectories. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 475(2225):20180903, 2019.

[106] Richard E. Cleve. Quantum information processing - quantum information theory (i), November 2021.

[107] B. M. Escher, R. L. de Matos Filho, and L. Davidovich. General framework for estimating the ultimate precision limit in noisy quantum-enhanced metrology. Nature Physics, 7(5):406–411, 2011.

[108] Rafał Demkowicz-Dobrzański, Jan Kołodyński, and Mădălin Guţă. The elusive heisenberg limit in quantum-enhanced metrology. Nature Communications, 3(1), 2012.

[109] Man-Duen Choi. Completely positive linear maps on complex matrices. Linear Algebra and its Applications, 10(3):285–290, 1975.

[110] Sergey Bravyi and Alexei Kitaev. Universal quantum computation with ideal clifford gates and noisy ancillas. Physical Review A, 71(2), 2005.

[111] Scott Aaronson and Daniel Gottesman. Improved simulation of stabilizer circuits. Phys. Rev. A, 70:052328, Nov 2004.

[112] Harry Levine, Alexander Keesling, Giulia Semeghini, Ahmed Omran, Tout T. Wang, Sepehr Ebadi, Hannes Bernien, Markus Greiner, Vladan Vuletić, Hannes Pichler, and Mikhail D. Lukin. Parallel implementation of high-fidelity multiqubit gates with neutral atoms. Phys. Rev. Lett., 123:170503, Oct 2019.

[113] Xiu Gu, Jorge Fernández-Pendás, Pontus Vikstal, Tahereh Abad, Christopher Warren, Andreas Bengtsson, Giovanna Tancredi, Vitaly Shumeiko, Jonas Bylander, Göran Johansson, and Anton Frisk Kockum. Fast multiqubit gates through simultaneous two-qubit gates. PRX Quantum, 2:040348, Dec 2021.

[114] Yosep Kim, Alexis Morvan, Long B. Nguyen, Ravi K. Naik, Christian Jünger, Larry Chen, John Mark Kreikebaum, David I. Santiago, and Irfan Siddiqi. High-fidelity three-qubit iToffoli gate for fixed-frequency superconducting qubits. Nature Physics, 18(7):783–788, May 2022.

[115] Jorge Miguel-Ramiro, Zheng Shi, Albie Chan, Luca Dellantonio, Christine A. Muschik, and Wolfgang Dür. SQEM: Superposed quantum error mitigation. Preprint: arxiv, 2023.

[116] Muhammad Kashif and Saif Al-Kuwari. Physical realization of measurement based quantum computation. arXiv:2302.00318, 2023.

[117] J.R. Johansson, P.D. Nation, and Franco Nori. Qutip 2: A python framework for the dynamics of open quantum systems. Computer Physics Communications, 184(4):1234–1240, 2013.

[118] Cirq Developers. Cirq, December 2022. See full list of authors on Github: https://github .com/quantumlib/Cirq/graphs/contributors.

[119] Matthew Neeley, Radoslaw C Bialczak, M Lenander, Erik Lucero, Matteo Mariantoni, AD O'connell, D Sank, H Wang, M Weides, J Wenner, et al. Generation of three-qubit entangled states using superconducting phase qubits. Nature, 467(7315):570–573, 2010.

[120] Radoslaw C Bialczak, Markus Ansmann, Max Hofheinz, Erik Lucero, Matthew Neeley, Aaron D O'Connell, Daniel Sank, Haohua Wang, James Wenner, Matthias Steffen, et al. Quantum process tomography of a universal entangling gate implemented with josephson phase qubits. Nat. Phys., 6(6):409–413, 2010.

[121] Lorenza Viola and Seth Lloyd. Dynamical suppression of decoherence in two-state quantum systems. Phys. Rev. A, 58:2733–2744, Oct 1998.

[122] Lorenza Viola, Emanuel Knill, and Seth Lloyd. Dynamical decoupling of open quantum systems. Phys. Rev. Lett., 82:2417–2421, Mar 1999.

[123] Lorenza Viola, Seth Lloyd, and Emanuel Knill. Universal control of decoupled quantum systems. Phys. Rev. Lett., 83:4888–4891, Dec 1999.

[124] H. Y. Carr and E. M. Purcell. Effects of diffusion on free precession in nuclear magnetic resonance experiments. Phys. Rev., 94:630–638, May 1954.

[125] M. J. D. Powell. A direct search optimization method that models the objective and constraint functions by linear interpolation. In S. Gomez and J.-P. Hennart, editors, Advances in Optimization and Numerical Analysis, volume 275 of Mathematics and Its Applications, pages 51–67. Springer, 1994.

[126] D. R. Jones, C. D. Perttunen, and B. E. Stuckman. Lipschitzian optimization without the Lipschitz constant. Journal of Optimization Theory and Applications, 79:157–181, 1993.

[127] Steven G. Johnson. The NLopt nonlinear-optimization package. https://github.com/stevengj/nlopt, 2007.

[128] Miroslav Urbanek, Benjamin Nachman, Vincent R. Pascuzzi, Andre He, Christian W. Bauer, and Wibe A. de Jong. Mitigating depolarizing noise on quantum computers with noise-estimation circuits. Physical Review Letters, 127(27), 2021.

[129] The Qiskit Nature developers and contributors. Qiskit Nature 0.6.0, April 2023.

[130] Sergey B. Bravyi and Alexei Yu. Kitaev. Fermionic quantum computation. Ann. Phys., 298(1):210–226, 2002.

[131] Jiaxin Zhang. Modern Monte Carlo methods for efficient uncertainty quantification and propagation: A survey. arXiv:2011.00680, 2020.

[132] Nicolai Friis, Vedran Dunjko, Wolfgang Dür, and Hans J. Briegel. Implementing quantum control for unknown subroutines. Physical Review A, 89(3), 2014.

APPENDICES

Appendix A

Overview of Measurement-Based Quantum Computing

In this appendix, we give a more detailed description of the MB-QC protocol. At the outset, one prepares an entangled resource state – a graph or cluster containing input (I), body (B) and output (O) qubits connected via CZ-edges. The body and output qubits are initialized in $|+\rangle$, while the input qubits encode the input state $|\psi_{\text{in}}\rangle$. The initial state of the resource $|G\rangle$ is therefore:

$$|G\rangle = \prod_{\{i,j\}\in E} CZ_{i,j}\left(|\psi_{\text{in}}\rangle_I \bigotimes_{m\in\{B,O\}} |+\rangle_m \right), \tag{A.1}$$

where E is the set of prescribed edges. This is a $+1$ eigenstate of the operators $\hat{S}_n$ (i.e. $\hat{S}_n|G\rangle = |G\rangle$), where

$$\hat{S}_n = \hat{X}_n \prod_{k\in\mathcal{N}(n)} \hat{Z}_k, \tag{A.2}$$

and $\mathcal{N}(n)$ are the neighbours of qubit n (i.e. those connected to qubit n via a CZ-edge).

The main computation is driven by a series of projective measurements on the input and body qubits, typically in the eigenbasis of the Pauli operators $\hat{X}$, $\hat{Y}$, $\hat{Z}$, or the rotated basis $R(\theta) = \left\{ \frac{|0\rangle + e^{i\theta}|1\rangle}{\sqrt{2}}, \frac{|0\rangle - e^{i\theta}|1\rangle}{\sqrt{2}} \right\}$. Conventionally, all input qubits I are chosen to be measured in the X-basis [2]. Each measurement applied to a qubit removes it from the resource state and modifies the remaining connections depending on the prescribed basis. Concurrently, $|\psi_{\text{in}}\rangle$ is propagated through each body qubit (known as feedforward), and

upon reaching the output qubits it will have undergone a linear transformation V:

$$V : |\psi_{\text{in}}\rangle \rightarrow |\psi_{\text{out}}\rangle : \mathcal{H}_I \rightarrow \mathcal{H}_O, \tag{A.3}$$

where $\mathcal{H}_I$ ($\mathcal{H}_O$) denotes the Hilbert space spanned by the input (output) qubits, and $|I| = |O|$ for a unitary transformation. To ensure the result $|\psi_{\text{out}}\rangle$ is deterministic, i.e., the same for each possible set of random measurement outcomes, corrective operations are applied. These comprise byproduct Pauli operators $\hat{U}_\Sigma \in \{\hat{I}, \hat{X}, \hat{Y}, \hat{Z}\}^{\otimes|O|}$ acting on $|\psi_{\text{out}}\rangle$ and angle adaptations $M_i(R(\theta)) \rightarrow M_i(R(\pm\theta))$ for qubits i measured in non-Clifford bases (i.e. $\theta \not\equiv 0 \bmod \pi/2$). The angle adaptations require knowledge of prior outcomes, implying that a temporal (causal) order for measuring the adaptive qubits must exist for a pattern to yield deterministic results. In other words, an adaptive qubit may be measured only after all measurements that it depends on have taken place. The projective measurements ($\hat{P}$), the byproduct operators and the angle adaptations form an MB *pattern* realizing V, with

$$|\psi_{\text{out}}\rangle = V|\psi_{\text{in}}\rangle = \hat{U}_\Sigma \hat{P}|G\rangle. \tag{A.4}$$

Elementary one- and two-logical qubit gates correspond to well-known patterns [2], which are sufficient for universal computation. An arbitrary gate sequence may be achieved by concatenating each output of a pattern to the corresponding input of the next [2].

Example. To illustrate the MB-QC procedure, we will derive and verify the MB pattern corresponding to the $\hat{Z}^{\otimes 3}(\theta)$ operation, which is combined with rotation gates to form the $\hat{\mathcal{P}}^{\otimes 3}$ gadget. In the GB paradigm, it realizes the unitary:

$$\begin{aligned}
\hat{U}_{GB}(\theta) &= CX_{1,2}CX_{2,3}RZ_3(\theta)CX_{2,3}CX_{1,2} \\
&= diag(e^{-i\theta/2}, e^{i\theta/2}, e^{i\theta/2}, e^{-i\theta/2}, e^{i\theta/2}, e^{-i\theta/2}, e^{-i\theta/2}, e^{i\theta/2}).
\end{aligned} \tag{A.5}$$

From the gate sequence in Eq. (A.5), we may construct an equivalent MB pattern via concatenation, which yields a 31-qubit pattern [Fig. A.1]. (Note CX can be expressed in terms of CZ and H operations.) To reduce the number of initial qubits, we apply graph-theoretic simplifications to the pattern via the ZX-calculus (Ref. [10]), which enables a reduction from 31 to 15 qubits. The simplified pattern is shown in Fig. A.1, where the edges are:

$$\begin{aligned}
E = \{&(4,5), (6,7), (13,8), (9,4), (9,12), (12,5), (5,14), (15,10), (4,11), (1,11), (6,5), (3,7), \\
&(9,6), (6,10), (4,8), (2,9)\}.
\end{aligned} \tag{A.6}$$

The relation between input and output states is therefore:

$$|\psi_{out,MB}\rangle_{13,14,15} = \hat{P}\hat{U}_\Sigma|G\rangle_{1,2,...,15} = \hat{P}\hat{U}_\Sigma\left[\prod_{\{i,j\}\in E} CZ_{i,j}\left(|G_{\psi_{in,MB}}\rangle_{1,2,3} \otimes |+\rangle_{4,5,...,15}^{\otimes 12}\right)\right],$$
(A.7)

where $|G\rangle$ is the term in square brackets. By convention, we measure the input qubits first, followed by the non-adaptive (Clifford) and adaptive (non-Clifford) qubits. Since the inputs and non-adaptive qubits are not affected by the temporal ordering, they may be measured simultaneously. To calculate $\hat{U}_\Sigma$, we solve eigenvalue equations (Eq. (82) of Ref. [2]) for each qubit i. These equations are a direct consequence of Eq. (A.2), and determine whether to apply $\hat{X}_i$, $\hat{Z}_i$ or both on qubit i if it measures 1 (denoted by s_i). A similar step is performed to solve for the adaptive corrections on qubit 12, which indicate which outcomes of qubit 1–11 require the measurement angle of qubit 12 to be flipped (i.e. $\theta \to -\theta$). Here, we must obey the temporal ordering by considering *only* qubits that have been measured up to that point (i.e. qubits 1–11). For the pattern in Fig. A.1, we find the projective measurements and byproducts to be:

$$\hat{P} = M_{12}\big(R(\theta(-1)^{s_5+s_8+s_{10}})\big)\prod_{k=1}^{11}M_k(X),$$
(A.8)

$$\hat{U}_\Sigma = \hat{Z}_{15}^{s_3+s_6+s_{12}}\hat{X}_{15}^{s_7+s_{10}}\hat{Z}_{14}^{s_2+s_{12}}\hat{X}_{14}^{s_5+s_9}\hat{Z}_{13}^{s_1+s_4+s_{12}}\hat{X}_{13}^{s_8+s_{11}},$$
(A.9)

where M_k is the measurement operator on qubit k and $s_k \in \{0,1\}$ its corresponding outcome. Because of the outcome dependence of $\hat{U}_\Sigma$, they must be applied *after* the measurements and therefore require commutation over the operations in Eq. (A.5), which V implements. Here, we note that the diagonal form of the unitary implies that $\hat{U}_\Sigma$ remains unchanged after commutation [32]; however, the adaptive corrections on qubit 12 are affected. Eqs. (A.7) and (A.8) become:

$$|\psi_{out,MB}\rangle_{13,14,15} = \hat{U}_\Sigma\hat{P}'|G\rangle_{1,2,...,15},$$
(A.10)

$$\hat{P}' = M_{12}\big(R(\theta(-1)^{s_7+s_9+s_{11}})\big)\prod_{k=1}^{11}M_k(X),$$
(A.11)

where we have used the commutation relations in Eqs. (72) and (73) of Ref. [2]. For simplicity, we consider a scenario where each qubit of the input state is a Pauli eigenvector $|+\rangle$; however, any valid measurement pattern can process an arbitrary input. We discuss how to address such cases in App. F.

69
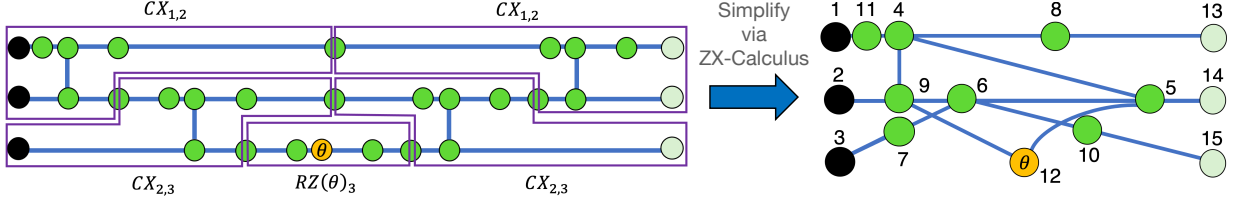
FIG. A.1: **MB pattern for $Z^{\otimes 3}(\theta)$, obtained by concatenating smaller patterns corresponding to each operation in Eq.** (A.5)**.** To reduce the number of initial qubits, the pattern is simplified via the techniques of Ref. [10]. Black and light green circles denote input and output qubits respectively, blue lines denote CZ-edges, green circles denote nonadaptive qubits measured in the X-basis, and orange circles denote adaptive qubits measured in the $R(\theta)$-basis. The numbers label the qubits and their measurement order (excluding outputs).

One may now proceed and perform Eq. (A.11) directly. However, since the majority of qubits (1–11) are non-adaptive, they can be efficiently simulated classically via the stabilizer formalism of Ref. [111]. To do this, we employ the tableau representation of stabilizer states and track the input state as it undergoes Clifford operations and measurements. For all qubits initialized in $|+\rangle$, the stabilizer state is given by:

$$\{(+\hat{X}_1)...(+\hat{X}_{15})\}. \tag{A.12}$$

After performing all CZs and non-adaptive (X-basis) measurements in Eq. (A.11), the canonical form stabilizers over the remaining adaptive and output qubits (assuming an outcome of 0 for qubits 1–11) are:

$$\{(+\hat{Z}_1), (+\hat{Z}_2), ..., (+\hat{Z}_{11}), (+\hat{X}_{12}\hat{X}_{15}), (+\hat{Z}_{12}\hat{Z}_{13}\hat{Z}_{14}\hat{Z}_{15}), (+\hat{X}_{13}\hat{X}_{15})\}, \tag{A.13}$$

where elementary row operations were used to reduce each of qubits 1–11 to a single $+Z$. These stabilizers, corresponding to the post-measured state $|0\rangle$ may be subsequently traced out from the overall state. Using the Gottesman-Knill theorem [111], we then work out a unitary $\hat{U}_{\text{Cliff}}$ that produces a *reduced* graph state stabilized by the operators above. To achieve the form of Eq. (A.2) (which describes the connectivity of graphs), we express Eq. (A.13) as a generator matrix over qubits 12–15 and reduce the stabilizer ($\hat{X}$) portion to row-echelon form, where each manipulation incurs a local Clifford (LC) operation H or

70

S (see Ref. [111]). The resulting stabilizers and circuit are:

$$
\begin{matrix}
\hat{Z}_1 \\
\hat{Z}_2 \\
\hat{Z}_3 \\
\hat{Z}_4 \\
\hat{X}_1 \\
\hat{X}_2 \\
\hat{X}_3 \\
\hat{X}_4
\end{matrix}
\begin{bmatrix}
0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
1 & 0 & 1 & 1
\end{bmatrix}
\xrightarrow{H_4}
\begin{matrix}
\hat{Z}_1 \\
\hat{Z}_2 \\
\hat{Z}_3 \\
\hat{Z}_4 \\
\hat{X}_1 \\
\hat{X}_2 \\
\hat{X}_3 \\
\hat{X}_4
\end{matrix}
\begin{bmatrix}
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 \\
1 & 1 & 1 & 0 \\
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\tag{A.14}
$$

$$
\rightarrow \{(+\hat{X}_1\hat{Z}_4), (+\hat{X}_2\hat{Z}_4), (+\hat{X}_3\hat{Z}_4), (+\hat{X}_4\hat{Z}_1\hat{Z}_2\hat{Z}_3)\} \tag{A.15}
$$

$$
\implies \hat{U}_{\text{Cliff}} = H_4 CZ_{1,4} CZ_{1,3} CZ_{1,2}, \tag{A.16}
$$

where H_4 is a LC-operation and we have relabelled qubits 13, 14, 15, 12 as 1, 2, 3, 4 respectively. Consequently, V reduces to a 4-qubit ($3 + 1$ ancilla) pattern, and we have:

$$
\begin{aligned}
|\psi_{out,MB}\rangle_{1,2,3} &= V_{reduced}\hat{U}_{\text{Cliff}}|+\rangle_{1,2,3,4}^{\otimes 4} \\
&= (\hat{Z}_3\hat{Z}_2\hat{Z}_1)^{s_4} M_4\big(R(\theta(-1)^{s_4})\big)\hat{U}_{\text{Cliff}}|+\rangle_{1,2,3,4}^{\otimes 4}.
\end{aligned}
\tag{A.17}
$$

One observes that the graph structure corresponding to Eq. (A.15) resembles that shown in Fig. 2.1(a). If qubit 4 is measured with outcome 1 ($s_4 = 1$), we obtain the final output:

$$
\begin{aligned}
|\psi_{out,MB}\rangle_{1,2,3} &= \frac{1}{\sqrt{8}}\hat{Z}_3\hat{Z}_2\hat{Z}_1\Big[e^{-i\theta/2}\sum_{x\in\{b_1...b_3\in\vec{b}|b_1+...+b_3=0\}}|x\rangle - e^{i\theta/2}\sum_{x'\in\{b_1...b_3\in\vec{b}|b_1+..+b_3=1\}}|x'\rangle\Big] \\
&= \frac{1}{\sqrt{8}}\Big[e^{-i\theta/2}\sum_{x\in\{b_1...b_3\in\vec{b}|b_1+...+b_3=0\}}|x\rangle + e^{i\theta/2}\sum_{x'\in\{b_1...b_3\in\vec{b}|b_1+...+b_3=1\}}|x'\rangle\Big],
\end{aligned}
\tag{A.18}
$$

where $\vec{b} = \{0,1\}^{\otimes 3}$. It is straightforward to verify that the state is equivalent (up to a global phase) to $\hat{U}_{GB}$ acting on $|+\rangle^{\otimes 3}$.

In summary, MB-QC constitutes a natural method of computation for platforms that implement highly entangled graph or cluster states, including photonic and spin qubits. Despite the larger number of ancillary qubits employed in comparison to GB-QC, the method is beneficial when the resource state contains a large Clifford portion, as shown in

the example above. By performing the majority of the pattern classically, one drastically reduces the size of the resource state to be implemented. On a practical level, this can yield fewer computational steps, and in noisy cases, opportunities for the overall resource to decohere.

Appendix B

Hybrid VQE Demonstrations – Methods and Techniques

In this appendix, we discuss the numerical program and associated methods used to perform VQE simulations on IBM Quantum systems. We also detail the error mitigation and suppression techniques implemented to combat noise and decoherence effects prevalent on real hardware. A schematic of the entire program is depicted in Fig. B.1.

Setup. The program takes in three main inputs, which are specified by the user:

1. A target Hamiltonian $\hat{\mathcal{H}}$ describing the physical system, expressed as a sum of Pauli interactions [Eq. (2.7)].

2. A parameterized ansatz circuit that outputs $|\psi_a^{(\theta)}\rangle$ to compute the mean energy $E = \langle\psi_a|\hat{\mathcal{H}}|\psi_a\rangle$.

3. A starting guess for the variational parameters $(\vec{\theta}^{(i)})$.

To assess VQE performance, we use the above to calculate the exact GS ($|\psi_0\rangle$) and GS energy $E_0 = \langle\psi_0|\mathcal{H}|\psi_0\rangle$ via exact diagonalization (ED). We note that this is feasible only for small to intermediate system sizes.

Measurement protocol. To calculate the total energy, we group all observables in $\hat{\mathcal{H}}$ into n commuting sets g_n, where all commuting observables within a set are measured _simultaneously_ by a single circuit (for a total of n). For example, in the $\mathbb{Z}_2$ demonstration, we require two circuits — one measuring $XXXX$ and the other measuring $ZZZZ$, since they commute with all terms in $\hat{\mathcal{H}}_\square$ and $\hat{\mathcal{H}}_\triangle$ respectively.

Choosing measurement shot numbers. Given the iterative nature of VQE, a practical approach involves minimizing the allocated budget, which encompasses the number of optimizer iterations/function evaluations and the total number of measurement shots over all circuit executions. While there are standard choices for the former, selecting an appropriate budget for the latter while balancing performance is less straightforward. Here, we take an adaptive approach based on the size of the first energy gap $E_g = E_1 - E_0$. We treat E_g as an indicator of the optimization difficulty in that the smaller it is, the more statistical counts are needed to resolve the difference. In this sense, we may regard E_g as the standard deviation over all i commuting sets $g_i \in \hat{\mathcal{H}}$, weighted by their respective shots N_{g_i}:

$$E_g = \sqrt{\frac{\langle \hat{\mathcal{H}}_{g_1} \rangle^2}{N_{g_1}} + \frac{\langle \hat{\mathcal{H}}_{g_2} \rangle^2}{N_{g_2}} + ... + \frac{\langle \hat{\mathcal{H}}_{g_i} \rangle^2}{N_{g_i}}}. \tag{B.1}$$

We also assume that the variation due to g_i is weighted by the sum of the n coefficients corresponding to the n Pauli terms in the group:

$$\langle \hat{\mathcal{H}}_{g_i} \rangle^2 = \sum_n (g_i^{(n)})^2. \tag{B.2}$$

To minimize the L.H.S, we require that all terms in Eq. (B.1) are equal:

$$\frac{\sum g_1^2}{N_{g_1}} = \frac{\sum g_2^2}{N_{g_2}} = ... = \frac{\sum g_i^2}{N_{g_i}}. \tag{B.3}$$

For practical purposes, we run each circuit with the same number of shots, and so the problem reduces to optimizing the total number $N_{tot} = N_{g_1} + N_{g_2} + ... + N_{g_i}$. One can employ a classical optimizer to solve for each N_{g_i}; however, in the case of two groups, the problem may be solved directly and we find that:

$$N_{g_1} = RN_{g_2}; \quad R = \frac{\sum g_1^2}{\sum g_2^2} \tag{B.4}$$

$$\implies N_{g_1} = \frac{R}{1+R} N_{tot} \implies E_g = \sqrt{2 \left(\frac{1+R}{RN_{tot}} \right) \sum g_1^2} \tag{B.5}$$

$$\implies N_{tot} = 2 \left(\frac{1+R}{R(E_g)^2} \right) \sum g_1^2. \tag{B.6}$$

We then extend Eq. (B.6) directly to the general case, and replace the prefactor with a generic scaling factor s that can be chosen depending on the range of shots desired. We

also drop the weighting factor $(1+R)/R$ and match N_{tot} to the commuting group with the largest coefficient sum. Thus, we have:

$$N_{tot} = \left\lceil \frac{s}{(E_g)^2} \max_{g \in G} \left(\sum_i (c_g^{(i)})^2 \right) \right\rceil, \tag{B.7}$$

where G is the set of all commuting groups belonging to $\hat{\mathcal{H}}$ and $c_g^{(i)}$ is the coefficient of the ith Pauli in group g. To ensure a representative enough sample, we impose that N_{tot} is at least 250. Likewise, N_{tot} is capped at 5×10^4 to avoid large processing demands on the hardware. Any calculated N_{tot} falling outside these ranges are set to the closer of the two limits.

Readout mitigation. This is a standard technique to reduce errors from readout measurements [119, 120], which yield the counts data for calculating observable expectation values. The main principle involves capturing the extent of the readout errors via a *calibration matrix* M. Each element in M is the probability of initializing a computational basis state $|x\rangle$ and subsequently measuring $|x'_{\text{meas}}\rangle$ (i.e. $|\langle x'_{\text{meas}}|x\rangle|^2$). This requires the execution of 2^n basis circuits, where n is the number of qubits in $\hat{\mathcal{H}}$. For example, M has the following structure for $n = 3$ (acquired from a sample run on `ibm_perth`):

	$|000\rangle$	$|001\rangle$	$|010\rangle$	$|011\rangle$	$|100\rangle$	$|101\rangle$	$|110\rangle$	$|111\rangle$
$\langle 000|$	0.95446	0.02772	0.01029	0.00000	0.00980	0.00000	0.00000	0.00000
$\langle 001|$	0.02178	0.95248	0.00000	0.01437	0.00000	0.00596	0.00000	0.00000
$\langle 010|$	0.01386	0.00000	0.97942	0.01848	0.00000	0.00000	0.00648	0.00000
$\langle 011|$	0.00000	0.01386	0.00617	0.95277	0.00000	0.00000	0.00000	0.00386
$\langle 100|$	0.00594	0.00000	0.00000	0.00000	0.95686	0.02187	0.01296	0.00000
$\langle 101|$	0.00000	0.00396	0.00000	0.00000	0.02157	0.94433	0.00000	0.01544
$\langle 110|$	0.00000	0.00000	0.00000	0.00000	0.00784	0.00199	0.96544	0.00579
$\langle 111|$	0.00000	0.00000	0.00000	0.01027	0.00000	0.02187	0.01296	0.96525

$$\tag{B.8}$$

For low noise levels, M is expected to be predominantly diagonal and close to the identity (as is shown above). Furthermore, if ancillary qubits are present (i.e. in the Pauli gadget demonstrations), we determine a separate calibration matrix over n qubits for each possible set of ancillary outcomes.

If C_{ideal} is the counts distribution corresponding to the exact E_0, then MC_{ideal} leads to an approximation of the noisy distribution C_{noisy}. Naturally, we wish to perform the reverse. Näively, this is achieved by applying $M^{-1}C_{noisy}$, where M^{-1} is the inverse of M; however, it can yield negative counts (quasiprobabilities). We remedy this by performing

a least-square fit to the closest valid distribution. After the calibration matrix is determined, it is stored and applied to the counts distribution from every subsequent circuit evaluation. This is most ideal when the hardware-specific parameters (i.e. error rates, decoherence times) are constant; however, as drifts in the parameters over time are expected, we determine M more than once (i.e. after each data point) to ensure the errors are accurately captured throughout the course of the VQE run. In all demonstrations, we employ 10^4 shots for each basis circuit used to determine M.

VQE algorithm. As shown in Fig. B.1, the VQE proceeds as a closed feedback loop between the quantum processor and classical optimizer. Prior to circuit execution, we apply various techniques to suppress noise:

1. Performing the circuits requires one to specify a mapping between virtual and physical hardware qubits. Since the latter is restricted by the qubit topology, SWAP operations (each incurring three CXs) are added to match the connectivity of the considered circuit. With CXs being the most expensive operation on IBM Quantum systems, we choose mappings that lead to the fewest number of added SWAPs. At the time of the simulation run, we also ensure that the chosen qubits have reasonably small CX errors and long decoherence times.

2. Resets and delays are added to the beginning of each circuit. This improves the likelihood of each qubit being properly initialized in the ground state $|0\rangle$.

3. On the pulse level, we schedule circuits to incorporate dynamical decoupling (DD) [121, 122, 123]. This mitigation technique inserts extra pulses (equivalent to the identity) whenever a particular qubit is idle. Although it introduces further noise, the periodicity of the applied pulses enable undesired state rotations from system-environmental interactions to be reversed. Here, we employ X-X pulses, which are commonly employed in spin-refocusing techniques to increase the coherence time of qubits [124].

During the VQE, the ansatz circuit is sent to and executed on the IBM Quantum hardware, which outputs results in the form of counts data. From the counts, we compute the total energy E and contributions $\langle \hat{O}_1 \rangle, \langle \hat{O}_2 \rangle ... \langle \hat{O}_N \rangle$ from specific observables of interest. The value of E is then sent to a classical optimizer, which modifies $\vec{\theta}$ with the goal of minimizing E. The above process repeats until a stopping condition for convergence is met by the optimizer, at which point the final energy E_{opt} and parameters $\vec{\theta}^{(opt)}$ are returned.

All circuits are performed on IBM Quantum systems, with readout measurements chosen depending on the desired basis. Furthermore, all circuits employing Pauli gadgets
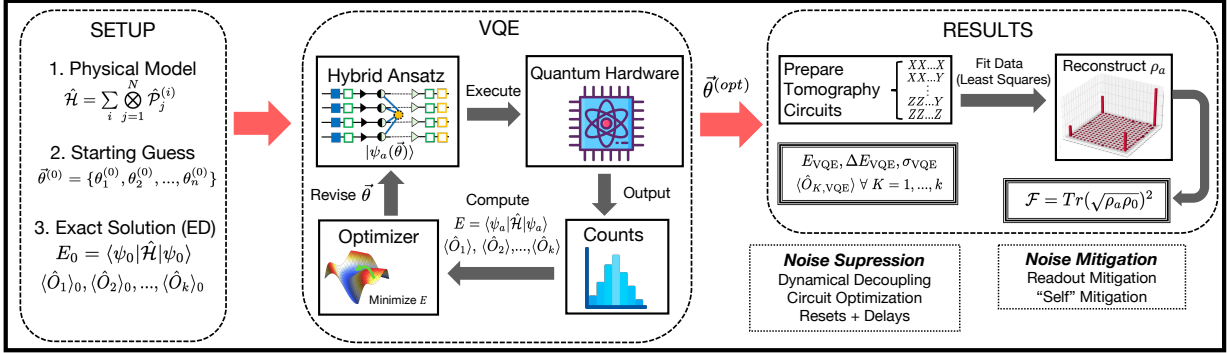
FIG. B.1: **Complete workflow for the numerical VQE program.** Readout mitigation and all noise suppression techniques are employed during execution of each variational circuit, while self-mitigation occurs immediately after $\vec{\theta}^{(opt)}$ is determined.

are performed *dynamically*. These incorporate added `if_test` instructions in the circuit, which act corrective byproduct ($\hat{U}_\Sigma$) and adaptive operations (see App. A) on the main qubits based on the outcomes of the ancillary qubits measured previously.[1] This enables one in practice to perform MB-QC with a circuit (i.e. as opposed to clusters or graphs). By convention, the byproduct operator for a Pauli gadget only acts nontrivially when the ancilla measurement outcome is '1'. Fig. B.2 depicts the dynamic circuit used in the $\mathbb{Z}_2$ demonstration.

We employ the COBYLA and DIRECT methods for the classical optimization. COBYLA [125] enables a rapid and expansive search of the energy landscape with few iterations, while DIRECT [126] enables a more systematic search, identifying promising areas that may contain the global minimum. When using the latter, we follow the recommendation in Ref. [127] and perform a second, local minimization (i.e. COBYLA) afterwards to refine the search in the most promising region. While SPSA is often recommended for noisy simulations [80, 92], we avoid its use as it generally requires large iteration numbers to converge. Indeed, our circuits have reasonably small depths and few parameters such that they achieve a low-enough level of noise for other optimizers to be effective.

Self-mitigation. At the conclusion of the VQE run, we employ the self-mitigation technique [13, 77, 128]. It specifically addresses CX errors by evaluating the ansatz circuit in two ways. First, there is a physical run using the circuit parameterized with $\vec{\theta}^{(opt)}$, then there is a mitigation run which involves a modified version of the physical as reference. In both runs, each CX is dressed with randomly chosen single-qubit operations (see Ref. [77]

---

[1]The adaptive bases discussed in App. A are not required for the Pauli gadget, provided the byproduct operators are applied immediately after each gadget [32].
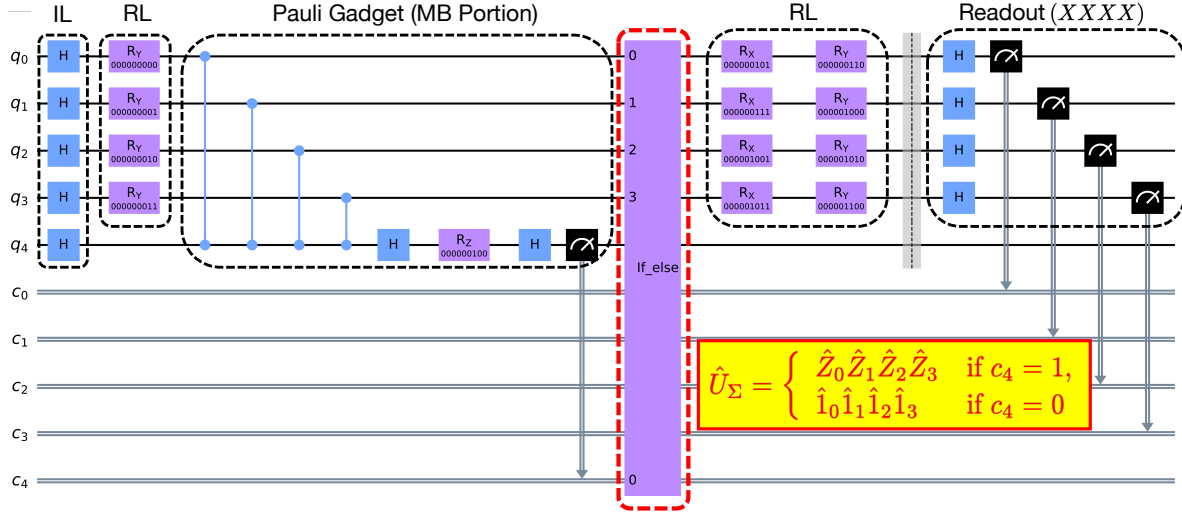
FIG. B.2: **Sample dynamic circuit performed in the $\mathbb{Z}_2$ VQE demonstration.** The qubits and corresponding classical registers for storing measurement outcomes are labelled by q and c respectively. After the mid-circuit measurement of the ancilla (q_4), the dynamic `if_else` instruction (highlighted in red) implements the byproduct operators $\hat{U}_\Sigma = (\hat{Z}_0 \hat{Z}_1 \hat{Z}_2 \hat{Z}_3)^{c_4}$ corresponding to the $\hat{Z}^{\otimes 4}$ operation (see App. A). The measurement basis of the circuit is $XXXX$.

for details). These do not alter the action of the CX operations (they only change the basis), but instead transform the CX noise from being coherent to incoherent in nature. The correction acts as a scale factor on any given Pauli-string observable $\langle \hat{O} \rangle$, based on the results of evaluating the physical and mitigation runs. It is given by:

$$\langle \hat{O} \rangle_{\text{phys, true}} = \langle \hat{O} \rangle_{\text{phys, meas}} \times \left(\frac{\langle \hat{O} \rangle_{\text{mitig, true}}}{\langle \hat{O} \rangle_{\text{mitig, meas}}} \right)^\kappa, \tag{B.9}$$

where κ is the ratio of CX gate numbers between the physical and mitigation circuits, and "true" ("meas") refers to the exact (noisy) expectation values. To minimize added resources, all observables that commute with the measurement basis of the ansatz circuit are corrected with the same factor. For simplicity, we also formulate our mitigation circuits so that $\kappa = 1$ and $\langle \hat{O} \rangle_{\text{mitig, true}} = 1$. We achieve $\kappa = 1$ by including only the CXs present in the ansatz circuit [128]. In both cases, a suitable initial state is chosen to ensure $\langle \hat{O} \rangle_{\text{mitig, true}} = 1$, with LC-operations added where necessary.

State fidelity. In the final step, we perform full state tomography to determine the ansatz state corresponding to $\vec{\theta}^{(opt)}$. This requires the evaluation of 3^n circuits (i.e. all

n-qubit combinations of X, Y, and Z Paulis). In all demonstrations, we employ 10^4 shots for each circuit evaluation. If ancillas are present, their counts are marginalized over the counts of the readout qubits. The 3^n counts distributions are then fitted to a valid density matrix ρ_a using maximum likelihood estimation, which corresponds to the reconstructed ansatz state. This enables us to compute the state fidelity $\mathcal{F}$ between ρ_a and the exact GS (ρ_0) as

$$\mathcal{F} = \text{Tr}(\sqrt{\rho_a \rho_0})^2, \tag{B.10}$$

where Tr is the trace operation. As with exact diagonalization, calculating $\mathcal{F}$ via tomography is practical only for small system sizes owing to the exponential scaling.

Appendix C

Hybrid VQE Demonstrations – Supplementary Information

This appendix presents target Hamiltonians, ground state properties, and raw data from the hybrid VQE demonstrations on IBM Quantum systems (Ch. 2). All demonstrations employed single-layer ansatz modifications ($L = 1$) or single Pauli gadgets ($g = 1$).

Notes:

1. For all data presented, we refer to App. B for a discussion of how the measurement shots were selected. All scaling factors s in Eq. (B.7) were chosen heuristically based on prior trial-and-error VQE simulation runs, combined with the optimization difficulty of the considered model.

2. N_{CX} refers to the number of CX gates in a VQE circuit after its qubits were mapped to the respective IBM system.

3. Unless otherwise stated, all reported energies are unitless and may be expressed in absolute units (e.g. eV, J, Ha) using a scaling factor with appropriate dimensions.

1D QCD: $\hat{\mathcal{H}}$ defined by Eq. (2.9); see Ref. [77] for ground state properties.

PC: $\hat{\mathcal{H}}$ defined by Eq. (2.4) ($M = 2, N = 1$); see App. E for a perturbative analysis of ground state properties.

$\mathbb{Z}_2$ LGT, single plaquette:

$$\hat{\mathcal{H}}_{\mathbb{Z}_2} = \lambda \hat{X}_1 \hat{X}_2 \hat{X}_3 \hat{X}_4 + \frac{1}{\lambda} \sum_{i=1}^{4} \hat{Z}_i \tag{C.1}$$

$$|\Psi_0\rangle = \sqrt{\frac{1}{2} + \frac{2}{\sqrt{16+\lambda^4}}} \left(\frac{4 - \sqrt{16+\lambda^4}}{\lambda^2} |0\rangle^{\otimes 4} + |1\rangle^{\otimes 4} \right), \quad E_0 = -\sqrt{\frac{16}{\lambda^2} + \lambda^2} \tag{C.2}$$

LiH:

$$|\Psi_0\rangle \approx 0.9877 \, |0011\rangle - 0.1154 \, |1100\rangle, \quad E_0 = -7.881072044030926 \tag{C.3}$$

Hamiltonian derived from Qiskit VQE tutorial [41, 129] (shown in Table C.1), with parity transformation to map fermionic to spin qubits [130]. State simplifications include freezing core orbitals [80] to reduce the number of initial qubits from 12 to 4.

Table C.1: 4-qubit Hamiltonian for the LiH molecule. All Pauli terms are weighted by the coefficients shown and summed.

$IIIZ$	$IIZX$	$IIIX$	$IIXX$	$IIYY$	$IIZZ$	$IIXZ$	$IIXI$	$IIZI$	$IZII$	$ZXII$	$IXII$
-0.0938	-0.00318	0.00318	-0.00125	0.00125	-0.212	0.0192	0.0192	0.358	0.0938	0.00318	0.00318
$XXII$	$YYII$	$ZZII$	$XZII$	$XIII$	$ZIII$	$IZIZ$	$IZZX$	$IZIX$	$IZXX$	$IZYY$	$IXIZ$
-0.00125	0.00125	-0.212	-0.0192	0.0192	-0.358	-0.122	0.0121	-0.0121	0.0317	-0.0317	0.0121
$ZXIZ$	$IXZX$	$ZXZX$	$IXIX$	$ZXIX$	$IXXX$	$ZXXX$	$IXYY$	$ZXYY$	$YYIZ$	$XXIZ$	$YYZX$
0.0121	-0.00327	-0.00327	0.00327	0.00327	-0.00865	-0.00865	0.00865	0.00865	0.0317	-0.0317	-0.00865
$XXZX$	$YYIX$	$XXIX$	$YYXX$	$XXXX$	$YYYY$	$XXYY$	$ZZIZ$	$ZZZX$	$ZZIX$	$ZZXX$	$ZZYY$
0.00865	0.00865	-0.00865	-0.031	0.031	0.031	-0.031	0.0559	0.00187	-0.00187	0.0031	-0.0031
$XIIZ$	$XZIZ$	$XIZX$	$XZZX$	$XIIX$	$XZIX$	$XIXX$	$XZXX$	$XIYY$	$XZYY$	$ZIIZ$	$ZIZX$
0.0128	-0.0128	-0.00235	0.00235	0.00235	-0.00235	-0.00798	0.00797	0.00797	-0.00797	0.113	-0.0108
$ZIIX$	$ZIXX$	$ZIYY$	$IZZZ$	$IZXZ$	$IZXI$	$IXZZ$	$ZXZZ$	$IXXZ$	$ZXXZ$	$IXXI$	$ZXXI$
0.0108	-0.0336	0.0336	-0.0559	-0.0128	-0.0128	-0.00187	-0.00187	0.00235	0.00235	0.00235	0.00235
$YYZZ$	$XXZZ$	$YYXZ$	$XXXZ$	$YYXI$	$XXXI$	$ZZZZ$	$ZZXZ$	$ZZXI$	$XIZZ$	$XZZZ$	$XIXZ$
-0.0031	0.0031	0.00798	-0.00798	0.00798	-0.00798	0.0845	-0.00899	-0.00899	-0.00899	0.00899	0.00661
$XZXZ$	$XIXI$	$XZXI$	$ZIZZ$	$ZIXZ$	$ZIXI$	$IZZI$	$IXZI$	$ZXZI$	$YYZI$	$XXZI$	$ZZZI$
-0.00661	0.00661	-0.00661	0.0604	0.011	0.011	0.113	-0.0108	-0.0108	-0.0336	0.0336	-0.0604
$XIZI$	$XZZI$	$ZIZI$	$IIII$								
-0.011	-0.011	-0.113	-7.012								

Table C.2: Optimized energies, operator expectation values, fidelities, and errors for the $\mathbb{Z}_2$ VQE demonstration ($s = 100$).

λ	System	Optimizer	Shots, Iters.	N_{CX}	κ	$E_{\mathrm{VQE}}^{(i)}$	$E_{\mathrm{VQE}}^{(f)}$	E_0	E_g
0.5	ibm_peekskill	COBYLA	13857, 50	7	1	-3.958	-7.427	-8.016	3.984
0.63	ibm_peekskill	COBYLA	16267, 50	7	1	-3.162	-6.036	-6.380	3.144
0.85	ibm_peekskill	COBYLA	22187, 50	7	1	-2.197	-4.547	-4.782	2.280
1.12	ibm_peekskill	COBYLA	34372, 50	7	1	-1.651	-3.436	-3.743	1.635
1.52	ibm_peekskill	COBYLA	50000, 50	7	1	-1.157	-2.735	-3.039	1.029
1.98	ibm_peekskill	COBYLA	50000, 50	7	1	-0.793	-2.548	-2.829	0.606
2.33	ibm_peekskill	COBYLA	50000, 50	7	1	-0.824	-2.659	-2.894	0.411
2.65	ibm_peekskill	COBYLA	50000, 50	7	1	-0.587	-2.784	-3.050	0.294
2.88	ibm_peekskill	COBYLA	50000, 50	7	1	-0.604	-2.896	-3.197	0.235
3.3	ibm_peekskill	COBYLA	50000, 50	7	1	-0.392	-3.104	-3.516	0.160

λ	$\langle \hat{\square} \rangle_{\mathrm{VQE}}$	$\langle \hat{\square} \rangle_0$	$\langle \hat{\triangle} \rangle_{\mathrm{VQE}}$	$\langle \hat{\triangle} \rangle_0$	$\mathcal{F}$	$\sigma_{E,\mathrm{VQE}}$	$\sigma_{\square,\mathrm{VQE}}$	$\sigma_{\triangle,\mathrm{VQE}}$	$|\Delta E|/E_g$
0.5	0.0321	-0.0624	-0.936	-0.998	0.862	0.0289	0.00482	0.0285	0.148
0.63	-0.0468	-0.0987	-0.922	-0.995	0.850	0.0178	0.00560	0.0169	0.109
0.85	-0.166	-0.178	-0.936	-0.984	0.826	0.0153	0.00676	0.0137	0.103
1.12	-0.262	-0.299	-0.880	-0.954	0.811	0.0109	0.00665	0.00870	0.188
1.52	-0.363	-0.500	-0.830	-0.866	0.815	0.00976	0.00763	0.00608	0.296
1.98	-0.481	-0.700	-0.790	-0.714	0.776	0.0112	0.0100	0.00514	0.464
2.33	-0.572	-0.805	-0.773	-0.593	0.818	0.0143	0.0135	0.00493	0.573
2.65	-0.755	-0.869	-0.518	-0.495	0.768	0.0160	0.0150	0.00552	0.904
2.88	-0.868	-0.901	-0.286	-0.434	0.792	0.0197	0.0189	0.00579	1.282
3.3	-0.8559	-0.939	-0.230	-0.345	0.812	0.0206	0.0199	0.00510	2.569

Table C.3: Optimized energies, operator expectation values, fidelities, and errors for the QCD VQE demonstration ($s = 10$).

$\tilde{m}^a$	System[b]	Optimizer	Shots, Iters.	N_{CX}	κ	$E_{VQE}^{(i)}$	$E_{VQE}^{(f)}$	E_0	E_g
-1.0	ibm_lagos	COBYLA	750, 50	4	1	-3.574	-6.240	-6.259	2.878
-0.5	ibm_lagos	COBYLA	1522, 50	4	1	-2.179	-3.220	-3.395	1.771
-0.2	ibm_lagos	COBYLA	3953, 50	4	1	-1.322	-1.626	-1.791	0.965
-0.05	ibm_lagos	COBYLA	9680, 55	4	1	-0.831	-0.952	-1.122	0.639
0.01	ibm_lagos	COBYLA	10402, 56	4	1	-0.704	-0.763	-0.924	0.610
0.05	ibm_peekskill	COBYLA	9380, 57	4	1	-0.518	-0.623	-0.822	0.639
0.1	ibm_peekskill	COBYLA	7290, 50	4	1	-0.299	-0.570	-0.723	0.719
0.2	ibm_lagos	COBYLA	3953, 50	4	1	1.228	-0.384	-0.591	0.965
0.5	ibm_lagos	COBYLA	1131, 50	4	1	2.073	-0.335	-0.395	1.771
1.0	ibm_lagos	COBYLA	454, 50	4	1	3.446	-0.164	-0.259	2.878

$\tilde{m}$	$\langle\hat{N}\rangle_{VQE}$	$\langle\hat{N}\rangle_0$	$\langle\hat{N'}\rangle_{VQE}$	$\langle\hat{N'}\rangle_0$	$\mathcal{F}$	$\sigma_{E,VQE}$	$\sigma_{N,VQE}$	$\sigma_{N',VQE}$	$|\Delta E|/E_g$
-1.0	5.817	5.822	3.662	3.768	0.902	0.0368	0.00258	0.0242	0.00673
-0.5	5.638	5.580	3.454	3.469	0.834	0.0245	0.00692	0.0170	0.0992
-0.2	5.414	4.963	3.247	2.815	0.888	0.0120	0.00147	0.0103	0.171
-0.05	4.518	3.725	2.250	1.780	0.791	0.00741	0.00164	0.00817	0.266
0.01	3.067	2.849	1.060	1.170	0.884	0.00461	0.00150	0.00404	0.264
0.05	2.591	2.275	0.751	0.814	0.792	0.00598	0.000951	0.00194	0.311
0.1	1.441	1.702	0.808	0.497	0.808	0.00845	0.000730	0.00716	0.214
0.2	0.548	1.037	0.00223	0.197	0.771	0.0152	0.00320	0.00902	0.214
0.5	0.443	0.420	0.00905	0.0298	0.923	0.0273	0.00352	0.0140	0.0363
1.0	0.00737	0.178	-0.0801	0.00468	0.893	0.0499	0.00716	0.00648	0.0332

a At $x = 0.8$.
b Owing to logistical challenges with the IBM Quantum platform, data points were collected over multiple systems.

Table C.4: Optimized energies, fidelities, and errors for the LiH VQE demonstration.[a] All energies are expressed in Hartrees (Ha).

System	Optimizer	Shots, Iters.	N_{CX}	κ	$E_{VQE}^{(i)}$	$E_{VQE}^{(f)}$	E_0	E_1	E_2	$\mathcal{F}^{(i)}$	$\mathcal{F}^{(f)}$	$\sigma_{E,VQE}$	$|\Delta E|/E_0$
ibm_perth	COBYLA	10000, 99	7	1	−6.967	−7.794[b]	−7.881	−7.766	−7.748	0.0673[c]	0.782[c]	0.000534	0.0110

[a] At an interatomic distance of 1.6 Å.

[b] Since VQE convergence behavior is analyzed here, we do not apply self-mitigation to the final energy.

[c] See Fig. 2.6(e).

Table C.5: Optimized energies and errors for the 2×1 PC VQE demonstration ($s = 25$).[a]

ξ	System[b]	Optimizer[c]	Shots, Iters[c]	N_{CX}^{d}	κ^{d}	$E_{VQE}^{(i)}$	$E_{VQE}^{(f)}$	E_0	E_g	$\sigma_{E,VQE}$	$|\Delta E|/E_g$
0.01	ibm_perth	DIRECT+COBYLA	946, 50 + 50	27	43/27	−1.525	−7.750	−8.001	1.991	0.227	0.126
0.0316	ibm_lagos	DIRECT+COBYLA	960, 50 + 50	27	36/27	−1.516	−7.800	−8.009	1.977	0.105	0.106
0.1	ibm_perth	DIRECT+COBYLA	965, 50 + 50	27	40/27	−1.284	−7.465	−8.083	1.983	0.169	0.312
0.316	ibm_perth	DIRECT+COBYLA	781, 50 + 50	27	40/27	−1.216	−8.422	−8.632	2.316	0.199	0.091
1.0	ibm_lagos	DIRECT+COBYLA	1653, 50 + 50	27	36/27	−1.755	−10.880	−11.465	2.217	0.117	0.264
1.5	ibm_lagos	DIRECT+COBYLA	10172, 50 + 50	27	36/27	−1.850	−13.688	−13.822	1.156	0.0696	0.116
2.0	ibm_perth	DIRECT+COBYLA	50000, 50 + 50	27	40/27	−0.668	−17.828	−16.245	0.120	0.0743	0.402[e]
3.0	ibm_lagos	DIRECT+COBYLA	11722, 50 + 50	27	36/27	−2.269	−23.588	−23.083	1.918	0.227	0.263
4.0	ibm_lagos	DIRECT+COBYLA	4755, 50 + 50	27	36/27	−1.914	−30.585	−30.062	3.938	0.654	0.133
6.0	ibm_perth	DIRECT+COBYLA	2546, 50 + 50	27	40/27	−2.179	−45.345	−44.042	7.959	2.594	0.164
8.0	ibm_lagos	DIRECT+COBYLA	1981, 50 + 50	26	39/26	−3.106	−61.001	−58.031	11.969	6.316	0.248
10.0	ibm_lagos	DIRECT+COBYLA	1729, 50 + 50	27	36/27	−2.168	−80.911	−72.025	15.975	8.107	0.556

[a] Since 7 qubits are used in this demonstration, performing full state tomography is impractical and therefore the fidelities $\mathcal{F}$ were not calculated.

[b] See note in Table C.3.

[c] See discussion in Sec. B regarding optimization methods.

[d] On account of the larger number of qubits and their fixed connectivity on IBM Quantum systems, we employed the noise adaptive layout method to map qubits onto hardware [41]. The mapping is chosen automatically by Qiskit's transpiler (based on current calibration data) and leads to variations in the number of CX gates incurred by each circuit in a VQE run. Here, we reflect the full range of CX numbers during self-mitigation by calculating κ from the physical and mitigation circuits with the largest and smallest CX numbers respectively, over all measurement bases. In other words, we employ the maximum possible κ.

[e] Calculated using the gap of the second excited state $E_2 - E_1 = 3.939$, since the first excited state and the GS are nearly degenerate.

Table C.6: Initial and final (optimized) variational parameters for each VQE demonstration.

Model	Free Param.	$\vec{\theta}^{(i)}$	$\vec{\theta}^{(f)}$
$\mathbb{Z}_2$	$\lambda = 0.5$	$\{3.07, 5.56, 1.11, 4.17\}$	$\{3.274, 6.232, 2.023, 4.508\}$
	$\lambda = 0.63$	$\{3.07, 5.56, 1.11, 4.17\}$	$\{3.674, 6.119, 2.946, 4.138\}$
	$\lambda = 0.85$	$\{3.07, 5.56, 1.11, 4.17\}$	$\{3.119, 6.063, 1.11, 4.962\}$
	$\lambda = 1.12$	$\{3.07, 5.56, 1.11, 4.17\}$	$\{3.179, 5.948, 1.152, 4.974\}$
	$\lambda = 1.52$	$\{3.07, 5.56, 1.11, 4.17\}$	$\{3.210, 5.858, 2.710, 4.626\}$
	$\lambda = 1.98$	$\{3.07, 5.56, 1.11, 4.17\}$	$\{3.451, 5.741, 2.654, 4.668\}$
	$\lambda = 2.33$	$\{3.07, 5.56, 1.11, 4.17\}$	$\{3.210, 5.627, 2.766, 4.637\}$
	$\lambda = 2.65$	$\{3.07, 5.56, 1.11, 4.17\}$	$\{3.003, 5.302, 2.728, 4.654\}$
	$\lambda = 2.88$	$\{3.07, 5.56, 1.11, 4.17\}$	$\{3.302, 5.002, 2.790, 4.630\}$
	$\lambda = 3.3$	$\{3.07, 5.56, 1.11, 4.17\}$	$\{2.994, 4.984, 2.726, 4.595\}$
QCD	$\tilde{m} = -1$	$\{\pi, \pi/2, \pi, \pi/2\}$	$\{4.261, 1.551, 3.210, 1.728\}$
	$\tilde{m} = -0.5$	$\{\pi, \pi/2, \pi, \pi/2\}$	$\{4.142, 1.826, 3.142, 1.829\}$
	$\tilde{m} = -0.2$	$\{\pi, \pi/2, \pi, \pi/2\}$	$\{4.009, 1.682, 3.029, 1.887\}$
	$\tilde{m} = -0.05$	$\{\pi, \pi/2, \pi, \pi/2\}$	$\{3.640, 1.571, 3.142, 1.848\}$
	$\tilde{m} = 0.01$	$\{\pi, \pi/2, \pi, \pi/2\}$	$\{3.184, 1.502, 3.101, 1.795\}$
	$\tilde{m} = 0.05$	$\{\pi, \pi/2, \pi, \pi/2\}$	$\{3.280, 1.722, 3.293, 1.551\}$
	$\tilde{m} = 0.1$	$\{\pi, \pi/2, \pi, \pi/2\}$	$\{3.679, 2.131, 3.142, 1.571\}$
	$\tilde{m} = 0.2$	$\{\pi/2\} \times 4$	$\{2.105, 1.581, 2.897, 2.185\}$
	$\tilde{m} = 0.5$	$\{\pi/2\} \times 4$	$\{2.183, 1.671, 3.217, 1.757\}$
	$\tilde{m} = 1$	$\{\pi/2\} \times 4$	$\{0.785, 4.328, 2.955, 0.955\}$
PC	$\xi = 0.01$	$\{\pi\} \times 11$	$\{5.278, 3.113, 3.118, 3.099, 3.103, 3.100,$ $3.117, 1.047, 3.113, 3.115, 3.108\}$
	$\xi = 0.0316$	$\{\pi\} \times 11$	$\{1.051, 3.137, 3.127, 3.159, 3.043, 3.221,$ $3.144, 6.272, 3.146, 3.147, 3.142\}$
	$\xi = 0.1$	$\{\pi\} \times 11$	$\{1.047, 3.142, 3.159, 2.990, 3.142, 3.142,$ $3.142, 6.021, 3.142, 3.142, 3.142\}$
	$\xi = 0.316$	$\{\pi\} \times 11$	$\{1.852, 3.142, 3.142, 3.142, 3.142, 3.142,$ $3.142, 5.236, 4.745, 3.142, 3.142\}$
	$\xi = 1.0$	$\{\pi\} \times 11$	$\{1.032, 3.142, 3.094, 3.141, 3.130, 3.215,$ $3.335, 6.283, 3.122, 3.120, 3.124\}$
	$\xi = 1.5$	$\{\pi\} \times 11$	$\{1.046, 3.388, 2.844, 2.994, 3.146, 3.036,$ $3.003, 6.282, 2.980, 3.142, 3.209\}$
	$\xi = 2.0$	$\{\pi\} \times 11$	$\{1.651, 3.142, 1.964, 3.142, 3.227, 3.219,$ $3.124, 6.160, 3.142, 3.142, 3.142\}$
	$\xi = 3.0$	$\{\pi\} \times 11$	$\{1.679, 3.142, 1.711, 3.142, 3.142, 2.978,$ $2.685, 5.775, 1.384, 6.283, 6.283\}$
	$\xi = 4.0$	$\{\pi\} \times 11$	$\{1.832, 3.142, 1.832, 3.142, 3.166, 3.166,$ $3.171, 6.018, 4.696, 4.718, 3.142\}$
	$\xi = 6.0$	$\{\pi\} \times 11$	$\{1.702, 3.976, 1.369, 3.017, 2.552, 2.447,$ $2.539, 6.254, 0.352, 5.622, 3.498\}$
	$\xi = 8.0$	$\{\pi\} \times 11$	$\{1.703, 2.287, 2.094, 3.237, 3.136, 3.211,$ $3.178, 6.283, 2.942, 5.658, 3.717\}$
	$\xi = 10.0$	$\{\pi\} \times 11$	$\{1.499, 3.778, 1.360, 2.496, 2.355, 2.939,$ $2.273, 6.078, 4.990, 5.584, 3.703\}$
LiH	$ID = 1.6$ Å	$\{0.1\} \times 9$	$\{0.483, 1.713, -1.475, 4.600, 2.326,$ $1.048, -0.126, -0.126, -0.126\}$

85

Appendix D

Hybrid VQE Demonstrations – Error Analysis

In this appendix, we elaborate on the method used to compute observable errors with statistical correlations and after self-mitigation is applied (App. B). For a given observable $\hat{O} \in \hat{\mathcal{H}}$ measured with N shots, the variance is given by:

$$\text{Var}(\hat{O}) = \frac{1}{N} \sum_i N^{(i)} (O^{(i)} - \langle \hat{O} \rangle)^2, \tag{D.1}$$

where $O^{(i)}$ is the ith unbiased estimator (i.e. a computational basis measurement outcome) of $\hat{O}$ with frequency $N^{(i)}$, $\sum_i N^{(i)} = N$, and $\langle \hat{O} \rangle = \sum_i N^{(i)} O^{(i)} / N$ is the average over all estimators. The uncorrelated error is then calculated as the standard deviation $\sqrt{\text{Var}(\hat{O})}$.

To lower resource requirements, we exploit the property that any set of commuting observables in $\hat{\mathcal{H}}$ can be measured *simultaneously* (i.e. with a single circuit). Specifically, if $\hat{O}$ is an n-qubit Pauli string $\hat{P}_1 \hat{P}_2 ... \hat{P}_n \in \{\hat{I}, \hat{X}, \hat{Y}, \hat{Z}\}^{\otimes n}$, then any n-qubit Pauli $\hat{Q}_1 \hat{Q}_2 ... \hat{Q}_n$ where $\hat{Q}_i \in \{\hat{P}_i, I\}$ will commute with $\hat{O}$. These observables are grouped into a set G from which we designate a single circuit capable of measuring each of them during readout. We note that this is guaranteed since each $\hat{O} \in G$ must act either $\hat{I}$ or the same $\hat{X}, \hat{Y}$ or $\hat{Z}$ on a given qubit. Thus, one may set each qubit basis to the corresponding non-identity Pauli (if present in at least one $\hat{O}$, else it remains unmeasured) and this constitutes a valid N-qubit measurement basis for the circuit.

The grouping requires us to account for correlations that exist between observables in G. In the context of error analysis, these manifest as covariances since each observable

is influenced by the statistics of the same counts distribution. Partitioning $\hat{\mathcal{H}}$ into K commuting groups G_k, we obtain:

$$\hat{\mathcal{H}} = \sum_{j_1=1}^{|G_1|} c_{j_1} \hat{O}_{j_1} + \sum_{j_2=1}^{|G_2|} c_{j_2} \hat{O}_{j_2} + ... + \sum_{j_K=1}^{|G_K|} c_{j_K} \hat{O}_{j_K}, \tag{D.2}$$

with coefficients $c_1, ..., c_{|G_k|}$ and measurement basis $\hat{O}_k$ corresponding to the group G_k. Note that a particular $\hat{O}_{j_k}$ may belong to multiple G; its expectation value is then computed from the combined counts of multiple circuits, which contributes to increased estimation accuracy at no added cost. Let $\hat{O}'$ denote such an observable (with coefficient c) belonging to $k \leq K$ groups. Then its contribution to each G_k is $cR_k\hat{O}'$, where $R_k = N_k/(\sum_{k=1}^{K} N_k)$. To compute the covariance, we consider all bitwise products between $\hat{O}_a$ and $\hat{O}_b$ (cross-terms) within each G. We then have the set $G^\times$, which we define as the set G augmented with the cross-terms, i.e.:

$$G^\times = G \cup \{\hat{O}_a \hat{O}_b \,\forall\, a, b \leq |G| : a \neq b\}. \tag{D.3}$$

Therefore, the variance of $\hat{\mathcal{H}}$ with correlations included (denoted with a tilde) is:

$$\widetilde{\mathrm{Var}}(\hat{\mathcal{H}}) = \sum_{k=1}^{K} \frac{\widetilde{\mathrm{Var}}(\hat{O}_k)}{N_k}, \tag{D.4}$$

where,

$$\widetilde{\mathrm{Var}}(\hat{O}_k) = c_k^2 \,\mathrm{Var}(\hat{O}_k) + \sum_{\substack{\ell=1 \\ \hat{O}_k \neq \hat{O}_\ell}}^{|G_k^\times|} \left[c_\ell^2 R_\ell^2 \mathrm{Var}(\hat{O}_\ell) + 2c_k c_\ell R_k \big(\underbrace{\langle \hat{O}_k \hat{O}_\ell \rangle - \langle \hat{O}_k \rangle \langle \hat{O}_\ell \rangle}_{\mathrm{Cov}(\hat{O}_k, \hat{O}_\ell)} \big) \right], \tag{D.5}$$

and $\mathrm{Cov}(\hat{O}_k, \hat{O}_\ell)$ denotes the co-variance between $\hat{O}_k$ and $\hat{O}_\ell$.

Now, we may proceed to calculate the errors after self-mitigation. Since $\langle \hat{O} \rangle_{\mathrm{phys,\,true}}$ is a function of $\langle \hat{O} \rangle_{\mathrm{phys,\,meas}}$ and $\langle \hat{O} \rangle_{\mathrm{mitig,\,meas}}$, we must propagate their respective errors.

Applying the variance formula to Eq. (B.9), with $\langle \hat{O}_{k,\text{mitig, true}} \rangle$ fixed at 1, we obtain:

$$\widetilde{\text{Var}}(\hat{O}_{k,\text{phys, true}}) \tag{D.6}$$

$$= \left(\frac{\partial \langle \hat{O}_k \rangle_{\text{phys, true}}}{\partial \langle \hat{O}_k \rangle_{\text{phys, meas}}} \right)^2 \widetilde{\text{Var}}(\hat{O}_{k,\text{phys, meas}}) + \left(\frac{\partial \langle \hat{O}_k \rangle_{\text{phys, true}}}{\partial \langle \hat{O}_k \rangle_{\text{mitig, meas}}} \right)^2 \widetilde{\text{Var}}(\hat{O}_{k,\text{mitig, meas}}) \tag{D.7}$$

$$= \left(\frac{1}{\langle \hat{O}_k \rangle_{\text{mitig, meas}}} \right)^{2\kappa} \widetilde{\text{Var}}(\hat{O}_{k,\text{phys, meas}}) + \kappa^2 \left(\frac{\langle \hat{O}_k \rangle_{\text{phys, meas}}}{(\langle \hat{O}_k \rangle_{\text{mitig, meas}})^{\kappa+1}} \right)^2 \widetilde{\text{Var}}(\hat{O}_{k,\text{mitig, meas}}), \tag{D.8}$$

with $\sqrt{\widetilde{\text{Var}}(\hat{O}_{k,\text{ phys, true}})}$ being the errors reported in the simulation data tables (App. C). Here, we remark that the propagation formula linearly approximates $\langle \hat{O} \rangle_{\text{phys, true}}$, and therefore assumes the errors from the physical and mitigation runs to be uncorrelated and relatively small. This is a fair assumption since their respective observables are calculated from separate circuits and retain the same counts distribution after correcting the expectation values. However, if self-mitigation is instead used to correct each frequency in the counts distribution, then a reshaping can occur and potentially amplify the errors. These situations would require alternate techniques (e.g. Monte Carlo [131]) that analyze how the observables' error distributions transform under propagation, owing to their nonlinear relations with the function.

Appendix E

Perturbative Analysis of the $M \times N$ Planar Code

In this appendix, we analyze the $M \times N$ planar code on a more physical level by studying how the local perturbations act on its qubits. To achieve this, we employ time-independent perturbation theory to approximate the GS and associated energy as a function of the perturbation strength ξ. Since the free part of the Hamiltonian [Eq. (2.3)] is $\hat{\mathcal{H}}_0 = \hat{\mathcal{H}}_\square + \hat{\mathcal{H}}_+$, its corresponding GS energy is given by the contributions from all its corresponding plaquette and star operators:

$$E_0^{(0)} = - \underbrace{MN}_{\# \text{ plaquettes}} - \underbrace{(M-1)(N-1)}_{\# \text{ 4-body star}} - \underbrace{2(M+N-2)}_{\# \text{ 3-body star}} - \underbrace{4}_{\# \text{ 2-body star}} = \sum_{i=1}^{n_q+1} (-1), \quad \text{(E.1)}$$

where each operator contributes a minimum eigenvalue of –1 and the superscript (0) indicates the unperturbed (uncorrected) case. On the other hand, the associated GS arises from the contribution of all self- and multi-plaquette interactions:

$$|\Psi_0^{(0)}\rangle = \frac{1}{\sqrt{2^{MN}}} \left[1 + \sum_{p_1=1}^{MN} \prod_{p_1} \square_{p_1} + \sum_{\substack{p_1,p_2=1 \\ p_1<p_2}}^{MN} \prod_{p_1,p_2} \square_{p_1}\square_{p_2} + ... + \square_{p_1}\square_{p_2}...\square_{p_{MN}} \right] |0\rangle^{\otimes n_q}. \quad \text{(E.2)}$$

Here, n_q is the number of qubits and each $\square_p$ is a 4-body plaquette operator $\hat{X}_A\hat{X}_B\hat{X}_C\hat{X}_D$, where A, B, C, D labels the four qubits forming the edges of plaquette p. Since the $M \times N$ lattice possesses a graph state representation, we may cast Eq. (E.2) in a simpler form involving its stabilizers. Denoting Q as the set of all qubits (i.e. $\{1, 2,...,n_q\}$), E_{grph} as

89

the set of all CZ edges (acting on qubits a and b), and $LC \subset Q$ as the subset of qubits requiring local Clifford (H) operations, we have:

$$|\Psi_0^{(0)}\rangle = \left(\bigotimes_q H_q \right) \left(\prod_{E_{grph}} CZ_{a,b} \right) |+\rangle^{\otimes n_q}; \quad q \in LC \tag{E.3}$$

From this GS expression, one may systematically generate all excited states by applying combinations of $\hat{X}$ and $\hat{Z}$ operators, which serve to flip the constituent eigenvalues of the GS from -1 to $+1$. These states form a complete, mutually orthonormal basis, and are given as:

$$|\Psi_n^{(0)}\rangle = \left(\bigotimes_{j=1}^{n_q} (\hat{X}_j^{\vec{x}_j})^k (\hat{Z}_j^{\vec{x}_j})^{1-k} \right) |\Psi_0^{(0)}\rangle \quad (n = 1, 2...), \tag{E.4}$$

where $\vec{x} \in \{0, 1\}^{n_q} : \vec{x} \neq 0^{n_q}$, and $k = \begin{cases} 0, & j \notin LC \\ 1, & j \in LC \end{cases}$. Based on Eq. (E.4), we can derive the action of the free Hamiltonian on an excited state, and obtain an exact form for the excited state energies (each energy corresponding to a specific $\vec{x}$):

$$\hat{\mathcal{H}}_0 |\Psi_n^{(0)}\rangle = \sum_{i=1}^{n_q+1} \hat{\mathcal{H}}_{0,i} \left(\bigotimes_{j=1}^{n_q} (\hat{X}_j^{\vec{x}_j})^k (\hat{Z}_j^{\vec{x}_j})^{1-k} \right) |\Psi_0^{(0)}\rangle \tag{E.5}$$

$$= \sum_{i=1}^{n_q+1} \bigotimes_{j=1}^{n_q} \hat{A}_j^i |\Psi_0^{(0)}\rangle; \quad \hat{A}_j^i \in \{\hat{I}_j, \hat{X}_j, \hat{Z}_j, \hat{Z}_j\hat{X}_j, \hat{X}_j\hat{Z}_j\} \tag{E.6}$$

$$\implies E_n^{(0)} = \sum_{i=1}^{n_q+1} \prod_{j=1}^{n_q} a_j^i, \quad \text{where :} \tag{E.7}$$

$$a_i^j = \begin{cases} -1, & \hat{A}_j^i \in \{(\hat{X}_j\hat{Z}_j; \ j \notin LC), (\hat{Z}_j\hat{X}_j; \ j \in LC)\} \\ +1, & \hat{A}_j^i \in \{\hat{I}_j, (\hat{Z}_j; \ j \notin LC), (\hat{X}_j; \ j \in LC)\} \end{cases}, \tag{E.8}$$

and $\hat{\mathcal{H}}_{0,i}$ is the ith term in $\hat{\mathcal{H}}_0$. We can now proceed with the perturbative analysis, where we determine the first and second order corrections in the GS energy and GS respectively.

The first-order energy correction vanishes since:

$$E_0^{(1)} = \langle \Psi_0^{(0)} | \hat{\mathcal{H}}_\triangle | \Psi_0^{(0)} \rangle,$$

$$= \xi \langle + |^{\otimes n_q} \left(\bigotimes_{j=1}^{n_q} \hat{Z}_j \right) | + \rangle^{\otimes n_q} = \xi \langle + | - \rangle^{\otimes n_q} = 0, \tag{E.9}$$

and is indicative of a non-polarized (singlet) state in which all the up and down spins are paired. For the second-order correction, we have:

$$
\begin{aligned}
E_0^{(2)} &= \sum_n \frac{|\langle \Psi_n^{(0)} | \hat{\mathcal{H}}_\triangle | \Psi_0^{(0)} \rangle|^2}{E_0^{(0)} - E_n^{(0)}} \\
&= \xi^2 \sum_{\vec{x}} \frac{|\sum_{j'} \langle + |^{\otimes n_q} (\prod_{E_{grph}} CZ_{a,b}) \hat{X}_{j'}^k (\prod_{E_{grph}} CZ_{a,b}) (\hat{Z}_{j'}^{1-k} \bigotimes_j \hat{Z}_j^{\vec{x}_j}) | + \rangle^{\otimes n_q}|^2}{\sum_i (-1 + \prod_j a_j^i)} \\
&= \xi^2 \sum_{\vec{x}} \frac{|\sum_{j'} \langle + |^{\otimes n_q} (\bigotimes_{j'' \in V_{j'' \to j'}} \hat{Z}_{j''}^k)(\hat{X}_{j'}^k \hat{Z}_{j'}^{\vec{x}_{j'}+1-k})(\bigotimes_{j \neq j'} \hat{Z}_j^{\vec{x}_j}) | + \rangle^{\otimes n_q}|^2}{\sum_i (-1 + \prod_j a_j^i)} \\
&= \xi^2 \sum_{\vec{x}} \frac{|\sum_{j'} \langle + |^{\otimes n_q} \hat{X}_{j'}^k (\hat{Z}_{j'}^{\vec{x}_{j'}+1-k} \bigotimes_{j \neq j'} \hat{Z}_j^{\vec{x}_j + ck}) | + \rangle^{\otimes n_q}|^2}{\sum_i (-1 + \prod_j a_j^i)},
\end{aligned}
\tag{E.10}
$$

where $V \subset Q$ is the set of all qubits j'' connected via CZ-edges to the qubit j', and $c = \begin{cases} 0, & j \notin V \\ 1, & j \in V \end{cases}$. The non-zero contributions in $E_0^{(2)}$ occur when each qubit is acted on by an even number of perturbative $\hat{Z}$ operators, which prevents orthogonality between $\hat{\mathcal{H}}_\triangle | \Psi_0^{(0)} \rangle$ and $| \Psi_n^{(0)} \rangle$. Since the perturbation is odd under parity transformation, its action above imposes interactions between antisymmetric states which is a characteristic of spin-1/2 systems. From Eq. (E.10), we may now determine the first order correction to the GS:

$$|\Psi_0^{(1)}\rangle = \sum_n \frac{\langle \Psi_n^{(0)} | \hat{\mathcal{H}}_\triangle | \Psi_0^{(0)} \rangle}{E_0^{(0)} - E_n^{(0)}} | \Psi_n^{(0)} \rangle \tag{E.11}$$

$$= \left[\xi \sum_{\vec{x}} \frac{\sum_{j'} \langle + |^{\otimes n_q} \hat{X}_{j'}^k (\hat{Z}_{j'}^{\vec{x}_{j'}+1-k} \bigotimes_{j \neq j'} \hat{Z}_j^{\vec{x}_j + ck}) | + \rangle^{\otimes n_q}}{\sum_i (-1 + \prod_j a_j^i)} \left(\bigotimes_j (\hat{X}_j^{\vec{x}_j})^k (\hat{Z}_j^{\vec{x}_j})^{1-k} \right) \right] | \Psi_0^{(0)} \rangle. \tag{E.12}$$

Applying Eqs. (E.10) and (E.12) to the 1×1 PC (single plaquette), the corrected GS energy and GS to their respective orders are:

$$E_0(\xi) \approx -5 - 8\xi^2 \tag{E.13}$$

$$|\Psi_0(\xi)\rangle \approx \left(\frac{1}{\sqrt{2}} - \xi\sqrt{2}\right)|0\rangle^{\otimes 4} + \left(\frac{1}{\sqrt{2}} + \xi\sqrt{2}\right)|1\rangle^{\otimes 4}. \tag{E.14}$$

For the 1×2 (2×1) PC, containing two horizontal (vertical) plaquettes:

$$E_0(\xi) \approx -8 - \frac{37}{4}\xi^2 \tag{E.15}$$

$$|\Psi_0(\xi)\rangle \approx \left(\frac{1}{2} - \frac{13}{8}\xi\right)|0\rangle^{\otimes 7} + \left(\frac{1}{2} + \frac{1}{8}\xi\right)(|1011010\rangle + |0101101\rangle) + \left(\frac{1}{2} + \frac{11}{8}\xi\right)|1110111\rangle. \tag{E.16}$$

From these equations, we observe that increasing the perturbation ξ shifts the GS energy downwards, while the GS shifts towards states which have qubits that are all or predominantly in the polarized $|1\rangle$ state. This is expected since the dominance of the terms in $\hat{\mathcal{H}}_\triangle$ (each containing a single $\hat{Z}$ operator) acting on $|1\rangle$ contributes eigenvalues $\propto -\xi$ for large ξ. With these expressions valid for small-to-intermediate ξ, the $\hat{Z}$-operator terms guarantee a crossover to $|1\rangle^{\otimes n_q}$ as $\xi \to \infty$. It is also worth mentioning the equal positive and negative contributions by the perturbation on $|\Psi_0\rangle$, which is due to its uniform nature. This preserves the underlying symmetry of the problem, and thus enables redundant parameters to be eliminated in the VQE implementation. We remark that if different perturbation strengths were applied to each qubit, then a symmetry breaking is expected and all free parameters must be optimized separately.

Appendix F

The `fast_wolf` Algorithm

In this appendix, we introduce the `fast_wolf` algorithm which performs efficient MB-QC with large numbers of *adaptive* qubits. Unlike non-adaptive qubits (which may be classically processed), adaptive qubits form the essential part of a resource state. In the standard MB-QC approach, all of them must be initialized and entangled before any measurements can occur, which leads to a clear exponential scaling and performance bottleneck. However, this is not a strict requirement as practical graph designs often do not possess all-to-all-connectivity. In these scenarios, information propagation is confined along smaller, local portions of the graph. The `fast_wolf` method takes advantage of this by processing the graph in a *piecewise* fashion. It processes these local parts one at a time and measures only qubit(s) in its immediate vicinity.

The method is also capable of handling arbitrary *non-Clifford* inputs to the MB pattern, which cannot be processed efficiently via the stabilizer formalism. These correspond to states which are not eigenstates of the Pauli operators $\hat{I}, \hat{X}, \hat{Y}, \hat{Z}$, such as $T|+\rangle$. To achieve this, we initially omit the arbitrary input qubits when classically simulating the non-adaptive portion (App. A). After obtaining the LC-reduced graph, we attach "connector" qubits (initialized in $|+\rangle$ or $|-\rangle$) to prescribed body qubits and then entangle it with an arbitrary input. The role of these connectors is to allow teleportation of the input state into the graph upon their measurement.

The `fast_wolf` simulation begins from the input qubits and entangles them with their nearest body-qubit neighbours. For non-Clifford inputs, the connectors are also entangled to the inputs, with both sets of qubits subsequently measured in the X-basis. From there, the body qubits are entangled with their respective neighbours and measured in the prescribed basis. This process continues until only the output qubits remain, at which point the byproduct operators $\hat{U}_\Sigma$ are applied to $|\psi_{\text{out}}\rangle$ based on the outcomes from the adap-

tive basis measurements and earlier classical simulation. Throughout, we act local Clifford (H, S) operations on a particular qubit (as determined from the classical simulation) immediately after it is entangled with its neighbours.

We make two further remarks. First, since only a *single* adaptive qubit is measured at a time, any angle flip in the basis depends solely on the outcome of the measurement, although an overall temporal ordering must still be followed to realize a deterministic pattern. Second, it is often convenient to determine all the steps involved in the `fast_wolf` procedure beforehand, storing them as a prescription list which can be reused if the same graph is executed multiple times (i.e. during a VQE run).

We present the pseudocode for the `fast_wolf` algorithm on the next page. It is divided into two main functions — `simulate_cliff` performs the efficient classical simulation, and `simulate_fast_wolf` is the piecewise execution of the graph itself:

Algorithm 1 `fast_wolf`

Inputs:
P: A valid measurement pattern with input (I), output (O), nonadaptive $(B^{(NA)})$, and adaptive body $(B^{(A)})$ qubits
$|\psi_{\text{in}}\rangle$: an arbitrary state over I
Outputs:
$|\psi_{\text{out}}\rangle$: the final state over O

1: **function** SIMULATE_FAST_WOLF$(P, |\psi_{\text{in}}\rangle)$
2:      $G, LC \leftarrow$ SIMULATE_CLIFF$(P, |\psi_{\text{in}}\rangle)$
3:      $|\Psi_G\rangle \leftarrow 1$
4:      $|\Psi_G\rangle \leftarrow |\Psi_G\rangle \otimes |\Psi_i\rangle \; \forall i \in I$ $\triangleright$ Construct $|\Psi_G\rangle$ from $I \cup IC \cup \mathcal{N}(IC) \in G$
5:      **if** $IC \neq \varnothing$ **then**
6:          $|\Psi_G\rangle \leftarrow CZ_{i,j}|\Psi_G\rangle \; \forall (i,j) \in (IC, \mathcal{N}(\tilde{I})) : i \neq j$
7:          $|\Psi_G\rangle \leftarrow LC_k|\Psi_G\rangle \; \forall k \in IC$
8:          $|\Psi_G\rangle \leftarrow CZ_{m,n}|\Psi_G\rangle \; \forall (m,n) \in (I, IC) : m \neq n$
9:          Measure $n \; \forall n \in I \cup IC$ (in X-basis)
10:     **end if**
11:     $Q \leftarrow \{1, 2, ..., (|IC| + |B^{(A)}| + |O|)\}$
12:     **while** $|Q| > |O|$ **do** $\triangleright$ Additional body qubits remain
13:          **for** $q \in Q$ **do**
14:              $|\Psi_G\rangle \leftarrow |\Psi_G\rangle \otimes |\Psi_r\rangle$
15:              $|\Psi_G\rangle \leftarrow CZ_{q,r}|\Psi_G\rangle \; \forall r \in \mathcal{N}(q)$
16:              **if** $q \in B^{(A)} \cup O$ **then** $\triangleright$ Measuring q must be valid w.r.t temporal ordering
17:                  $|\Psi_G\rangle \leftarrow LC_q|\Psi_G\rangle$
18:                  Measure q in prescribed basis
19:                  $Q \leftarrow Q - q$ $\triangleright$ Remove measured qubit q
20:              **end if**
21:          **end for**
22:     **end while** $\triangleright$ Exit loop when only output qubits remain
23:     $|\psi_{\text{out}}\rangle \leftarrow |\Psi_G\rangle$
24:     $|\psi_{\text{out}}\rangle \leftarrow U_\Sigma^{(i)}|\psi_{\text{out}}\rangle \; \forall i \in O$ $\triangleright$ Apply byproduct operators $U_\Sigma^{(i)}$ to qubit i
25:     **return** $|\psi_{\text{out}}\rangle$
26: **end function**
27:
28: **function** SIMULATE_CLIFF$(P, |\psi_{\text{in}}\rangle)$
29:     Construct tableau $\mathcal{T}$ from $I \cup B^{(NA)} \in P$
30:     Measure $I \cup B^{(NA)}$ in $\mathcal{T}$
31:     Determine resulting graph G over $B^{(A)} \cup O$
32:     **if** $\hat{\mathcal{P}}|\psi_{\text{in}}\rangle \neq \pm|\psi_{\text{in}}\rangle \; \forall \hat{\mathcal{P}} \in \{\hat{I}, \hat{X}, \hat{Y}, \hat{Z}\}^{\otimes |I|}$ **then**:
33:          Add connectors $IC_q \; \forall q \in I$ to G $\triangleright$ Input is not a Pauli eigenstate (i.e. non-Clifford)
34:     **end if**
35:     Determine $LC_q \in \{H, S, CZ\}; \; \forall q \in IC \cup B^{(A)} \cup O$
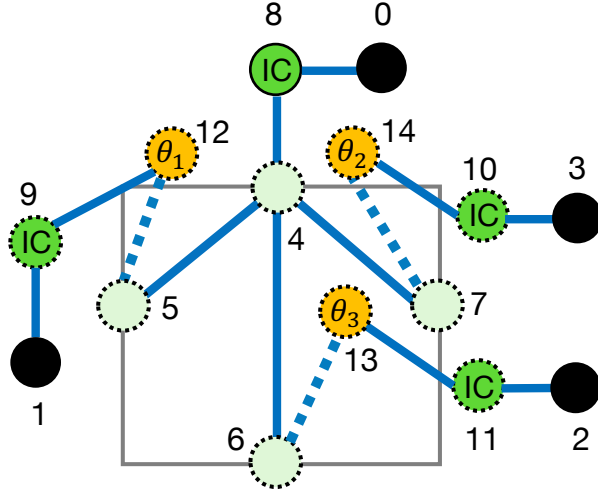36:   **return** G, LC
37: **end function**

FIG. F.1: **LC-reduced graph for the** $RY(\theta)$–CZ **modified** 1×1 **PC, with attached connector qubits (labelled by IC) to permit non-Clifford input states**. Connector, body, and output qubits are shown in green, orange, and light green respectively. Dotted circles around qubits indicate LC operations applied to them. Black circles indicate the arbitrary inputs, which undergo teleportation via the connectors into the main portion of the graph.

Example. To illustrate the `fast_wolf` algorithm, we apply it to the $RY(\theta)$–CZ modified 1×1 PC [Fig. 2.3(a)] with an arbitrary non-Clifford input $|\psi_{\text{in}}\rangle$ (defined over qubits 0 – 3). The corresponding LC-reduced graph is depicted in Fig. F.1. Performing the algorithm on the graph yields the following prescription steps:

```
1.  (init_input, {0, 1, 2, 3}), (init_connector, {8, 9, 10, 11}),
2.  (init_body, 13), (cz, 13, 11), (s, 11), (cz, 2, 11), (measure, 11),
3.  (init_body, 14), (cz, 14, 10), (s, 10), (cz, 3, 10), (measure, 10),
4.  (init_body, 12), (cz, 12, 9), (s, 9), (cz, 1, 9), (measure, 9),
5.  (init_out, 4), (cz, 4, 8), (cz, 0, 8), (measure, {8, 3, 2, 1, 0})
6.  (init_out, 5), (cz, 5, 12), (s, 12), (measure, 12),
7.  (init_out, 7), (cz, 7, 14), (s, 14), (measure, 14),
8.  (cz, 7, 4), (init_out, 6), (cz, 6, 13), (s, 13), (measure, 13),
9.  (cz, 7, 5), (cz, 7, 6), (s, 6), (s, 7), (s, 5), (s, 4), (permute, 3, 2,
0, 1),
```

where the input connectors (8–11), body qubits (12–14), and outputs (4–7) are initialized to either $|+\rangle$ or $|-\rangle$ depending on the outcomes from the efficient classical simulation,

and the qubits are labelled as in Fig. F.1. An extra permutation is prescribed at the end to ensure the ordering of qubits in $|\psi_{\text{out}}\rangle$ is consistent with that of $|\psi_{\text{in}}\rangle$. With the exception of the first initialization and final LC/permutation step, we consider a step to conclude whenever a measurement (or sequence of measurements) is performed. We obtain the state at the end of each step, and employ it as the input of the next.

Fig. F.2 shows a schematic of how the above steps are performed, in both their graph and circuit representations. Here, we have assumed all adaptive and non-adaptive outcomes are '0' so that no byproduct or adaptive corrections are required. From the figure, the advantages of the method are abundantly clear. We observe that at most 9 qubits are present at any time (vs. 15 in the standard approach), and no more than 3 qubits are entangled at once. One may also verify that for any choice of $|\psi_{\text{in}}\rangle$ and angles $\theta_1, \theta_2, \theta_3$, $\hat{U}_\Sigma |\psi_{\text{out}}\rangle$ is indeed equivalent to the action of the modified GB circuit [described by Eq. (H.1)] on $|\psi_{\text{in}}\rangle$.

Although the `fast_wolf` method is a practical way to simulate arbitrary inputs and complex patterns, it is important to note the challenges of the technique in experimental settings. For instance, the resource state must be entangled multiple times with qubits in arbitrary states. Given the challenges with generating and maintaining entanglement interactions in the lab, this can lead to greater imperfections in resource state preparation while affecting the computational fidelity. Moreover, it also remains to be seen whether the `fast_wolf` method provides a significant enough exponential size reduction. This can be done through a detailed bounds or complexity analysis. If the number of qubits remains large for a given pattern, one can resort to a concatenation of smaller patterns, with `simulate_fast_wolf` applied separately to each. However, this requires prior knowledge of the unitary decomposition, which may be unknown if one designs patterns in an arbitrary fashion (i.e. not from concatenating basic operations; see App. A).
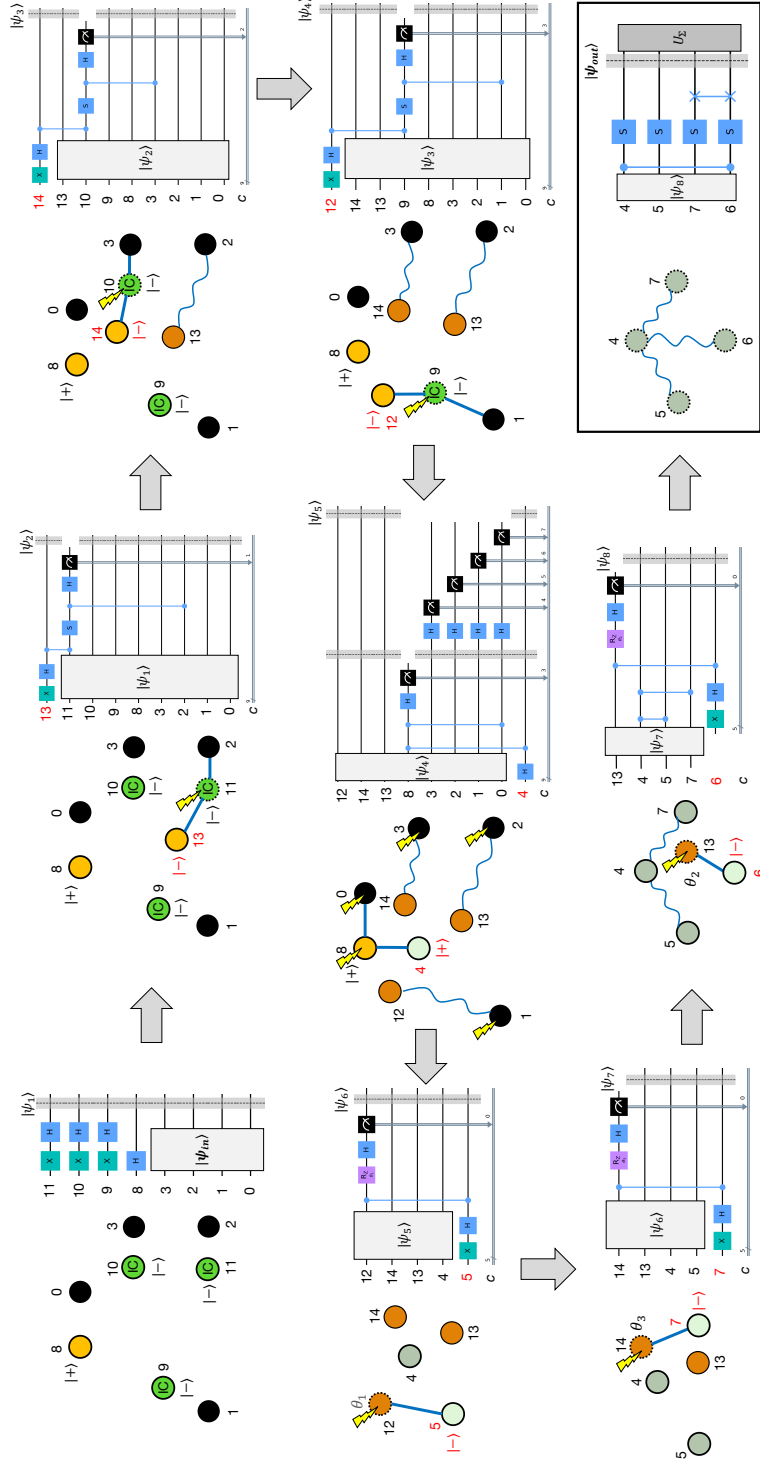
97

FIG. F.2: **Schematic diagram of the** `fast_wolf` **algorithm applied to the** $RY(\theta)$–CZ **modified** 1×1 **PC, with arbitrary non-Clifford inputs.** Arrows indicate the process flow, straight blue edges are CZ-operations applied, lightning symbols indicate measurements performed on the graph qubits, and wavy lines indicate the correlations that exist after measurement (which can vary in strength). Qubits initialized at a given step are highlighted with red labels and initial states, while color shade changes signify changes to their state after measurement.

Appendix G

Verifying Equivalence of the GB and MB Paradigms

In this appendix, we demonstrate that our ansatz-modification technique applied to the perturbed PC acts identically in the GB and MB paradigms. In other words, we verify that their corresponding unitary transformations are *equivalent*, that is: $\hat{U}_{GB} \equiv \hat{U}_{MB}$. For the purpose of analysis, we represent these paradigms as circuits and graphs respectively.

If $\hat{U}_{MB}$ is known beforehand, then verifying equivalence is trivial. However, it is often the case that the unitary is unknown; this can occur if the graph was designed deterministically, but in an arbitrary fashion (i.e. without resorting to concatenation of well-known patterns). In these situations, it is in fact possible to recover the operation by analyzing the relation between the pattern and its action on a unitary map. We use this technique, known as *phase map decomposition* (PMD) [9], as a general means of performing the verification. As a further check, we also present an alternative approach based on elements of our MB framework discussed in earlier appendixes.

For simplicity, we consider the $M = N = 1$ case (Figs. 2.3 and 2.4). Excluding the rotation layers, the modified circuit [Fig. 2.3(a)] realizes the overall unitary:

$$\hat{U}_{GB}(\theta_1, \theta_2, \theta_3) = H_2 H_3 H_4 \hat{U}_{1,4}^{am,GB}(\theta_3) \hat{U}_{1,3}^{am,GB}(\theta_2) \hat{U}_{1,2}^{am,GB}(\theta_1) H_1 H_2 H_3 H_4, \qquad \text{(G.1)}$$
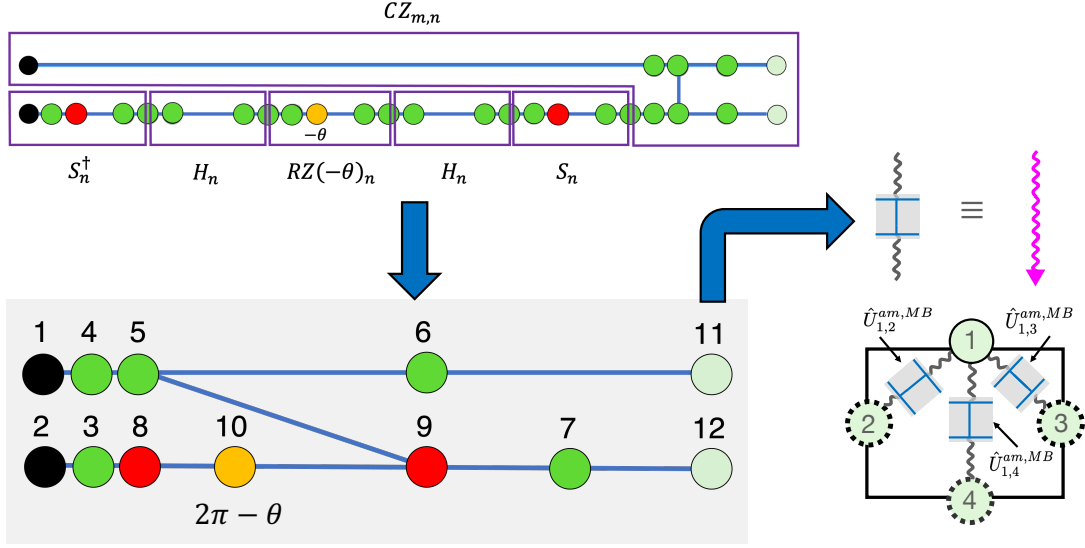
FIG. G.1: **MB pattern for** $RY(\theta)$–CZ, **obtained by concatenating smaller patterns corresponding to each operation in Eq.** (G.3). It is applied as a modification to each graph edge of the 1×1 perturbed PC. To reduce the number of initial qubits, the pattern is simplified via the techniques of Ref. [10]. Here, the pattern is regarded as a black-box unitary that is "inserted" into a graph edge to modify its degree of entanglement. Black and light green circles denote input and output qubits respectively, blue lines denote CZ-edges, green and red circles denote nonadaptive qubits measured in the X- and Y-basis respectively, and orange circles denote adaptive qubits measured in the $R(\theta)$-basis. The numbers label the qubits and their measurement order (excluding outputs).

where,

$$\hat{U}_{m,n}^{am,GB}(\theta) = CZ_{m,n}RY(\theta)_n = CZ_{m,n}(S^\dagger HRZ(-\theta)HS)_n \tag{G.2}$$

$$= \begin{pmatrix} \cos(\theta/2) & -\sin(\theta/2) & 0 & 0 \\ \sin(\theta/2) & \cos(\theta/2) & 0 & 0 \\ 0 & 0 & \cos(\theta/2) & -\sin(\theta/2) \\ 0 & 0 & -\sin(\theta/2) & -\cos(\theta/2) \end{pmatrix}. \tag{G.3}$$

To perform the analysis on graphs, we express the $RY(\theta)$–CZ modification as an MB pattern (shown in Fig. G.1) that is "inserted" into the edges of the 1×1 graph [Fig. 2.3(a)][1]

---

[1]Further processing of each modification yields the LC-reduced graph shown in Fig. 2.4(a).

The unitary of the modified graph may then be expressed in an similar form:

$$\hat{U}_{MB}(\theta_1, \theta_2, \theta_3) = H_2 H_3 H_4 \hat{U}_{1,4}^{am,MB}(\theta_3) \hat{U}_{1,3}^{am,MB}(\theta_2) \hat{U}_{1,2}^{am,MB}(\theta_1) H_1 H_2 H_3 H_4, \qquad \text{(G.4)}$$

where each edge modification between qubits m and n acts an identical unitary $\hat{U}_{m,n}^{am,MB}(\theta)$. As described above, the problem then boils down to determining this unitary. Here, we treat the initialization and LC operations externally of the pattern (as these can be simulated classically). We may also simplify the pattern obtained from direct concatenation via the ZX-calculus (Ref. [10]), which reduces the number of qubits from 28 to 12. As shown in Fig. G.1, it corresponds to:

$$V = \left[M_{10}(R(2\pi - \theta)) \prod_{i=8}^{9} M_i(Y) \prod_{i=1}^{7} M_i(X) \right] \hat{U}_\Sigma, \qquad \text{(G.5)}$$

$$|\psi_{\text{out}}\rangle_{11,12} = V \prod_{\{j,k\} \in E} CZ_{j,k}(|\psi_{\text{in}}\rangle_{1,2} \otimes |+\rangle_{3,4...,12}^{\otimes 10}), \qquad \text{(G.6)}$$

with $E = \{(1,4),(9,10),(5,6),(5,9),(7,9),(8,10),(3,8),(4,5),(7,12),(6,11),(2,3)\}$.

Method 1: Phase map decomposition. The PMD method regards the MB pattern as acting three distinct maps, each expressed as a matrix:

1. A dilation map $P_{I \to A}$ that expands the Hilbert space of the inputs (corresponding to the addition of ancillary qubits in $|+\rangle$).
2. A diagonal phase map $\Phi_{A \to A}$ that captures the entanglement (CZ-edges) prescribed by the pattern and correspondingly, any -1 and rotation phases incurred.
3. A contraction map $R_{A \to O}$ that reduces the Hilbert space to that defined over the output qubits (corresponding to the ancillary measurements).

The overall unitary transformation $\hat{U}$ is then described as the composition of these maps:

$$\hat{U} = R \circ \Phi \circ P. \qquad \text{(G.7)}$$

The method also assumes the *positive branch* has been taken by the pattern, which is defined as the unique set of measurement outcomes that do not require any corrections (i.e. each $U_\Sigma = \hat{I}$). In this case, it is the "all 0" outcome. Letting n label the computational basis states in the rows and columns, where $n' \equiv bin(n)$, we compute the three maps via the procedures outlined in Ref. [9]. We find that:

$$P_{I \to A}[., n] = |n'\rangle_{1,2} \otimes |+...+\rangle_{3,...,12}; \quad n = 0, 1..., 2^2 - 1, \qquad \text{(G.8)}$$

101

$$\Phi_{A\to A} = Ph(2\pi - \theta)_{10}Ph(\pi/2)_9Ph(-\pi/2)_8 \prod_{\{j,k\}\in E} CZ_{j,k} = diag(d_{1,1}, d_{2,2}, ..., d_{2^{12},2^{12}}),$$

$$\tag{G.9}$$

$$R_{A\to O}[., \, m] = \langle +... + |m'\rangle_{1,2,...,10} \otimes |m'\rangle_{11,12}; \quad m = 0, 1, ..., 2^{12} - 1, \tag{G.10}$$

where $Ph(\theta)$ is the phase operator defined as the matrix $\begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix}$, with

$$d_{n,n} = e^{-i[(2\pi-\theta)n'_{10}+(\pi/2)(n'_9-n'_8)]}(-1)^{\sum_{jk\in E} n'_j n'_k}. \tag{G.11}$$

Multiplying the matrices of Eqs. (G.8), (G.9), and (G.10) together, we finally obtain:

$$\therefore \hat{U}[n_R, n_P] = R_{A\to O}[n_R, .] \circ \Phi_{A\to A} \circ P_{I\to A}[., \, n_P] \tag{G.12}$$

$$\implies \hat{U}^{am,MB}(\theta) = \begin{pmatrix} \hat{U}[0,0] & \hat{U}[0,1] & \hat{U}[0,2] & \hat{U}[0,3] \\ \hat{U}[1,0] & \hat{U}[1,1] & \hat{U}[1,2] & \hat{U}[1,3] \\ \hat{U}[2,0] & \hat{U}[2,1] & \hat{U}[2,2] & \hat{U}[2,3] \\ \hat{U}[3,0] & \hat{U}[3,1] & \hat{U}[3,2] & \hat{U}[3,3] \end{pmatrix} \tag{G.13}$$

$$= 2^{-5}e^{-i\theta/2} \begin{pmatrix} \cos(\theta/2) & -\sin(\theta/2) & 0 & 0 \\ \sin(\theta/2) & \cos(\theta/2) & 0 & 0 \\ 0 & 0 & \cos(\theta/2) & -\sin(\theta/2) \\ 0 & 0 & -\sin(\theta/2) & -\cos(\theta/2) \end{pmatrix}. \tag{G.14}$$

Method 2: Measurement-based framework. As a further check, we may also determine $\hat{U}^{am,MB}$ via the MB-QC framework discussed in App. A. Since the framework works on the statevector level and not directly with operators, we must construct the unitary from its constituent columns and determine the relative phases that exist between them. We perform this method as a four-step procedure on Eq. (G.5):

1. First, we process the pattern up to the measurement of the adaptive qubit(s) (qubit 10), with the non-adaptive portion (1-9) simulated classically. The resulting state is:

$$|\psi\rangle_{10,11,12} = \left(\prod_{i=1}^{9} M_i(Z) \right) \left((H_{10} RZ(2\pi - \theta)_{10})(H_9 S_9^\dagger)(H_8 S_8) \bigotimes_{i=1}^{7} H_i \right)$$
$$\cdots \prod_{\{j,k\} \in E} CZ_{j,k} \left(|\psi_{\text{in}}\rangle_{1,2} \otimes |+\rangle_{3,4,...,12}^{\otimes 10} \right), \tag{G.15}$$

where the X, Y, and $R(\theta)$ measurements are re-expressed as computational (Z) basis measurements up to transformation operations.

2. Next, we learn the relative phase factor incurred on $|\psi\rangle_{10,11,12}$ when $|\psi_{\text{in}}\rangle_{1,2}$ is set to each two-qubit computational basis state (i.e. $|00\rangle, |01\rangle, |10\rangle, |11\rangle$). Each corresponding $|\psi\rangle$ forms one of the four orthonormal columns of the unitary matrix. For a basis state $|b\rangle$ defined over all pattern qubits, the action of each operation above (H, CZ, S, RZ) on their constituent $|0\rangle$ or $|1\rangle$ states incurs a specific phase on them. Denoting t and c as control and target qubits respectively, these operations transform the states as:

$$
\begin{aligned}
(+1)|b_t\rangle &\xrightarrow{H_t} (+1)|0_t\rangle + (-1)^{b_t}|1_t\rangle \\
(+1)|b_t\rangle &\xrightarrow{S_t} (+i)^{b_t}|b_t\rangle \\
(+1)|b_t\rangle &\xrightarrow{S_t^\dagger} (-i)^{b_t}|b_t\rangle \\
(+1)|b_t\rangle &\xrightarrow{CZ_{c,t}} \begin{cases} (+1)^{b_t}|b_t\rangle, & b_c = |0\rangle \\ (-1)^{b_t}|b_t\rangle, & b_c = |1\rangle \end{cases} \\
(+1)|b_t\rangle &\xrightarrow{RZ(\theta)_t} e^{(-1)^{1-b_t}(i\theta/2)|b_t\rangle},
\end{aligned} \tag{G.16}
$$

where $b_t \in \{0, 1\}$. Over all $|\psi\rangle$, we determine the first nonzero entries with identical amplitudes (from the top of the vectors)[2] and record the phase differences between them. In the case of Eq. (G.15), we find that for $\theta > 0$[3]:

---

[2]In other words, with the smallest binary ket labels. We select these entries for comparing phases since additional phase differences can exist further down the statevector, even if the topmost elements have the same phases. These entries generally correspond to ket labels with a greater number of '1's, and therefore have extra phases tacked on by the operations in Eq. (G.16).

$$|00\rangle_{1,2} \mapsto -2^{-3/2} (e^{\theta/2} + e^{-\theta/2})|000\rangle_{10,11,12} \implies \text{phase} = +1$$
$$|01\rangle_{1,2} \mapsto -(i)2^{-3/2} (e^{\theta/2} + e^{-\theta/2})|001\rangle_{10,11,12} \implies \text{phase} = +i$$
$$|10\rangle_{1,2} \mapsto -2^{-3/2} (e^{\theta/2} + e^{-\theta/2})|010\rangle_{10,11,12} \implies \text{phase} = +1 \tag{G.17}$$
$$|11\rangle_{1,2} \mapsto -(i)2^{-3/2} (e^{\theta/2} + e^{-\theta/2})|011\rangle_{10,11,12} \implies \text{phase} = +i,$$

where the phases are taken relative to the $|00\rangle$ input, and we have imposed that all measured qubits (1-9) yield an outcome of '0' (positive branch).

3. For each computational basis input, we then process the remaining adaptive portion in an MB fashion and obtain the corresponding output $|\psi_{\text{out}}\rangle$. The outputs are:

$$|00\rangle_{1,2} \mapsto |\psi_{out,MB}\rangle_{11,12} = \{-\cos(\theta/2), -\sin(\theta/2), 0, 0\}^T$$
$$|01\rangle_{1,2} \mapsto |\psi_{out,MB}\rangle_{11,12} = \{-i\sin(\theta/2), i\cos(\theta/2), 0, 0\}^T$$
$$|10\rangle_{1,2} \mapsto |\psi_{out,MB}\rangle_{11,12} = \{0, 0, -\cos(\theta/2), \sin(\theta/2)\}^T \tag{G.18}$$
$$|11\rangle_{1,2} \mapsto |\psi_{out,MB}\rangle_{11,12} = \{0, 0, -i\sin(\theta/2), -i\cos(\theta/2)\}^T.$$

4. Finally, we multiply each $|\psi_{\text{out}}\rangle$ by the relative phase factor corresponding to its basis input $|\psi_{\text{in}}\rangle$. We then collect all $|\psi_{\text{out}}\rangle$ together as the columns of a matrix, ordering them according to their input states (i.e. $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$, from left to right). This matrix is effectively $\hat{U}^{am,MB}(\theta)$. Performing this step on Eq. (G.18), we obtain:

$$\hat{U}^{am,MB}(\theta)$$
$$= \begin{pmatrix} (+1)(-\cos(\theta/2)) & (+i)(-i\sin(\theta/2)) & 0 & 0 \\ (+1)(-\sin(\theta/2)) & (+i)(i\cos(\theta/2)) & 0 & 0 \\ 0 & 0 & (+1)(-\cos(\theta/2)) & (+i)(-i\sin(\theta/2)) \\ 0 & 0 & (+1)(\sin(\theta/2)) & (+i)(-i\cos(\theta/2)) \end{pmatrix} \tag{G.19}$$

$$= (-1)\begin{pmatrix} \cos(\theta/2) & -\sin(\theta/2) & 0 & 0 \\ \sin(\theta/2) & \cos(\theta/2) & 0 & 0 \\ 0 & 0 & \cos(\theta/2) & -\sin(\theta/2) \\ 0 & 0 & -\sin(\theta/2) & -\cos(\theta/2) \end{pmatrix}. \tag{G.20}$$

Since Eqs. (G.14) and (G.20) are identical to Eq. (G.3) (up to global phases), it follows that $\hat{U}_{MB}(\theta_1, \theta_2, \theta_3) \equiv \hat{U}_{GB}(\theta_1, \theta_2, \theta_3)$. Thus, equivalency between the MB and GB paradigms has been shown.

---

[3]For $\theta < 0$, Eqs. (G.17) and (G.20) incur a global phase of $+1$ instead of -1.

Appendix H

Performing Coherent MBQC

In this appendix, we give further details regarding the treatment of the generalized cSWAPs in the MB version of the coherent superposition protocol. Recall from Eq. (3.14) that the operation swaps the target input $|\psi_0\rangle$ with the auxiliary state $|\phi_0\rangle_i$ when the d-level control is in $|i\rangle$; $i = 1, 2, ..., d-1$. For a control qubit ($d = 2$), the operation reduces to the Fredkin which can be translated into a MB cluster using its decomposition of seven T gates (Ref. [65]). For a control *qudit*, one may resort to a multi-qubit implementation as in the GB version (i.e. with $\log_2 d$ qubits). However, the direct sum structure of the ensuing implementation does not translate to known cluster states that encapsulate the whole operation. This motivates the following alternative approach, where we regard the control as an external GB degree of freedom but use it to swap entire MB patterns in a quantum switch-like manner. For simplicity, we will omit the m extra input registers needed to calculate the CJ fidelity, however the procedure described below remains unchanged.

Given an m-qubit input, the generalized cSWAP matrix consists of d block diagonals (each of size 2^{md}) of the form:

$$
\begin{pmatrix}
\bigotimes_j I_{|\psi_0^{(j)}\rangle} \bigotimes_{i,j} I_{|\phi_0^{(j)}\rangle_i} & & & \\
& \bigotimes_j \mathrm{SWAP}_{|\psi_0^{(j)}\rangle, |\phi_0^{(j)}\rangle_1} \bigotimes_{\substack{i,j \\ i \neq 1}} I_{|\phi_0^{(j)}\rangle_i} & & \\
& & \ddots & \\
& & & \bigotimes_j \mathrm{SWAP}_{|\psi_0^{(j)}\rangle, |\phi_0^{(j)}\rangle_d} \bigotimes_{\substack{i,j \\ i \neq d}} I_{|\phi_0^{(j)}\rangle_i}
\end{pmatrix}
$$

(H.1)

where $i = (1, 2, .., d)$ and $j = (1, 2, ..., m)$. We analyze the effect of noise on each block

separately (see below), by executing each of its SWAP and I operations as clusters [2]. For the former, we employ the decomposition $\text{SWAP}_{c,t} = CX_{c,t}CX_{t,c}CX_{c,t}$. Acting local noise in the form of Eq. (3.35), we obtain the effective action of the quantum map for some arbitrary ρ via the CJ-isomorphism. Specifically, we input the Bell state $|\Phi^+\rangle$ ($|\Phi_2^+\rangle$), apply noise to the cluster implementing I (SWAP), and obtain the output Choi matrix η. The action of the quantum map is given by:

$$\mathcal{E}(\rho) = \text{Tr}_A(I \otimes \rho_B^T)\eta_{AB}, \tag{H.2}$$

where A and B are the two systems denoted Alice and Bob, with Bob's system undergoing the noisy computation. We may then characterize the noise, first by perform a spectral decomposition of η_{AB} and then determining the Kraus operators corresponding to the channel. This is performed as such:

$$\eta_{AB} = \sum_n c_n|\phi_n\rangle\langle\phi_n| \rightarrow |\phi_n\rangle = [\phi_n^{(1)}, .., \phi_n^{(2^{|AB|})}]^T \tag{H.3}$$

$$\implies K_n = \frac{c_n}{\mathcal{N}} \begin{pmatrix} \phi_n^{(1)} & \cdots & \phi_n^{(2^{|AB|/2})} \\ \vdots & \ddots & \vdots \\ \phi_n^{(2^{|AB|/2}(2^{|AB|/2}-1))} & \cdots & \phi_n^{(2^{|AB|})} \end{pmatrix} \tag{H.4}$$

$$\implies \mathcal{E}(\rho) = \sum_n K_n\rho K_n^\dagger; \quad n = 1, 2, ...2^{|AB|}, \tag{H.5}$$

where we have undone the vectorization in Eq. (H.4), $\sum_n |c_n|^2 = 1$, and $\mathcal{N}$ is a normalization factor. If there are n operations in a block, we model the overall noise by composing each of their individual channels sequentially, i.e.: $\mathcal{E}_1(\rho) \circ \mathcal{E}_2(\rho)... \circ \mathcal{E}_n(\rho)$. For example, if the Kraus operators corresponding to noisy unitaries U_A and U_B are A and B, then:

$$\mathcal{E}_A(\rho) \circ \mathcal{E}_B(\rho) = \sum_{m,n}(B_nA_m)\rho(B_nA_m)^\dagger. \tag{H.6}$$

Now, let $C^{(i)}$ be the overall Kraus operators in the ith block of Eq. (H.1). Then, the state ρ_{cSWAP} after applying the generalized cSWAP to ρ (defined over the control in $|+\rangle_c$, and all inputs and auxiliaries), is calculated as:

$$\rho_{\text{cSWAP}} = \begin{pmatrix} \sum_n C_n^{(1)}\rho^{(1,1)}C_n^{(1)\dagger} & \sum_n C_n^{(1)}\rho^{(1,2)}C_n^{(2)\dagger} & \cdots & \sum_n C_n^{(1)}\rho^{(1,d)}C_n^{(d)\dagger} \\ \vdots & \ddots & \vdots \\ \sum_n C_n^{(d)}\rho^{(d,1)}C_n^{(1)\dagger} & \sum_n C_n^{(d)}\rho^{(d,2)}C_n^{(2)\dagger} & \cdots & \sum_n C_n^{(d)}\rho^{(d,d)}C_n^{(d)\dagger} \end{pmatrix}, \tag{H.7}$$

106

where $\rho^{(i,j)}$ is the ith row and j^{th} column partition of ρ (block matrix of size 2^{md}), and the off-diagonal elements describe interferences between the channels of each path. For bookkeeping purposes, each set of Kraus operators is arranged by their eigenvalues c_n from least to greatest, allowing us to pair the nth operators between any two sets. Furthermore, the hermicity of density matrices implies $\sum_n C_n^{(a)} \rho^{(a,b)} C_n^{(b)\dagger} = \sum_n C_n^{(b)} \rho^{(b,a)} C_n^{(a)\dagger}$, meaning that only the lower triangular portion of Eq. (H.7) needs to be computed. Despite the total number of Kraus operators in the approach exponentially scaling in md, we note that the generalized cSWAPs are performed only *twice* regardless of the size of the main computation U. This computational advantage is further bolstered by the fact that the size of the cluster implementing the cSWAPs scales linearly as $\mathcal{O}(md + \log_2 d)$.

Although direct manipulation of matrices is not possible in a physical experiment (instead, expensive process tomography is often required to characterize noise), it facilitates direct numerical analyses. We have also assumed that the noises between and within each block in Eq. (H.1) are uncorrelated. This is reasonable if the control system is spatially far away from the main system enacting the SWAPs (e.g. implemented on a different apparatus) such that the ensuing channels are weakly correlated. Related setups involving photonics and trapped ions have been explored in Ref. [132], where the control is treated as a black-box with no prior knowledge of the unitaries, and each unitary is realized on a separate physical apparatus.

As a clarifying note, if noise affects the generalized cSWAPs, then attention must be given to ensuring that noise acts exactly *once* on every cluster qubit in the protocol. Based on how the operations are concatenated in Fig. 3.4, we therefore act noise on:

- All qubits of the 1st generalized cSWAP cluster

- Body and output qubits of each cluster in U, except the last one

- Body qubits only for the last cluster in U

- All qubits of the 2nd generalized cSWAP cluster

For *noiseless* cSWAP runs, we simply act noise on every qubit for each operation in U. Finally, it may be observed that the results presented in Fig. 3.5 are generally better than their counterparts in Fig. 10 of Ref. [65]. This stems from simplified versions of the clusters used for the cNOT and T operations, which contain fewer qubits compared to the patterns prescribed in Ref. [2]. Likewise, for $d = 2$, the use of SWAPs involves significantly fewer cluster qubits than the Fredkin. However, the main conclusions and underlying physics remain unchanged.