Efficient Long-Range Entanglement Using Dynamic Circuits

Elisa Bäumer[1,*] Vinay Tripathi,[2,3] Derek S. Wang,[2] Patrick Rall,[4] Edward H. Chen,[5,6]
Swarnadeep Majumder,[2] Alireza Seif,[2] and Zlatko K. Minev[2]

[1] *IBM Quantum, IBM Research—Zurich, Rüschlikon 8803, Switzerland*
[2] *IBM Quantum, IBM T.J. Watson Research Center, Yorktown Heights, New York 10598, USA*
[3] *Department of Physics & Astronomy, University of Southern California, Los Angeles, California 90089, USA*
[4] *IBM Quantum, MIT-IBM Watson AI Lab, Cambridge, Massachusetts 02142, USA*
[5] *IBM Quantum, Almaden Research Center, San Jose, California 95120, USA*
[6] *IBM Quantum, Research Triangle Park, North Carolina 27709, USA*

(Received 31 October 2023; revised 23 May 2024; accepted 28 June 2024; published 22 August 2024)

Quantum simulation traditionally relies on unitary dynamics, inherently imposing efficiency constraints on the generation of intricate entangled states. In principle, these limitations can be superseded by nonunitary, dynamic circuits. These circuits exploit measurements alongside conditional feed-forward operations, providing a promising approach for long-range entangling gates, higher effective connectivity of near-term hardware, and more efficient state preparations. Here, we explore the utility of shallow dynamic circuits for creating long-range entanglement on large-scale quantum devices. Specifically, we study two tasks: controlled-NOT gate teleportation between up to 101 qubits by feeding forward 99 midcircuit measurement outcomes, and the preparation of Greenberger–Horne–Zeilinger states with genuine entanglement. In the former, we observe that dynamic circuits can outperform their unitary counterparts. In the latter, by tallying instructions of compiled quantum circuits, we provide an error budget detailing the obstacles that must be addressed to unlock the full potential of dynamic circuits. Looking forward, we expect dynamic circuits to be useful for generating long-range entanglement in the near term on large-scale quantum devices.

DOI: 10.1103/PRXQuantum.5.030339

I. INTRODUCTION

Quantum systems present two distinct modes of evolution: deterministic unitary evolution, and stochastic evolution as the consequence of quantum measurements. To date, quantum computations predominantly utilize unitary evolution to generate complex quantum states for information processing and simulation. However, due to inevitable errors in current quantum devices [1], the computational reach of this approach is constrained by the depth of the quantum circuits that can realistically be implemented on noisy devices. The introduction of nonunitary dynamic circuits, also called adaptive circuits or LAQCC (local alternating quantum classical computation) circuits [2], can not only implement more general quantum channels, but may also be able to overcome some of these limitations by employing midcircuit measurements and feed-forward operations. As classical computation and communication are viewed as essentially free compared to quantum operations, such conditional operations are a necessary ingredient for quantum error correction (see, e.g., Ref. [3]). In the near term, dynamic circuits present a promising avenue for generating long-range entanglement, a task at the heart of quantum algorithms [4,5]. This includes both implementation of long-range entangling gates that, due to local connectivity among the qubits in many quantum platforms, can require deep unitary quantum circuits, and preparation of many-qubit entangled [6,7] and topologically ordered quantum states [8–16].

From a physical standpoint, the entanglement needs to propagate across the entire range between the qubits. Given that the entanglement cannot spread faster than its information light cone [17,18], entangling two qubits that are a distance n apart requires a minimum two-qubit gate depth that scales as $\mathcal{O}(n)$, and even when assuming all-to-all connectivity, the generation of entanglement over n qubits necessitates a minimum two-qubit gate depth of $\mathcal{O}(\log n)$. Thus, the task becomes challenging when applying only unitary gates. Using dynamic circuits, the

*Contact author: eba@zurich.ibm.com

spread of information can be mostly conducted by classical calculations, which can be faster and with a higher fidelity than the quantum gates, and long-range entanglement can be created in a shallow quantum circuit [19–21], i.e., the depth of quantum gates is constant for any n.

While dynamic circuits have been explored in small-scale experiments [22–26], only recently have there been experimental capabilities on large-scale quantum devices. However, most demonstrations (with the exception of, e.g., Refs. [27–30]) have utilized postselection [31] or postprocessing [32,33] instead of feed forward to prepare entangled states. Such approaches enable the study of properties of the state prepared in isolation, but have limited applicability when the state preparation is part of a larger quantum information processing task.

Here, we explore the utility of shallow dynamic circuits for creating long-range entanglement on large-scale superconducting quantum devices. In Sec. II, we demonstrate an advantage with dynamic circuits by teleporting a long-range entangling controlled-NOT (CNOT) gate over up to 101 locally connected superconducting qubits. We also discuss how this approach can be generalized to more complex gates, such as the three-qubit Toffoli gate. Then, in Sec. III, we prepare a long-range entangled state, the Greenberger–Horne–Zeilinger (GHZ) state [6], with a dynamic circuit. We show that—with a composite error mitigation stack customized for the hardware implementation of dynamic circuits—we can prepare genuinely entangled GHZ states, but fall short of state-of-the-art system sizes achieved with unitary gates due to hardware limitations. We predict conditions under which dynamic circuits should be advantageous over unitary circuits based on our error budget calculation.

II. CNOT GATE TELEPORTATION

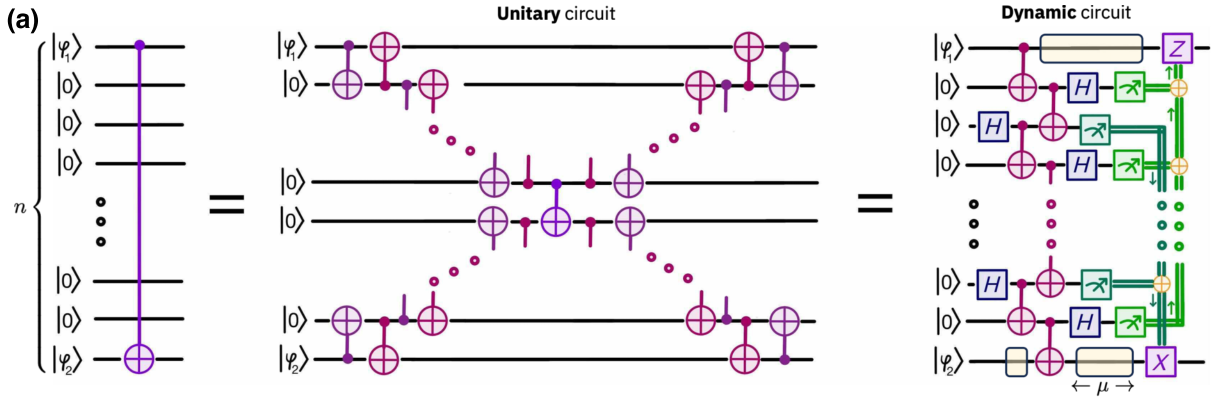
The limited connectivity between qubits in many quantum computational platforms can result in the compilation of nonlocal unitary circuits into deep and error-prone unitary circuits. A potential solution is the use of shallow dynamic circuits. The crucial ingredient for such protocols is long-range CNOT gates from the first to nth qubit, as shown on the left in Fig. 1(a). In the following, we demonstrate a regime under which dynamic circuits enable higher-fidelity long-range CNOT gates *via* gate teleportation. We first describe the dynamic circuit and compare to its equivalent unitary counterpart. We argue, using a simple error budget, that there exists a regime in which the dynamic circuit implementation has an advantage over the unitary one; see Fig. 1(b). Then, using up to 101 qubits on a superconducting processor, we demonstrate a crossover in the fidelity of CNOT gate teleportation, where dynamic circuits perform better for entangling qubits over longer ranges; see Fig. 1(c). This gate teleportation scheme enables an effective all-to-all connectivity in devices with a more limited connectivity, such as those on a heavy-hexagonal lattice. By using some of the qubits as ancillas for measurement and classical feed-forward operations, the ancilla qubits form a bus that connects all system qubits with each other. Therefore, by sacrificing some of the qubits in a large device with limited connectivity, we gain effective access to an all-to-all connected device with fewer qubits; see Fig. 1(d). As this effective all-to-all connectivity limits the parallelization of gates, the orange system qubits could be sacrificed as ancilla qubits as well to further parallelize gate execution with increased connectivity. In addition, a clever compilation could increase parallelization, as, e.g., shown in Fig. 10 in Appendix E, where a long-range controlled-controlled-Z (CCZ) gate could be implemented with two feed-forward operations rather than teleporting all six CNOT gates separately.

We describe the dynamic circuit for CNOT gate teleportation, shown on the right in Fig. 1(a) and derived in Appendix A 1. Importantly, this dynamic circuit can be straightforwardly extended for any number of qubits n (where n is the number of ancillas) such that the depth remains constant for any initial states $|\varphi_1\rangle$ ($|\varphi_2\rangle$) of the control (target) qubit. We expect the error to be dominated by the n midcircuit measurements, $n+1$ CNOT gates parallelized over two gate layers, and idle time mostly over the classical feed-forward time. Note that in this particular realization, each of the n ancilla qubits between the two system qubits must be in state $|0\rangle$. Therefore, during the course of the gate teleportation, the ancillas cannot also be used as memory qubits, further motivating the division of qubits into system and sacrificial ancilla qubits in Fig. 1(d).

We also present an equivalent, low-error unitary counterpart in the middle of Fig. 1(a). (In Appendix B, we propose several different unitary implementations of the long-range CNOT gate. Based on experimental results, as well as the noise model described in Appendix G that gives rise to the error budget described in Appendix B 2, we select this one.) In this unitary realization, the system qubits are connected by a bus of ancilla qubits that are initialized in and returned to the $|0\rangle$ state, just as in its dynamic counterpart. In our particular compilation, throughout the execution of the circuit, qubits that are not in the $|\phi_1\rangle$ or $|\phi_2\rangle$ state are in the $|0\rangle$ state. Doing so minimizes both decoherence and cross-talk errors intrinsic to our superconducting qubit design, as heuristically we learned that the noise affecting our qubits is primarily limited to amplitude damping, dephasing, and ZZ cross-talk errors on neighboring qubits, which implies essentially no idling errors on qubits in the $|0\rangle$ state. Therefore, relative to the dynamic version, there is no error due to idle time or midcircuit measurements, although there are about 4 times more CNOT gates.

A summary of the error budgets for the dynamic and unitary circuits is presented in Fig. 1(b). Based on this table, we expect that dynamic circuits should be advantageous

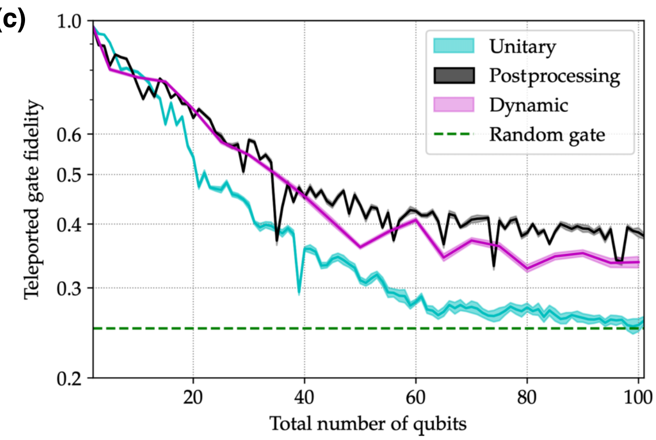
(a)



(b)

CNOT circuit	Idle time error (per CNOT gate time)	CNOT gate	Additional measurements
Unitary	0	$4n+1$	0
Post-processing	2	$n+1$	n
Dynamic	$2\mu+2$	$n+1$	n

(c)



(d)

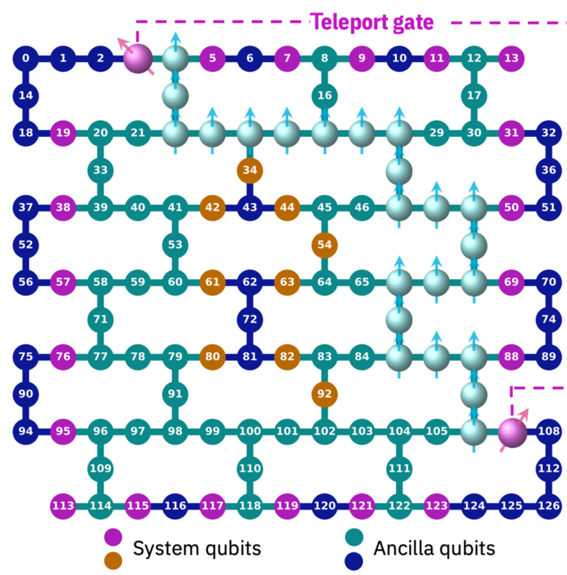


FIG. 1. Teleporting a CNOT gate for long-range entanglement. (a) Left: circuit for a long-range CNOT gate spanning a one-dimensional (1D) chain of n qubits subject to nearest-neighbor connections only. Middle: equivalent unitary decomposition into implementable CNOT gates; circuit depth $\mathcal{O}(n)$. Right: equivalent circuit employing measurements with feed-forward operations; circuit depth $\mathcal{O}(1)$. If the postmeasurement state is unused, feed-forward operations can be handled in postprocessing, eliminating the need for their experimental implementation. Yellow regions indicate the idle time during CNOT gates on other qubits as well as during measurement and feed forward (which is denoted by duration μ). (b) Error model inputs for unitary, measurement-based, and dynamic-circuit CNOT protocols comprise the total number of nonzero idle-block times, CNOT gates, and additional measurements. (c) Experimental results, where dynamic circuits offer improved fidelity for CNOT gate teleportation across a qubit chain $\gtrsim$10 qubits. (d) Map of a 127-qubit heavy-hexagonal processor, `ibm_sherbrooke`, overlaid with system configurations for long-range gate teleportation across a locally connected bus. To establish an effective all-to-all connectivity, we show one possible strategy of dividing the qubits into system (purple and orange) and sacrificial ancilla (turquoise and blue for extra connections) qubits. To parallelize gate execution with increased connectivity, orange qubits can be used as ancillas. We show how a particular long-range CNOT gate can be implemented through an ancilla bus marked as turquoise spins.

over unitary circuits if the additional n midcircuit measurements in the dynamic circuit introduce less error than the $3n$ extra CNOT gates in the unitary circuit, assuming that n is large enough such that the idling error μ incurred during measurement and classical feed forward in the dynamic circuit is relatively small. Importantly, we should note that these error analyses only consider the gate error on the two respective qubits, but not the error introduced on other qubits, which we expect to be much larger in the unitary case due to the linear depth. Thus, the constant-depth dynamic circuit might be even more advantageous than what we can see from the gate fidelity.

To determine the experimental gate fidelity, let our ideal unitary channel be $\mathcal{U}(\rho) := U\rho U^\dagger$ and its noisy version be $\tilde{\mathcal{U}}(\rho) := \mathcal{U}[\Lambda(\rho)]$, where Λ is the effective gate noise channel and ρ is a quantum state. The average gate fidelity of the noisy gate is $\mathcal{F}_{\text{avg}}(\mathcal{U},\tilde{\mathcal{U}}) = \int \mathrm{d}\psi\ \mathrm{Tr}[\mathcal{U}(\rho_\psi)\tilde{\mathcal{U}}(\rho_\psi)]$, where the Haar average is taken over the pure states $\rho_\psi = |\psi\rangle\langle\psi|$. This fidelity can be faithfully estimated from Pauli measurements on the system, using *Monte Carlo process certification* [34,35], as detailed in Appendix C 2.

The results from a superconducting quantum processor are shown in Fig. 1(c). The implementation details can be found in Appendix D 1. As expected, for a small number of qubits $n \lesssim 10$, the unitary implementation yields the best fidelities. However, for increasing n, it converges much faster to the fidelity of a random gate (0.25) than the dynamic circuit implementation, which converges to a value slightly below 0.4. These align well with the error budget analysis in Appendix B 2 and the noise model predictions depicted in Appendix G. Note that, in the limit of large n, the fidelities of the measurement-based scheme are limited by the Z and X corrections on $|\phi_1\rangle$ and $|\phi_2\rangle$ [see Fig. 1(a)]. A straightforward derivation using this noise model shows that the minimum possible process fidelity due to only incorrect Z and X corrections (without the fixed infidelity from the idle time and CNOT gates) is 0.25, which converts to a gate fidelity of 0.4.

The measurement-based protocol with postprocessing performs slightly better than the dynamic circuits as the former does not incur errors from the classical feed forward, allowing us to isolate the impact of classical feed forward from other errors, such as the $n+1$ intermediate CNOT gates and midcircuit measurements. Note, however, that the postprocessing approach is generally not scalable if further circuit operations follow the teleported CNOT gate due to the need to simulate large system sizes, further emphasizing the advantage of dynamic circuits as errors rooted in classical feed forward are reduced. Overall, we find that CNOT gates over large distances are more efficiently executed with dynamic circuits than unitary ones.

In Appendix E we show that these ideas can be generalized to teleporting multiqubit gates, such as the Toffoli or CCZ gate. Compiling them more efficiently than simply implementing multiple teleported CNOT gates, we expect their shallow implementation with dynamic circuits to be even more advantageous over their unitary counterpart, especially for large n.

III. STATE PREPARATION: GHZ

Dynamic circuits can also be used to prepare long-range entangled states. A prototypical example is the GHZ state [6], shown schematically in Fig. 2(a). While it can be created using only Clifford gates and thus can be simulated efficiently on a classical computer [36], it becomes nonsimulatable when followed by a sufficient number of non-Clifford gates in a larger algorithm, or when inserted as a crucial ingredient in, e.g., the efficient compilation of multiqubit gates [37,38].

Here, we show that GHZ states with long-range entanglement can be prepared with dynamic circuits. Although we do not see a clear advantage of dynamic circuits over unitary ones in this case, we provide a detailed description of the challenges that must be addressed to realize such an advantage.

For preparation of a GHZ state on a 1D n-qubit chain, in Fig. 2, we show the equivalence between the unitary circuit (left) and dynamic circuit (right). (For a detailed derivation, see Appendix A 2.) Notably, the unitary equivalent has a two-qubit gate depth that scales as $\mathcal{O}(n)$ with quadratically increasing idle time and $n-1$ total CNOT gates, while the depth of the dynamic circuits remains constant with linearly increasing idle time, $3n/2-1$ total CNOT gates, and $n/2-1$ midcircuit measurements [see Fig. 2(c)]. The dynamic circuit incurs less idle time and fewer two-qubit gate depths at the cost of increased CNOT gates and midcircuit measurements. Therefore, we expect dynamic circuits to be advantageous for large system sizes n and low errors in the midcircuit measurement. For a more detailed analysis of the error budget, see Appendix F 1.

We explore whether current large-scale superconducting quantum devices enable an advantage with dynamic circuits for preparation of the entangled GHZ state. To efficiently verify the preparation of a quantum state σ, we use the Monte Carlo state certification that samples from Pauli operators with nonzero expectation values, as implemented in Ref. [31] and described in detail in Appendix C 1. As the n-qubit GHZ state is a stabilizer state, we can randomly sample m of the 2^n stabilizers $\{S_i\}_{i=1,\dots,2^n}$ and approximate the fidelity by $F = (1/m)\sum_{k=1}^{m}\langle S_k\rangle_\sigma + \mathcal{O}(1/\sqrt{m})$.

The experimental results of GHZ state preparation with unitary and dynamic circuits are shown in Fig. 2(d). They all include measurement error mitigation on the final measurements [39]. The implementation details can be found in Appendix D 2. On the left, we show the results without dynamical decoupling. In the unitary case, we observe genuine multipartite entanglement, defined as state fidelity $F > 0.5$ [40], within a confidence interval of 95% up to seven qubits with a rapid decay in fidelity with increasing system size mainly due to errors in two-qubit gates and ZZ cross-talk errors during idling time [41]. As these errors are mostly coherent, they lead to an oscillation of the fidelity such that it increases again for higher qubit numbers. To suppress the coherent ZZ errors we apply dynamical decoupling (DD) pulses, as described below.

In the dynamic case, we observe genuine entanglement up to six qubits. Here, we do not find a crossover point after which dynamic circuits have an advantage over unitary circuits. We attribute the performance of dynamic circuits to several factors, including the fact that the current implementation results in an average classical feed-forward time that scales with the number of potential midcircuit measurement bitstring outcomes, which itself grows exponentially with system size. This limitation appears because the

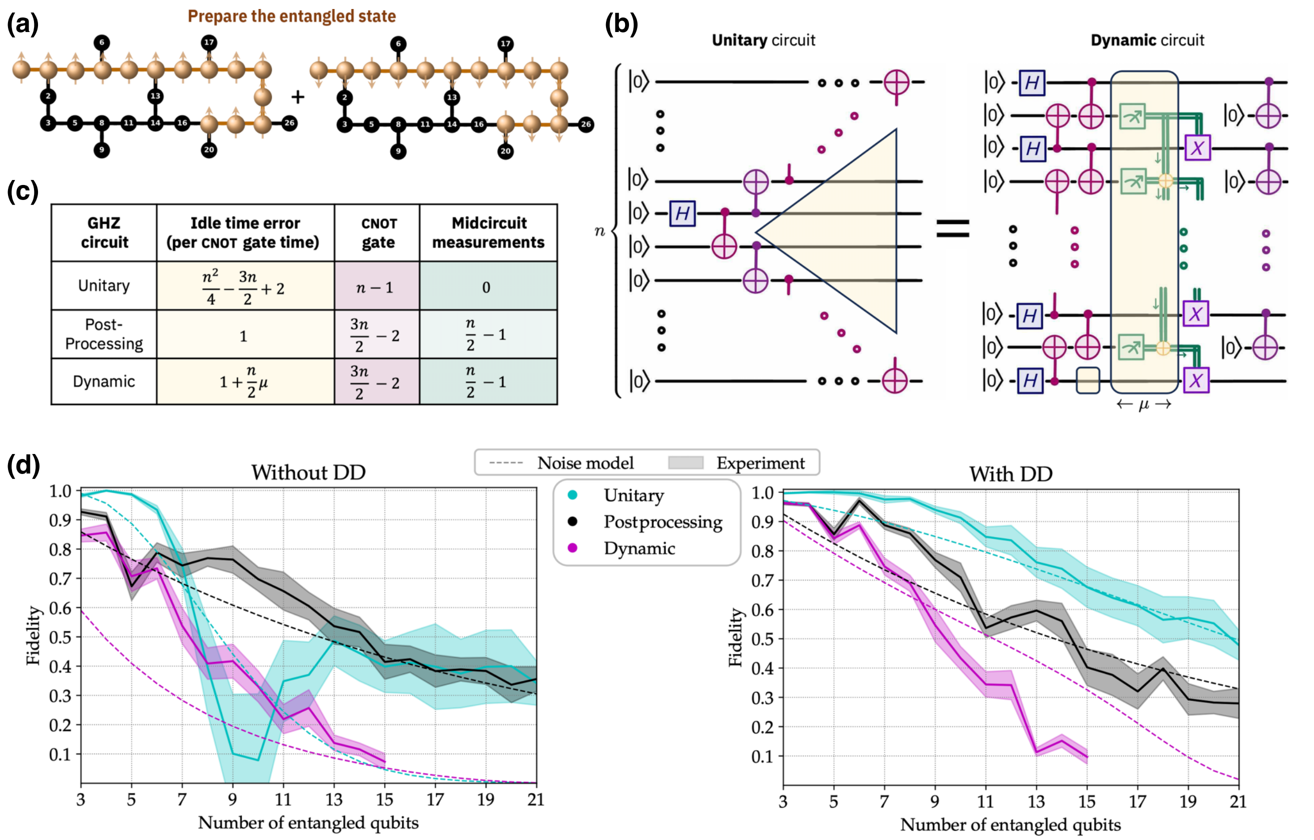


FIG. 2. Preparing long-range entangled states. (a) Illustration of a GHZ state with chosen qubit spins (spheres) in a superposition of “all up” and “all down” polarizations (arrows), overlaid on a quantum processor. (b) Circuits to prepare an n-qubit GHZ state using either a unitary (left) or dynamic (right) circuit. For a 1D qubit chain, the depth of the unitary (respectively, dynamic) circuit scales as $\mathcal{O}(n)$ [respectively, $\mathcal{O}(1)$]. If the final state is not directly used, the feed-forward operations can be implemented in classical postprocessing on the output bits (classically controlled-X gates and resets can be omitted). Yellow regions indicate the idle time during CNOT gates on other qubits as well as during measurement and feed forward (which is denoted by duration μ). (c) Error model inputs for the GHZ preparation circuits. The model incorporates the noisy components of the circuits: nonzero idle circuit periods (yellow), the number of CNOT gates (pink), and the number of midcircuit measurements (green). These parameters are used to derive an error model that yields a lower bound on the protocol fidelity, shown in the following panel. (d) Fidelity of preparing the GHZ state on quantum hardware using unitary, measurement-based postprocessing, or dynamic circuits in the absence or presence of dynamical decoupling (DD). Data shown with dots. Theory curves based on the error model parameters of panel (c) shown with dashed lines.

switch operator is currently testing each possible case of measurement outcomes sequentially, so on average checks half of the cases until it finds the correct one. With our future control software we expect to implement the correct feed-forward operations in constant time. By reducing the error induced by idle time during classical feed forward, we expect dynamic circuits to surpass unitary circuits at $\gtrsim 10$ qubits—we can see this by studying the postprocessing case, which is equivalent to the dynamic circuit implementation except that the classical logic is executed in postprocessing, not during execution of the quantum circuit itself.

On the right of Fig. 2(d), we show the results using DD [42,43]. We observe improved fidelities for both the unitary and dynamic circuit cases, but not for the postprocessing case as there is little error induced by idle times to quench with dynamical decoupling in the first place. For the unitary case, we observe genuine multipartite entanglement up to 17 qubits, more than twice as many compared to the unmitigated unitary case. This result is close to the state of the art on superconducting quantum processors and is limited by the fact that we do not leverage the 2D connectivity of the device, as in Ref. [44]. While the fidelities are improved with DD for dynamic circuits, the improvement is less dramatic. We attribute this difference to two reasons. First, the unitary circuit has a quadratic idling error term in contrast to a leading linear term for dynamic circuits, resulting in comparatively smaller improvement for dynamic circuits with dynamical decoupling. Second, with the current controls, we are not able to apply DD pulses during the classical feed-forward time, which is the main source of idling error in the dynamic circuit. As in the unmitigated case, we observe rapid decay of the fidelity with increasing system size. This can again be partially attributed to exponential growth of the classical feed-forward time. In the future, we expect to reduce this scaling to a constant, in which case we expect drastically improved performance and genuine entanglement up to about 15 qubits. Still, however, we do not expect to observe an advantage with dynamic circuits for preparation of GHZ states over unitary ones. To realize an advantage with dynamic circuits, we require a scenario where the quadratically scaling idle error of the unitary circuit dominates over sufficiently small CNOT and mid-circuit measurement errors; see Appendix F 2 for a more detailed analysis. We anticipate these conditions can be realized through a combination of hardware improvements and the extension of error mitigation techniques, such as probabilistic error cancelation [45,46], toward midcircuit measurements.

IV. CONCLUSION AND OUTLOOK

Dynamic circuits are a promising feature toward overcoming connectivity limitations of large-scale noisy quantum hardware. Here, we demonstrate their potential for efficiently generating long-range entanglement with two useful tasks: teleporting entangling gates over long ranges to enable effective all-to-all connectivity, and state preparation with the GHZ state as an example. For CNOT gate teleportation, we show a regime in which dynamic circuits result in higher fidelities on up to 101 qubits of a large-scale superconducting quantum processor. We leave incorporating this more efficient implementation of long-range entangling gates as a subroutine in another quantum algorithm to future work; potential studies can include simulating many-body systems with nonlocal interactions. As we demonstrate theoretically, gate teleportation schemes can be extended beyond CNOT gates to multiqubit ones, such as the CCZ gate. Its experimental implementation is also a promising project for the future. For state preparation, based on both unmitigated and mitigated hardware experiments, we expect to see the value of dynamic circuits once the classical postprocessing becomes more efficient and the midcircuit measurement errors can be reduced. We plan on revising the experiments as soon as these capabilities are available. We anticipate that further experiments with dynamic circuits and the development of noise models describing them will be vital contributions toward efficient circuit compilation, measurement-based quantum computation, and fault-tolerant quantum computation.

ACKNOWLEDGMENTS

We thank Diego Ristè, Daniel Egger, Alexander Ivrii, and Luke Govia for valuable discussions and feedback. We thank Emily Pritchett, Maika Takita, Abhinav Kandala, and Sarah Sheldon for their help. We also thank Thomas Alexander, Marius Hillenbrand, and Reza Jokar for their support with implementing dynamic circuits.

APPENDIX A: CIRCUIT DERIVATIONS

In the following we show the circuit equivalences of the CNOT gate teleportation [Fig. 1(a)] and the GHZ state preparation [Fig. 2(b)]. We are not claiming any novelty with this “proof,” but just wanted to show the reader how to derive them in an illustrative way. Before, let us start with some features that we will be using.

(a) The Bell state $(|00\rangle + |11\rangle)/\sqrt{2}$ can be illustrated as a so-called “cup,” as shown in Fig. 3(a), We can move gates along wires including along the cup, as in Fig. 3(b).

(b) *Principle of deferred measurement*: a controlled gate followed by a measurement of the controlled qubit results in the same outcome as first performing the measurement and then applying a classically controlled gate as in Fig. 3(c).

(c) While CNOT gates commute when they are conditioned on the same qubit or have the same target qubit, we get an extra gate when they act on the same qubit differently, as shown in Fig. 3(d).

1. Long-range CNOT gate

In Fig. 4 we illustrate a derivation of the CNOT gate teleportation, as exemplified for seven qubits, which can be straightforwardly extended to an arbitrary number of qubits. In the following, we provide explanations for each step of the derivation, labeled by roman numerals in the figure.

(i) In the first step, we observe that entangling, measuring, and resetting the ancilla qubits does not affect the circuit.

(ii) We insert CNOT gates that would cancel each other. From now on we omit writing down the reset of the ancilla qubits following the measurement.

(iii) We move the pink CNOT gates along the Bell states to the respective qubits above. Also, we add Hadamard gates to flip the direction of the orange CNOT gates (except for that at the bottom). Note that we can omit the Hadamard gates right before the measurements, as they are not affecting the other qubits anymore.

(iv) By moving the bottom orange CNOT gate “up” along the Bell state and passing a pink CNOT gate, we get the extra purple CNOT gate.

(v) Moving the new purple CNOT gate “up” along the Bell state, an extra gate appears that cancels with the initial long-range CNOT gate when pushed to the left (and then it is controlled on state $|0\rangle$, so can be omitted as well).

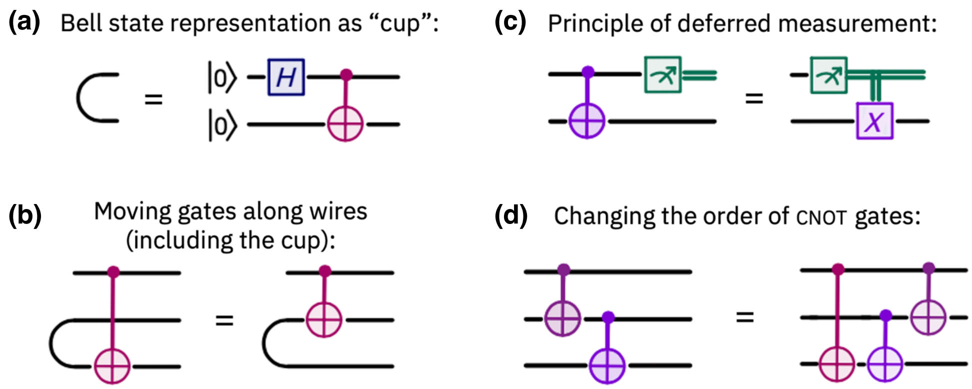


FIG. 3. Useful circuit identities that are used in the illustrative derivation of the CNOT gate teleportation and GHZ state preparation: (a) Bell state representation as "cup", along which we can move gates as shown in (b). (c) Principle of deferred measurement. (d) Commutation relation for CNOT gates.

(vi) Now we make use of the principle of deferred measurement.
(vii) In a final step we merge the classically conditioned gates. The orange ⊕ correspond to XOR gates, i.e., addition mod 2. We also represented the initial Bell states again with their circuit representation.

2. GHZ state preparation

In Fig. 5 we have illustrated a derivation of the GHZ state preparation, exemplary for seven qubits, but it can be straightforwardly extended to an arbitrary number of qubits. In the following, we provide explanations for each step of the derivation, labeled by roman numerals in the figure.

(i) Pushing every second CNOT gate to the very right introduces the extra pink CNOT gates.
(ii) We can omit CNOT gates that are conditioned on state $|0\rangle$.
(iii) As every second qubit is only involved at the very end, we can use those before and reset them.
(iv) A Bell state followed by a CNOT gate results in two uncorrelated qubits in states $|+\rangle$ and $|0\rangle$.
(v) We move the pink CNOT gates along the Bell states to the respective qubits above (they commute with the other CNOT gates they are "passing").
(vi) Pushing the pink CNOT gates to the left through the purple CNOT gates introduces the extra orange CNOT gates.
(vii) We make use of the principle of deferred measurement.
(viii) In a final step we merge the classically conditioned gates. As the classical calculation can be done extremely fast compared to quantum gates, we draw it as a vertical line. The orange ⊕ correspond to XOR gates, i.e., addition mod 2. We also represented the initial Bell states again with their circuit representation.

APPENDIX B: CNOT CIRCUITS

1. Unitary variants

In order to compare the dynamic circuit implementation to a solely unitary one, let us first consider different unitary strategies that might be more or less powerful in different regimes.

a. Strategy I: ancilla-based implementation

We can consider a similar setting as for dynamic circuits, where we place the system qubits in a way that they are connected by a bus of empty ancilla qubits. In this case, we need to swap the system qubits towards each other and back, so that the ancillas are empty in the end again. The swaps can be simplified since the ancillas are empty in the beginning. Here we can divide into different scenarios.

(i) *Circuit Ia.* To minimize the number of CNOT gates, we could swap the controlled qubit all the way to the target qubit and back, which results in the circuit depicted in Fig. 6. Here, a lot of gates cancel, so, given n ancilla qubits, the number of CNOT gates is $2n+1$. However, here the idle time of the qubits while they are not in state $|0\rangle$ equals n^2+2n times the CNOT gate time.
(ii) *Circuit Ib.* In order to decrease the idle time, we could essentially swap both the controlled qubit and the target qubit halfway and back, as illustrated in Fig. 6 (similar to some circuits presented in Refs. [47,48]). In that case, less gates "cancel," so, for n ancilla qubits, we get $3n+1$ CNOT gates, but the idle time reduces to $n^2/4+n$ times the CNOT gate time.
(iii) *Circuit Ic.* If we wanted to reduce the idle time even further, it might be beneficial to not cancel the CNOT gates in scenario 1b, but keep them to bring the swapped qubits back to state $|0\rangle$, as shown in Fig. 6. In that case, we have essentially no idle time (as qubits in state $|0\rangle$ are not prone to idling errors).

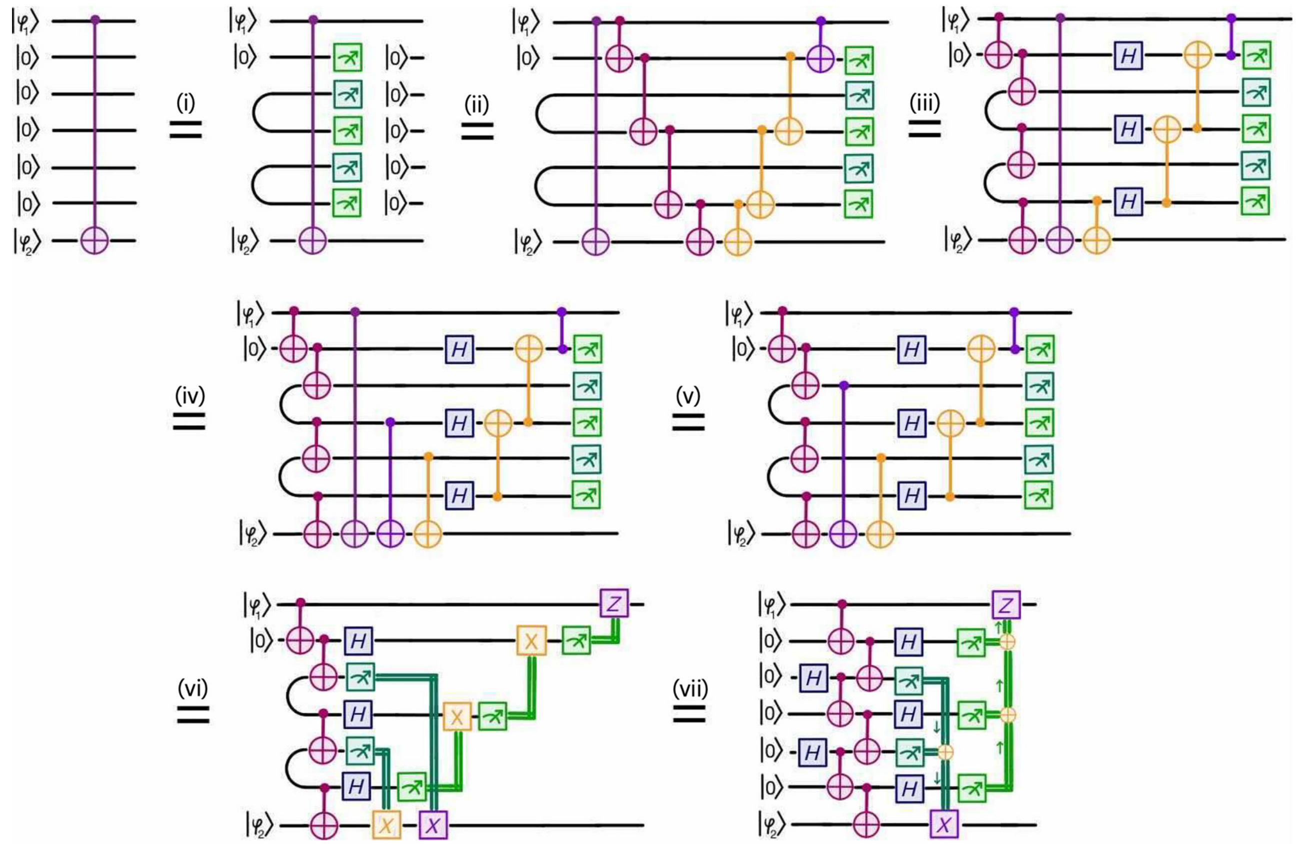


FIG. 4. Graphical derivation for reducing a long-range CNOT gate into gate teleportation executed with measurements and feed-forward operations, i.e., a dynamic circuit. Roman numerals indicate sequential step numbers described in main text.

Here, the number of CNOT gates increased to $4n + 1$ though.

b. Strategy II: SWAP-based implementation without ancillas

This is the case that happens if we just feed our circuit to the transpiler. Here we do not use any ancilla qubits, but only system qubits and apply swaps to move them around. The qubits can be at a different location in the end, so we do not need to swap back. The corresponding circuit is shown in Fig. 6. In this case we require $3\tilde{n} + 1$ CNOT gates and the idle time is $\frac{3}{2}\tilde{n}^2 - 2\tilde{n}$ times the CNOT gate time. However, it is important to note here that the number of qubits lying between the two qubits of interest $\tilde{n}$ is on average much shorter than the number of ancillas between two system qubits in the first scenario. Considering the connectivity illustrated in Fig. 1(c), the relation is approximately $n = 2\tilde{n} + 3$.

2. Error budget

Let us now compare the regimes in which we expect the different implementations to be most useful to demonstrate the benefit of dynamic circuits. In Appendix G we derive a simple noise model that allows us to compute the combined effect of different sources of decoherence as a single Pauli noise rate:

$$\lambda_{\text{tot}} = t_{\text{idle}}\lambda_{\text{idle}} + N_{\text{CNOT}}\lambda_{\text{CNOT}} + N_{\text{meas}}\lambda_{\text{meas}}. \quad \text{(B1)}$$

In Lemma 1 below we show that the final process fidelity is loosely lower bounded by $e^{-\lambda_{\text{tot}}}$. The quantity λ_{tot} combines the following noise sources.

(a) The total amount of time t_{idle} that qubits spend idle within the circuit, and a conversion factor λ_{idle} that quantifies the strength of decoherence. Time t_{idle} is expressed in multiples of the CNOT gate time (i.e., $t_{\text{idle}} = 3$ for three CNOT gate times). The time for a midcircuit measurement, including the additional time waiting for feedback, is defined as μ times the time for a CNOT gate.

(b) The total number of CNOT gates N_{CNOT} and an average Pauli noise rate λ_{CNOT} per CNOT gate.

(c) The total number of midcircuit measurements N_{meas} and an average Pauli noise rate λ_{meas} per measurement.

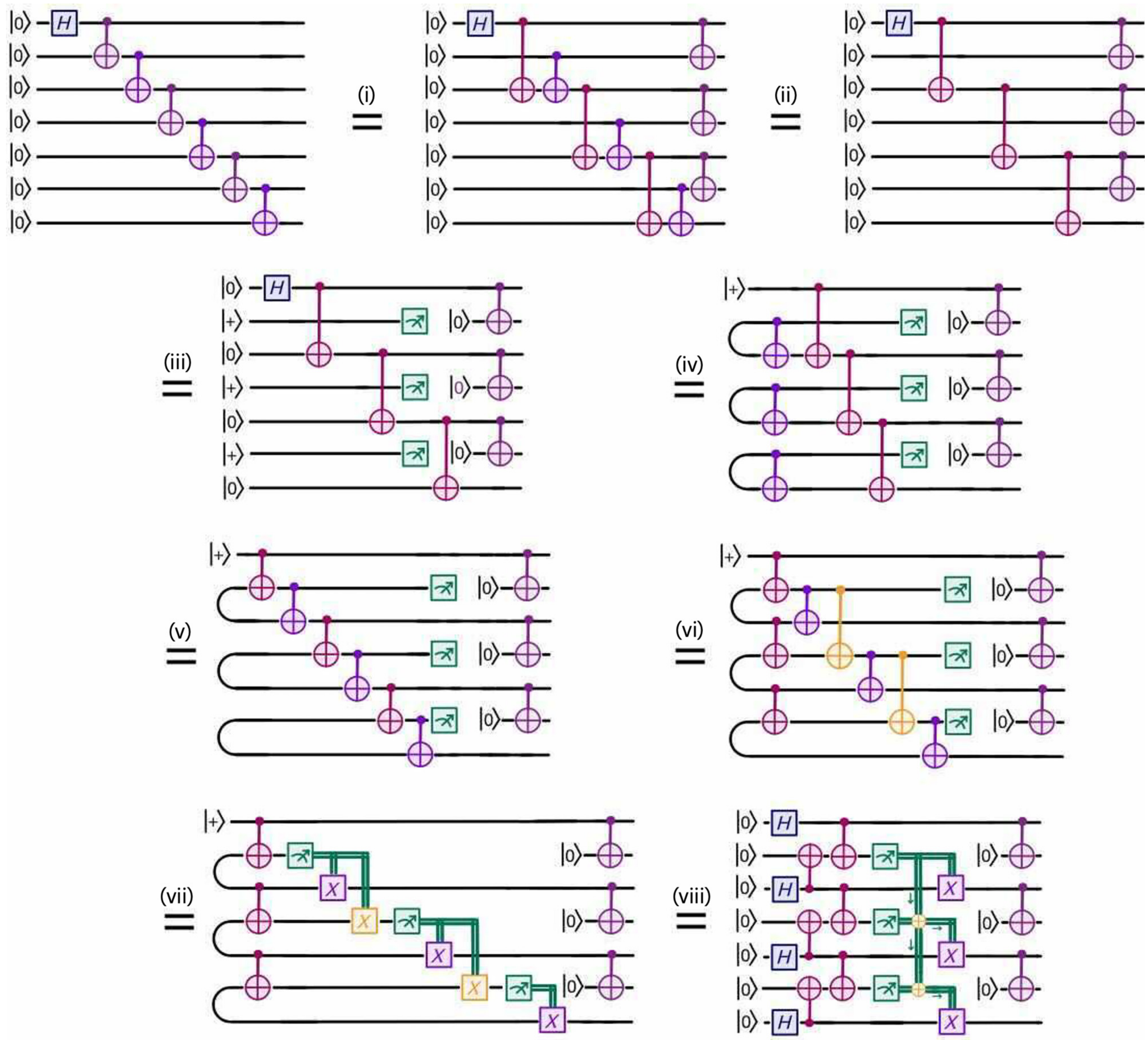


FIG. 5. Graphical derivation for the preparation of a GHZ state by converting its canonical but deep unitary circuit into a constant-depth circuit utilizing measurement and feed-forward operations—a dynamic circuit. Roman numerals indicate sequential step numbers described in the text.

In Table I, we have summarized the error budget for each of the cases.

Comparing the different unitary cases it becomes clear that, for large n, the unitary implementation Ic will be the best, as all other implementations have an error in the idling time that scales quadratically. This might be slightly counterintuitive, as it tells us that the extra $2n$ CNOT gates required for implementation Ic compared to implementation Ia are worthwhile not being canceled, as the full swap leaves the other qubits unentangled, resulting in a drastically decreased idling error, and as even without measurement and feed forward, it can still be beneficial to use ancilla qubits and thereby increase the distances. For small n, we need to keep in mind that, for the SWAP-based implementation (unitary II), the number of involved qubits $\tilde{n}$ is smaller than the number of qubits n needed for the same task in the ancilla-based implementation. Given the qubit division illustrated in Fig. 1(d), we achieve a ratio of 31 qubits connected to the bus, 30 qubits not connected to the bus, and 66 bus qubits, which is a ratio of roughly 1:2 for all-to-all connected qubits to bus qubits. Respecting the rescaled errors, unitary II would be the most promising implementation for small n. In addition to the CNOT errors and idling errors, for dynamic circuits, we also need to consider the error from the additional measurements, as well as a constant term μ that

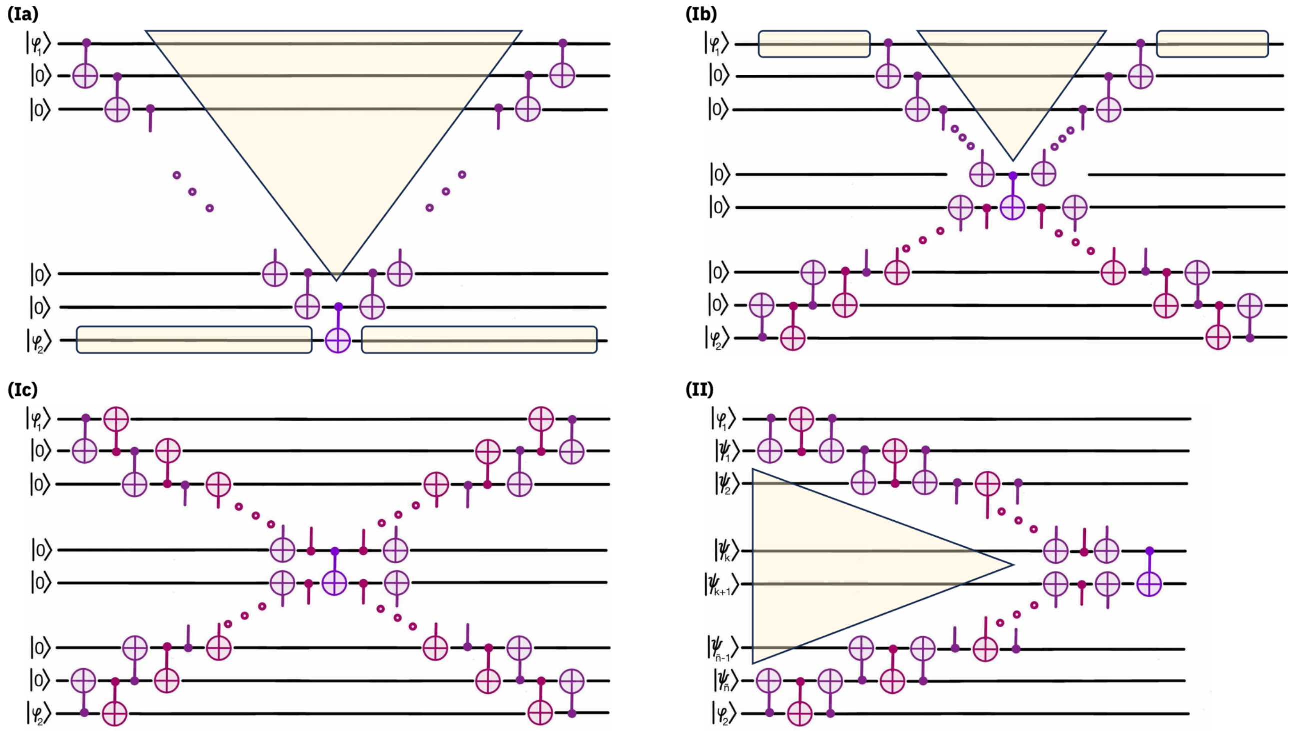


FIG. 6. Comparison of the different unitary implementations of a long-range CNOT gate. While the circuits in panels (Ia), (Ib), and (Ic) realize ancilla-based implementations, the circuit of panel (II) realizes a SWAP-based implementation without ancillas. The shaded regions indicate idle periods that accumulate errors.

comes from the idling error during measurement and feed forward.

Given this rough error analysis in Table I, we can infer that, for large n, dynamic circuits will be beneficial if the measurement of n qubits introduces less error than $3n$ CNOT gates, that is, when $\lambda_{\text{meas}} < 3\lambda_{\text{CNOT}}$. A sketch of how the fidelities for the different cases decrease with n is illustrated in Fig. 7. Note that these error analyses only take into account the error on the involved qubits though. Also considering the fact that there are potentially a lot of other m qubits "waiting" for this operation to be performed would add another idling error of $m(2n+1)$. So the fact that dynamic circuits can perform entangled gates between arbitrary qubits in constant depth instead of linear depth with only unitary operations speeds up the whole algorithm and therefore might be much more powerful than what we can see in the error on the respective qubits.

APPENDIX C: ESTIMATION OF THE STATE AND GATE FIDELITIES USING MONTE CARLO SAMPLING

1. State fidelity

In order to determine the fidelity of the experimentally prepared quantum state, denoted σ, we employ the *Monte Carlo state certification* method, which was introduced in Refs. [34,35]. We first briefly review the notion of fidelity between two quantum states.

TABLE I. Comparison of the error budgets of the unitary and dynamic circuit implementations in terms of the idle time, number of CNOT gates, midcircuit measurements, and two-qubit gate depth. Note that, as the number of involved qubits $\tilde{n}$ needed for the unitary implementation II is in general much smaller, we rescale it for the error budget with the relation $n \approx 2\tilde{n} + 3$.

Case	t_{idle}	N_{CNOT}	N_{meas}	Two-qubit gate depth
Unitary Ia	$n^2 + 2n$	$2n + 1$	0	$2n + 1$
Unitary Ib	$n^2/4 + n$	$3n + 1$	0	$2n + 1$
Unitary Ic	0	$4n + 1$	0	$2n + 1$
Unitary II	$\frac{3}{4}\tilde{n}^2 - \frac{3}{2}\tilde{n}$	$3\tilde{n} + 1$	0	$\frac{3}{2}\tilde{n} + 1$
Unitary II with normed n	$\approx \frac{3}{16}n^2 - \frac{15}{8}n + \frac{45}{16}$	$\approx \frac{3}{2}n - 2$	0	$\approx \frac{3}{4}n - \frac{5}{4}$
Dynamic circuits	$2\mu + 2$	$n + 1$	n	$2 + \mu$, or $O(1)$

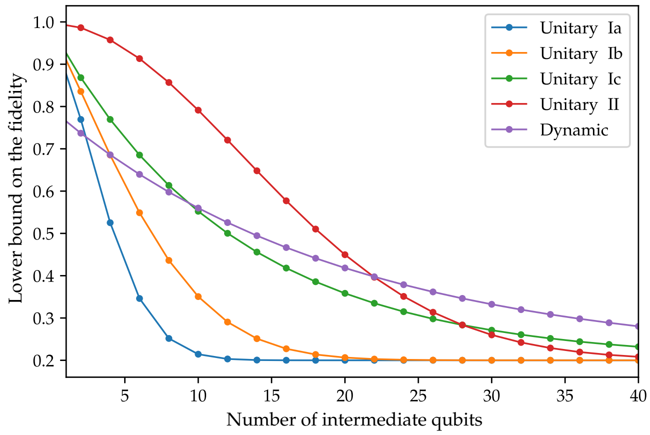


FIG. 7. Comparison of the process fidelities of the different unitary implementations as well as the dynamic circuit implementation considering the error budget indicated in Table I. In this figure we use $\mu = (t_{\text{meas}} + t_{\text{feed forward}})/t_{\text{CNOT}} \approx 3.65$, $\lambda_{\text{idle}} = 0.03$, $\lambda_{\text{CNOT}} = 0.02$, and $\lambda_{\text{meas}} = 0.03$.

a. Quantum state fidelity

Let us introduce the Uhlmann-Jozsa state fidelity between two general quantum states ρ and σ. These objects are elements of the space of valid density operators associated with the system Hilbert space, $\mathcal{H}$, i.e., $\rho, \sigma \in D(\mathcal{H})$. Assuming that one of them is a pure state $\sigma = |\phi\rangle\langle\phi|$, we can simplify the general expression as

$$\begin{aligned} F(\rho,\sigma) &:= \left[\operatorname{Tr}\left(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right)\right]^2 \\ &= \langle\phi|\rho|\phi\rangle \\ &= \operatorname{Tr}[\rho\sigma]. \end{aligned} \tag{C1}$$

If ρ is also a pure state $\rho = |\psi\rangle\langle\psi|$, the expression reduces to a simple overlap $F(\rho,\sigma) = |\langle\psi|\phi\rangle|^2$. We note that some authors define the square root of this as the fidelity.

b. Pauli decomposition

To connect to experimental measurements, let us decompose the quantum sates in the standard Pauli basis. The set of all Pauli operators on n qubits $\{I, X, Y, Z\}^{\otimes n}$ forms an orthogonal Hermitian operator basis. The inner product in operator space $L(\mathcal{H})$ between two Pauli operators $P_i, P_j \in \mathcal{L}(\mathcal{H})$ is $\langle P_i, P_j\rangle = \operatorname{Tr}(P_iP_j) = d\delta_{ij}$, where the dimension of the pure-state Hilbert space $d := \dim\mathcal{H} = 2^n$. In terms of this basis, any quantum state $\rho \in D(\mathcal{H})$, can be decomposed into

$$\rho = \sum_{i=0}^{4^n-1} \frac{\langle P_i, \rho\rangle}{\langle P_i, P_i\rangle} P_i = \frac{1}{d}\sum_{i=0}^{4^n-1} \rho_i P_i, \quad \text{with } \rho_i := \langle P_i, \rho\rangle = \operatorname{Tr}(P_i\rho),$$

where the Pauli expectation value of the state with respect to the ith Pauli operator is ρ_i—an easily measurable quantity. We can similarly define the expectation values of the Pauli operator P_i with respect to the prepared state σ and the desired state ρ as $\sigma_i := \langle P_i\rangle_\sigma = \operatorname{Tr}(\sigma P_i)$ and $\rho_i := \langle P_i\rangle_\rho = \operatorname{Tr}(\rho P_i)$, respectively.

c. Fidelity in terms of Pauli expectation values

The state fidelity between the measured σ and ideally expected pure ρ state [see Eq. (C1)] in terms of the Pauli decomposition of each is

$$F(\rho,\sigma) = \operatorname{Tr}[\rho\sigma] = \sum_i \frac{\rho_i\sigma_i}{d} = \sum_i \frac{\rho_i^2}{d}\frac{\sigma_i}{\rho_i}, \tag{C2}$$

where σ_i is an experimentally measured expectation value and ρ_i is a theoretically calculated one. Given this, we can now define the *relevance distribution* $r(P_i) := \rho_i^2/d$, such that $F(\rho,\sigma) = \sum_{i:\rho_i\neq 0} r(P_i)(\sigma_i/\rho_i)$.

d. Random sampling of expectation values

When sampling m random operators $\{P_k\}_{k=1,\ldots,m}$ according to the relevance distribution $r(P_k)$ and determining their expectation values σ_k, the estimated fidelity $\tilde{F} := \sum_{k=1}^m \sigma_k/\rho_k$ approximates the actual fidelity F with an uncertainty that decreases as $1/\sqrt{m}$. Note that there is also an uncertainty in estimating each σ_k, where, for an additive precision ϵ, roughly $(\epsilon\rho_k)^{-2}$ shots are required.

e. Random sampling of GHZ stabilizers

As the GHZ state is a stabilizer state, for each n, there are exactly 2^n nonzero Pauli operators P_i that each have eigenvalue ± 1. Note that some stabilizers of the GHZ state have a minus sign, e.g., $-YYX$. For the n-qubit GHZ state, by defining the set of stabilizers $\{S_i\}_{i=1,\ldots,2^n}$, we can express the fidelity in terms of only expectation values on the stabilizers

$$F(\rho,\sigma) = \frac{1}{2^n}\sum_{i=1}^{2^n}\langle S_i\rangle_\sigma. \tag{C3}$$

This expression can be approximated by randomly sampling m of the 2^n stabilizers, defining the unbiased estimator $\tilde{F} = (1/m)\sum_{k=1}^m \langle S_k\rangle_\sigma = F + \mathcal{O}(1/\sqrt{m})$, which converges with the number of random samples chosen to the ideal fidelity.

2. Gate and process fidelity

Similarly to the state fidelity, we use the *Monte Carlo process certification* following Ref. [35] to determine the average gate fidelity of our noisy CNOT gate.

a. Average gate fidelity

Consider the case in which we want to implement an ideal gate $\mathcal{U}(\rho) := U\rho U^\dagger$. However, instead, we can implement only a noisy gate $\tilde{\mathcal{U}}(\rho) := \mathcal{U}[\Lambda(\rho)]$, where Λ is some effective noise channel and ρ is a quantum state. What is the gate fidelity of noisy $\tilde{\mathcal{U}}$ relative to the ideal $\mathcal{U}$? For a single given pure state $\rho = |\phi\rangle\langle\phi|$, the state fidelity of the output of the ideal and noisy channels is

$$\begin{aligned} F(\mathcal{U},\tilde{\mathcal{U}};\rho) &= \left\{\mathrm{Tr}\left[\sqrt{\sqrt{\mathcal{U}(\rho)}\tilde{\mathcal{U}}(\rho)\sqrt{\mathcal{U}(\rho)}}\right]\right\}^2 \\ &= \mathrm{Tr}[\mathcal{U}(\rho)\tilde{\mathcal{U}}(\rho)] \\ &= \mathrm{Tr}[\rho\Lambda(\rho)], \end{aligned} \tag{C4}$$

which can be used to obtain the average gate fidelity devised by a uniform Haar average over the fidelity of the ideal and noisy output states, with $\rho_\psi = |\psi\rangle\langle\psi|$,

$$\begin{aligned} \mathcal{F}_{\mathrm{avg}}(\mathcal{U},\tilde{\mathcal{U}}) &= \int d\psi\, F(\mathcal{U},\tilde{\mathcal{U}};\rho_\psi) \\ &= \int d\psi\ \mathrm{Tr}[\mathcal{U}(\rho)\tilde{\mathcal{U}}(\rho)] \\ &= \mathrm{Tr}\left[\int d\psi\ |\psi\rangle\langle\psi|\,\Lambda(|\psi\rangle\langle\psi|)\right]. \end{aligned} \tag{C5}$$

To estimate $\mathcal{F}_{\mathrm{avg}}(\mathcal{U},\tilde{\mathcal{U}})$, we use the process (or entanglement) fidelity as a more experimentally accessible quantity.

b. Process fidelity

Compared to the gate fidelity, the process fidelity is more readily estimated. It can in turn serve as a direct proxy to the gate fidelity. To make the connection, recall that the Choi-Jamiolkowski isomorphism [49] maps every quantum operation Λ on a d-dimensional space to a density operator $\rho_\Lambda = (\mathbb{I}\otimes\Lambda)|\phi\rangle\langle\phi|$, where $|\phi\rangle = (1/\sqrt{d})\sum_{i=1}^{d}|i\rangle\otimes|i\rangle$. For a noise-free, ideal unitary channel $\mathcal{U}$ and its experimental, noisy implementation $\tilde{\mathcal{U}}$, the process fidelity $\mathcal{F}_{\mathrm{proc}}$ is the state fidelity of the respective Choi states $\rho_{\mathcal{U}}$ and $\rho_{\tilde{\mathcal{U}}}$:

$$\mathcal{F}_{\mathrm{proc}}(\mathcal{U},\tilde{\mathcal{U}}) := F(\rho_{\mathcal{U}},\rho_{\tilde{\mathcal{U}}}). \tag{C6}$$

From this fidelity, the gate fidelity can be extracted using the following relation derived in Ref. [50]:

$$\mathcal{F}_{\mathrm{gate}}(\mathcal{U},\tilde{\mathcal{U}}) = \frac{d\mathcal{F}_{\mathrm{proc}}(\rho_{\mathcal{U}},\rho_{\tilde{\mathcal{U}}})+1}{d+1}. \tag{C7}$$

c. Estimating the process fidelity

As described in Ref. [35], instead of a direct implementation of $(\mathbb{I}\otimes\tilde{\mathcal{U}})|\phi\rangle\langle\phi|$ followed by measuring random Pauli operators on all qubits, we follow the more practical approach, where $\tilde{\mathcal{U}}$ is applied to the complex conjugate of a random product of eigenstates of local Pauli operators $P_i\otimes P_j$, followed by a measurement of random Pauli operators $P_k\otimes P_l$. This leads to the same expectation values

$$\begin{aligned} \rho_{ijkl} &:= \mathrm{Tr}[(P_i\otimes P_j\otimes P_k\otimes P_l)(\mathbb{I}\otimes\mathcal{U})|\phi\rangle\langle\phi|] \\ &= \mathrm{Tr}[(P_k\otimes P_l)\mathcal{U}(P_i\otimes P_j)^*]/d. \end{aligned} \tag{C8}$$

The operators are then sampled according to the relevance distribution

$$r_{ijkl} := r(P_iP_jP_kP_l) = \frac{\rho_{ijkl}^2}{\tilde{d}}. \tag{C9}$$

Note that $\tilde{d}$ corresponds to the dimension of the Choi state, i.e., here $\tilde{d} = 16$. For $\Lambda(\rho) = \textsc{cnot}\rho\textsc{cnot}^\dagger$, there are only 16 combinations of Pauli operators with a nonzero expectation value ρ_{ijkl}: $\rho_{ijkl} = -1$ for $P_iP_jP_kP_l \in \{YYXZ, XZYY\}$ and $\rho_{ijkl} = 1$ for the remaining 14. Thus, the relevance distribution is uniform amongst those with $r = \frac{1}{16}$ and we can just take the average expectation value of those 16 operators.

APPENDIX D: EXPERIMENTAL DETAILS

1. CNOT gate teleportation

We perform the long-range gate teleportation experiments on `ibm_sherbrooke`, a 127-qubit superconducting quantum processor. The line of 101 qubits chosen for the experiments is indicated in Fig. 8(a). The cumulative distribution of their T1 and T2 coherence times as well as of their different error rates are shown in Figs. 8(b) and 8(c), with the corresponding median values also indicated. The two-qubit gate time is 0.5 μs, the readout time 1.2 μs, and the feed-forward time roughly 0.7 μs.

2. GHZ state preparation

We perform the GHZ state preparation experiments on `ibm_peekskill`, a 27-qubit superconducting quantum processor. The line of 21 qubits chosen for the experiments is indicated in Fig. 9(a). The cumulative distribution of their T1 and T2 coherence times as well as of their different error rates are shown in Figs. 9(b) and 9(c), with the corresponding median values also indicated. The two-qubit gate time is 0.6 μs, the readout time 0.9 μs, and the feed-forward time roughly 0.7 μs.

APPENDIX E: TOFFOLI OR CCZ gate

Dynamic circuits can also be applied to more efficiently compile multiqubit gates. As an example, we describe how the CCZ or Toffoli gate up to two single-qubit Hadamard gates can be implemented by optimizing multiple teleported CNOT gates. Compilation of the unitary circuit on

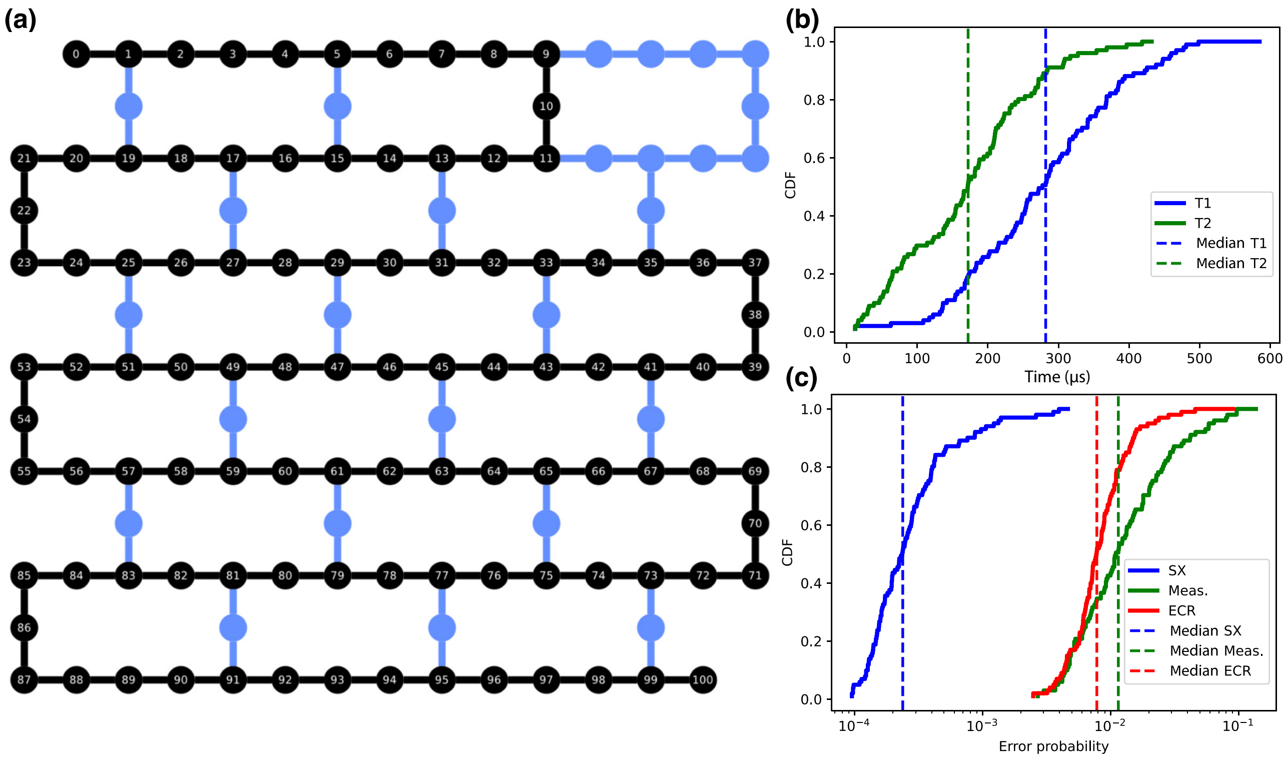


FIG. 8. Implementation details. In (a), we show the device layout of `ibm_sherbrooke`, with the 101 qubits chosen for our dynamic circuits marked in black. In (b) and (c), we plot the cumulative distribution function (CDF) of the T1 and T2 coherence times, the single-qubit gate (SX), readout (Meas.), and two-qubit echoed cross-resonance gate (ECR) error rates of the chosen qubits, as well as the corresponding median values.

a 1D chain of $n+3$ qubits using CNOT gates naïvely requires a two-qubit gate depth of $\mathcal{O}(n)$. Using dynamic circuits, we can implement this long-range entangling gate in shallow depth. Naïvely, one could successively implement each CNOT gate of the typical Toffoli decomposition [shown at the top of Fig. 10(a)] using the gate teleportation described previously. However, involving an ancillary qubit between the three system qubits to merge the teleported gates, as shown at the bottom of Fig. 10(a), allows for a more efficient implementation with the dynamic circuit; see Fig. 10(b). In total, this formulation requires $n+1$ measurements, $n+6$ CNOT gates, and five feed-forward operations divided across two sequential steps. Notably, as most qubits are projectively measured early in the circuit, the idling error should be low. Thus, we expect this shallow implementation with dynamic circuits to be advantageous over its unitary counterpart, especially for large n.

APPENDIX F: ERROR ANALYSIS FOR GHZ STATES

1. Error budget

As in Appendix B 2, we leverage Eq. (B1) to estimate the total noise λ_{tot} of a quantum circuit as motivated by the model discussed in Appendix G. There, we show that $e^{-\lambda_{\text{tot}}}$ gives a lower bound on the *process* fidelity of the circuit. For GHZ states however, we are interested in the *state* fidelity, so the bound from Lemma 1 no longer applies in a rigorous sense. However, we find that the same model can still provide useful intuition if we accept that the model parameters $\lambda_{\text{CNOT}}, \lambda_{\text{meas}}$ no longer have a direct interpretation in terms of worst-case Pauli-Lindblad noise or a combination of amplitude- and phase-damping noise, respectively. See Appendix G for details.

For the unitary approach, we require n CNOT gates to entangle n qubits. For simplicity, we assume (and implement) only a one-dimensional connectivity chain in our protocols and the following numbers correspond to an even number n (only constant terms change when considering odd n). To minimize the idling time, we start in the middle and apply CNOT gates simultaneously towards both ends. This leads to an idle time of $n^2/4 - \frac{3}{2}n + 2$ times the CNOT gate time, as displayed in Table II. In the dynamic circuit approach we require $\frac{3}{2}n - 2$ CNOT gates in total, while the idling time is $\mu n/2 + 1$ times the CNOT gate time, where μ corresponds to the measurement and feed-forward time (as a multiple of the CNOT gate time). However, here we also need to consider the errors of the additional $n/2 - 1$ measurements. As the error coming from the CNOT gates and the measurements is usually substantially larger than

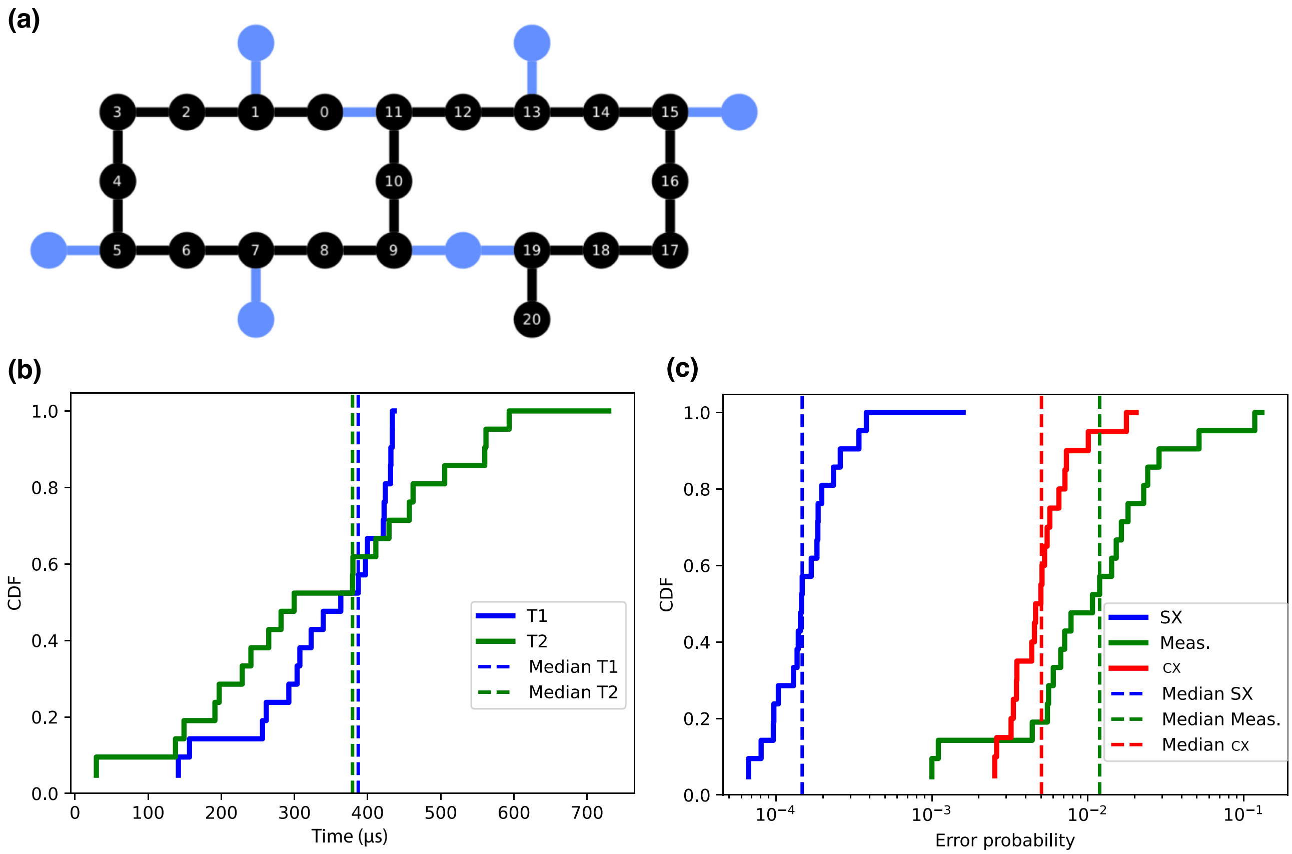


FIG. 9. Implementation details. In (a), we show the device layout of `ibm_peekskill`, with the 21 qubits chosen for our dynamic circuits marked in black. In (b) and (c), we plot the cumulative distribution of the T1 and T2 coherence times, the single-qubit gate (SX), readout (Meas.), and two-qubit controlled-X (CX) error rates of the chosen qubits, as well as the corresponding median values.

the error from the idling time, we expect that, for small n, the standard unitary preparation succeeds. However, as the idling time there scales as $\mathcal{O}(n^2)$ in contrast to all errors in the measurement-based approach scaling only as $\mathcal{O}(n)$, we expect a crossover for large n, where the implementation with dynamic circuits will become more beneficial. The error budget is summarized in Table II.

2. Expected crossover for lower midcircuit measurement errors

In Fig. 11 we determine the expected crossover in performance from unitary to dynamic circuits for varying midcircuit measurement and CNOT gate errors. We use the values of $t_{\text{idle}}, N_{\text{CNOT}}$, and N_{meas} shown in Table II to predict how many qubits are required to see and the state fidelity at the crossover, or where the performance of dynamic circuits becomes higher than that of its unitary counterpart, as a function of the midcircuit measurement errors. Note that in this noise model we assume that we can eliminate all ZZ errors by applying dynamical decoupling. We keep the idling error constant at $\lambda_{\text{idle}} = 0.001$ and consider different CNOT errors $\lambda_{\text{CNOT}} \in \{0.001, 0.01, 0.02\}$. We can reach a fidelity > 0.5 for a CNOT error of $\lambda_{\text{CNOT}} = 0.01$ with midcircuit measurement errors $\lambda_{\text{meas}} \lesssim 0.003$ and for a CNOT error $\lambda_{\text{CNOT}} = 0.001$ with midcircuit measurement errors $\lambda_{\text{meas}} \lesssim 0.012$

APPENDIX G: PAULI-LINDBLAD NOISE MODEL

In this appendix we present a simple framework for computing lower bounds on fidelities using the Pauli-Lindblad noise model discussed in Ref. [46]. Pauli-Lindblad noise channels have several nice properties that we can use to simplify calculations, and also allow us to reduce estimates of the noise properties of our hardware to relatively few parameters.

Normally, Pauli-Lindblad noise is the workhorse of probabilistic error cancelation—an error mitigation scheme that leverages characterization of noise in order to trade systematic uncertainty for statistical uncertainty. But we are more interested in using Pauli-Lindblad noise as a tool for capturing the behavior of fidelity as a function of circuit size with an appropriate balance of rigor and simplicity.

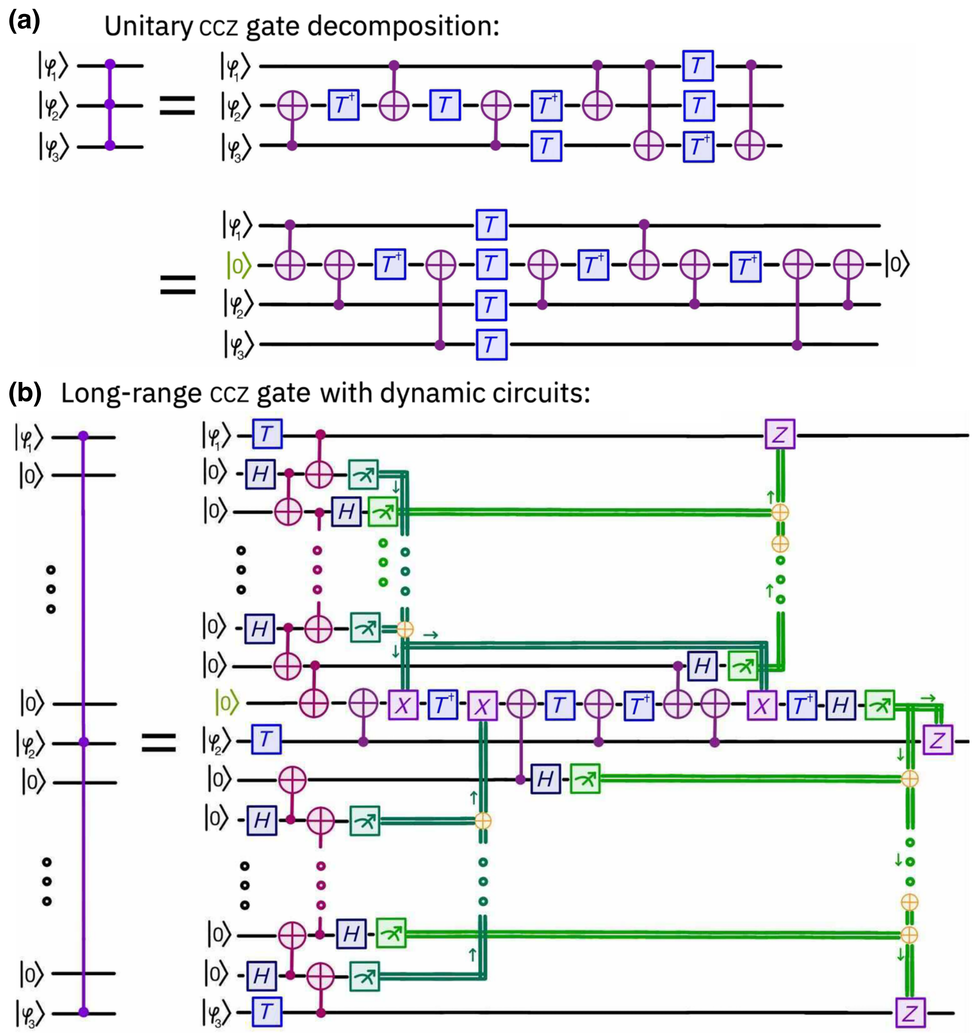


FIG. 10. CCZ gate with a (a) unitary circuit and (b) dynamic circuit over long ranges.

As such, our central goal in this section is to develop mathematical tools that allow us develop a Pauli-Lindblad representation of various noise sources such as decoherence and gate noise and to find a method to combine all of this noise into a fidelity for the entire process. In particular, we aim to give a justification for modeling noise via the quantity λ_{tot} as in Eq. (B1). This is achieved by Lemma 1 below, which states that $e^{-\lambda_{\text{tot}}}$ gives a lower bound on the process fidelity.

We leave the majority of our mathematical exposition without proof for sake of brevity, but present the proof of Lemma 1 at the end of this appendix.

1. Pauli-Lindblad noise

Pauli-Lindblad noise is a quantum channel defined as follows. Let $\mathcal{P}$ be the n-qubit Pauli group modulo phase, and consider some $P \in \mathcal{P}$. Then, for some noise rate $\lambda \in \mathbb{R}^+$, the noise channel Γ_P^λ is given by

$$\Gamma_P^\lambda(\rho) = (1-\omega)\rho + \omega P\rho P^\dagger, \quad \text{where } \omega := \frac{1-e^{-2\lambda}}{2}. \tag{G1}$$

TABLE II. Comparison of the error budgets of the unitary and dynamic circuit implementations in terms of idle time, the number of CNOT gates, midcircuit measurements, and two-qubit gate depth.

Case	t_{idle}	N_{CNOT}	N_{meas}	Two-qubit gate depth
Unitary	$n^2/4 - 3n/2 + 2$	$n-1$	0	$n-1$
Dynamic circuits	$1 + \mu n/2$	$3n/2 - 2$	$n/2 - 1$	$3 + \mu$, or $O(1)$

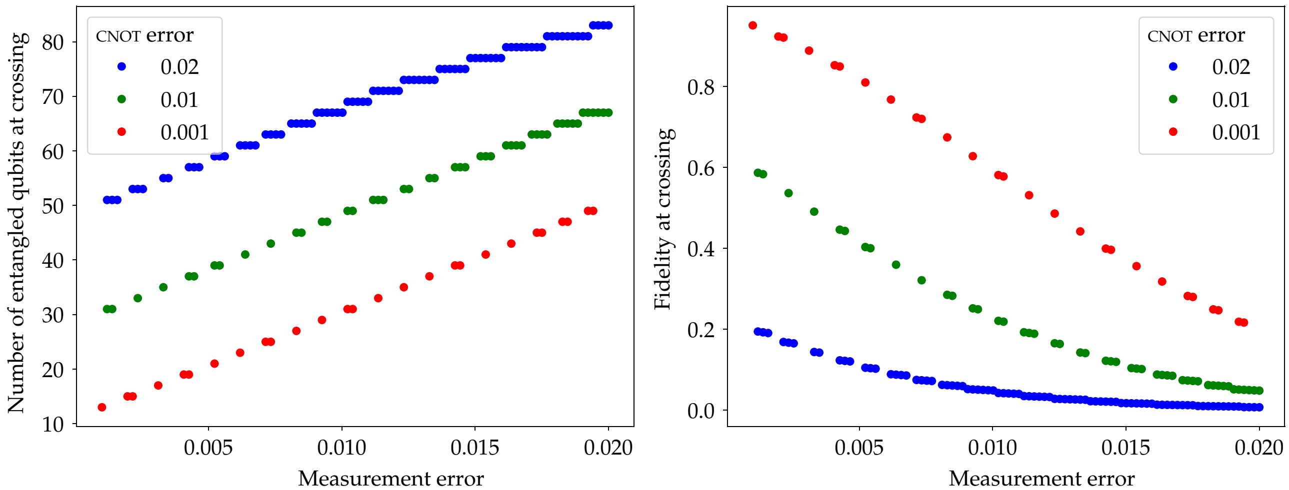


FIG. 11. Noise-model predictions that indicate how many qubits are required to see a crossover and what the corresponding fidelity would be as a function of the midcircuit measurement errors.

This is essentially applying P with probability ω. Pauli noise channels also have a representation as time evolution with respect to a simple Lindbladian: for $P \in \mathcal{P}$, let $\mathcal{L}_P(\rho) := P\rho P - \rho$. This way $\Gamma_P^\lambda = e^{\lambda \mathcal{L}_P}$.

The main justification for why we can restrict to Pauli noise channels is twirling. Conjugating an arbitrary noise channel by a random Pauli matrix yields a channel that is always expressible as a product of Pauli noise. Although our experiments do not feature twirling, even for untwirled circuits, we expect the Pauli-Lindblad noise to capture the first-order noise behavior.

Another reason why we expect our noise model to only capture the behavior to first order is that we assume that the noise rates are the same for all qubits. All CNOT gates and idle times are assumed to contribute the same amount of noise. But this is not a realistic representation of our hardware—in actuality different qubits have different coherence times and gate qualities also vary. When we consider circuits on many qubits, we expect these differences to average out.

Let Λ be a quantum channel. Then let $\tilde{\Lambda}$ be its Pauli-twirled version given by

$$\tilde{\Lambda} := \frac{1}{|\mathcal{P}|} \sum_{P \in \mathcal{P}} P \Lambda(P\rho P) P. \tag{G2}$$

For $Q \in \mathcal{P}$, twirled channels $\tilde{\Lambda}$ satisfy $\tilde{\Lambda}(Q) = c_Q Q$ for some coefficients c_Q. For every $\tilde{\Lambda}$, there exist noise rates λ_P for $P \in \mathcal{P}/\{I\}$ such that $\tilde{\Lambda} = \prod_P \Gamma_P^{\lambda_P}$. These noise rates satisfy

$$c_Q = \exp\left\{-2 \sum_P (\lambda_P \cdot 1_{PQ=-QP})\right\}. \tag{G3}$$

A central convenience of Pauli noise channels is that they do not interfere with each other when propagated: Pauli noise channels commute, $\Gamma_P^{\lambda_P} \Gamma_Q^{\lambda_Q} = \Gamma_Q^{\lambda_Q} \Gamma_P^{\lambda_P}$, and the noise rates can be added together when the Pauli operator is the same, $\Gamma_P^{\lambda_1} \Gamma_P^{\lambda_2} = \Gamma_P^{\lambda_1 + \lambda_2}$.

2. Combining noise channels into a single fidelity

Say we are trying to compute the overall amount of noise in a particular quantum circuit that has been appropriately twirled. Gates and the idle time of the qubits all contribute some amount of Pauli noise. We propagate all of the Pauli noise to the end of the circuit, thereby removing any noise that does not affect certain midcircuit measurements. Finally, we must tally up the noise Pauli operators on the resulting quantum state.

One metric for measuring the error on the final state is the trace distance, or diamond norm if we are considering a channel. For a single Pauli noise source, we have the simple relation that, for any P, we have $|\Gamma_P^\lambda - I|_\diamond = 1 - e^{-2\lambda}$. To generalize this to multiple Pauli operators, a simple approach could be to just apply the triangle inequality to all of the different Pauli operators. But, it turns out we can do much better using the following bound on the process fidelity.

Lemma 1. Consider a channel $\Lambda = \prod_P \Gamma_P^{\lambda_P}$ for some rates λ_P. Then $\mathcal{F}_{\text{proc}}(\Lambda, \mathcal{I}) \geq \exp(-\sum_P \lambda_P)$.

This bound is still pretty loose, but it is very simple and does better than adding up diamond norms. This can be seen by, for example, looking at the channel $\prod_{i=1}^N \Gamma_{P_i}^{c/N}$. Lemma 1 gives $\mathcal{F}_{\text{proc}} \geq \exp(-c)$, while adding up diamond norms and converting them to a fidelity bound gives $\mathcal{F}_{\text{proc}} \geq 1 - \frac{1}{2}N(1 - e^{-2c/N})$. The latter is looser for $N \geq 2$ and for any c.

Lemma 1 also has the key advantage that it makes computation of the overall noise rate very simple: just add up all the noise rates. This allows us to simply tally the total idle time and count the number of CNOT gates to obtain the total amount of noise, as in Appendix B 2.

An issue with using Lemma 1 is that it becomes increasingly loose in the limit of large $\sum_P \lambda_P$. The quantity $\exp(-\sum_P \lambda_P)$ vanishes in this limit, but, in general, we have $\mathcal{F}_{\text{proc}}(\Lambda, \Lambda') \geq 1/d$ for all Λ, Λ'. When we have only one source of Pauli noise Γ_P^λ then not even the lower limit of $1/d$ can be reached as $\lambda \to \infty$. Unfortunately, we see no way of overcoming this limitation while preserving the mathematical elegance of this tool: we would like to simply consider the quantity $\sum_P \lambda_P$. The reason for this shortcoming is that we do not account for cancelations between Pauli errors—we discuss the details of the derivation at the end of this appendix.

Another limitation of this analysis is that it completely ignores crosstalk. Every gate is assumed to behave independently. Assuming independent errors corresponds to a worst-case analysis analogous to the union bound, so we would expect the bounds resulting from Lemma 1 to still roughly capture average error from crosstalk by accounting for it as T_2 dephasing noise, an error that we include when modeling experiments without dynamical decoupling.

3. Propagating noise to the end of the circuit

Next, we discuss how to move all the noise sources to the end of the circuit. This is particularly easy since we are considering Clifford circuits. Once all the noise is in one place, we can use Lemma 1 to combine it into a single fidelity.

With $\mathcal{U}\cdot := U \cdot U^\dagger$ as before, an elementary calculation shows that $\mathcal{U}\Gamma_P^\lambda = \Gamma_{\mathcal{U}(P)}^\lambda \mathcal{U}$, so Pauli-Lindblad noise propagated through a unitary Clifford circuit is still Pauli-Lindblad noise.

Our circuits also feature several adaptive gates, propagation through which can be achieved as follows. Let Λ_{disc} be the channel that traces out the first of two qubits. Then $\Lambda_{\text{disc}}\Gamma_{P\otimes Q}^\lambda = \Gamma_Q^\lambda \Lambda_{\text{disc}}$. Similarly, let $\Lambda_{\text{corr},P}$ be the channel that measures the first qubit and applies a correction P onto the second qubit. If P and Q commute then $\Lambda_{\text{corr},P}\Gamma_{Q\otimes R}^\lambda = \Gamma_R^\lambda \Lambda_{\text{corr},P}$. Otherwise, $\Lambda_{\text{corr},P}\Gamma_{Q\otimes R}^\lambda = \Gamma_{PR}^\lambda \Lambda_{\text{corr},P}$.

Now that we have established how to move noise to the end of the circuit and to tally it into a bound on the fidelity; all that remains is to show how to bring various noise sources into Pauli-Lindblad form.

4. Decoherence noise

We begin with decoherence noise that affects idling qubits. We consider depolarizing, dephasing, and amplitude damping noise.

Conveniently, depolarizing and dephasing noise are already Pauli noise channels. A depolarizing channel $\Lambda_{\text{dep},q}$ replaces the input ρ with the maximally mixed state with probability $1 - q$:

$$\Lambda_{\text{dep},q}(\rho) = q\rho + (1-q)\frac{I}{2^n}. \tag{G4}$$

We find that $\Lambda_{\text{dep},q} = \prod_{P\in\mathcal{P}_n/\{I\}} \Gamma_P^\lambda$ with $q = \exp(-4^n\lambda)$.

The phase damping process is given by the Lindbladian with $L_0 = |0\rangle\langle 0|$ and $L_1 = |1\rangle\langle 1|$:

$$\mathcal{L}_{\text{ph}} = \sum_{i\in\{0,1\}} L_i \rho L_i^\dagger - \frac{1}{2}\{L_i^\dagger L_i, \rho\}. \tag{G5}$$

Since $\mathcal{L}_{\text{ph}} = \frac{1}{2}(Z\rho Z - \rho)$, it satisfies $e^{\lambda\mathcal{L}_{\text{ph}}} = \Gamma_Z^{\lambda/2}$. We can easily compute λ from a phase damping experiment: since $\langle +|\,\Lambda_{\text{damp}}^\lambda(|+\rangle\langle +|)\,|+\rangle = \frac{1}{2}(1+e^{-\lambda})$, we have $\lambda = t/T_2$.

The amplitude damping channel is not a Pauli-Lindblad channel, and must be twirled in order to bring into Pauli-Lindblad form. The amplitude damping process $\mathcal{L}_{\text{damp}}$ is given by $L = |0\rangle\langle 1|$ with

$$\mathcal{L}_{\text{damp}}(\rho) = L\rho L - \tfrac{1}{2}\{L^\dagger L, \rho\}. \tag{G6}$$

If we let $\Lambda_{\text{damp}}^\lambda := e^{\lambda\mathcal{L}_{\text{damp}}}$ then we have $\tilde{\Lambda}_{\text{damp}}^\lambda = \Gamma_X^{\lambda/4}\Gamma_Y^{\lambda/4}$. Similarly, λ can be obtained from an amplitude damping experiment: since $\langle 1|\,\Lambda_{\text{damp}}^\lambda(|1\rangle\langle 1|)\,|1\rangle = e^{-\lambda}$, we straightforwardly have $\lambda = t/T_1$.

If we have both dephasing and amplitude damping noise, we can combine the two together as follows. For some T_1, T_2, consider the combined noise channel $\Lambda_{\text{noise}}^t = \exp(t\mathcal{L}_{\text{damp}}/T_1 + t\mathcal{L}_{\text{ph}}/T_2)$. Then

$$\tilde{\Lambda}_{\text{noise}}^t = \Gamma_X^{t/4T_1}\Gamma_Y^{t/4T_1}\Gamma_Z^{t/2T_2}. \tag{G7}$$

This follows from the fact that $\mathcal{L}_{\text{damp}}$ and $\mathcal{L}_{\text{ph}}$ commute.

5. Noise from unitary gates

In principle, we could perform experiments, as in Ref. [46], to determine the exact Pauli rates for each unitary, as is necessary for probabilistic error cancelation. However, two-qubit gates like the CNOT gate have 15 noise parameters corresponding to the $4^2 - 1$ nontrivial two-qubit Pauli operators. For our purposes, we would prefer to model CNOT noise using just a single number.

One approach could be to just assume that the CNOT noise is simply depolarizing noise. In this case, all 15 Pauli noise rates are equal and can be connected to the process fidelity. Say that we aim to implement an ideal unitary U, but our hardware can only implement $\bar{\mathcal{U}} = \mathcal{U}\Lambda_{\text{dep},q}$ up to a known fidelity $F(\mathcal{U},\bar{\mathcal{U}})$. Then $q = [4^n F(\mathcal{U},\bar{\mathcal{U}}) - 1]/(4^n - 1)$.

However, it turns out that spreading out the error uniformly over all the Pauli operators is rather cumbersome because it requires propagating every possible Pauli error.

A more tractable approach is to just consider the worst-case Pauli error. In that case, For any unitary U and $P \in \mathcal{P}$, we have $F(\mathcal{U}, \mathcal{U}\Gamma_P^{\lambda}) = (1 + e^{-2\lambda})/2$.

6. Conclusions

We have derived a rigorous justification for a rather simple strategy for deriving theoretical predictions of noisy superconducting quantum hardware. Expressions for noise as a function of circuit size can be derived simply by counting the amount of idle time, CNOT gates, and the number of midcircuit measurements. The model has very few parameters, which are simply the Pauli-Lindblad noise rates corresponding to each of these operations (sometimes per unit time). These different noise rates are added up and converted to a fidelity via Lemma 1.

The advantage of a rigorous derivation is that we can directly see the ways in which this model fails to tightly capture the actual error. A central issue is that Lemma 1 does not take into account cancelation between various noise sources, causing the fidelity to approach zero in the limit of a high rate. This is despite the fact that the worst possible process fidelity is nonzero. Another oversimplification is that we do not capture the fact that not all possible Pauli noise rates can affect a given observable. We also cannot capture correlations between errors, as may be the case with crosstalk, and instead take a worst-case approach reminiscent of the union bound. All of these reasons indicate that this model should produce relatively loose lower bounds.

Proof of Lemma 1. Say that $\Lambda(\rho) = \sum_{P,Q} c_{P,Q} P\rho Q$. Then $\mathcal{F}_{\text{proc}}(I, \Lambda) = \mathcal{F}_{\text{proc}}(I, \tilde{\Lambda}) = c_{I,I}$.

The proof proceeds with two loose lower bounds that notably fail to capture cancelations between different error sources. Given $\Lambda = \prod_P \Gamma_P^{\lambda_P}$, recall that $\Gamma_P^{\lambda_P}(\rho) = (1 - \omega_P)\rho + \omega_P P\rho P^{\dagger}$. Expanding out Λ, we see that

$$c_{I,I} \geq \prod_P (1 - \omega_P) = \prod_P \frac{1 + e^{-2\lambda_P}}{2}. \tag{G8}$$

Next, observing that $(1 + e^{-2x})/2 \geq e^{-x}$ for $x > 0$,

$$\cdots \geq \prod_P e^{-\lambda_P} = \exp\left(-\sum_P \lambda_P\right). \tag{G9}$$

This completes the proof. ■

7. Convergence to 0.4

In the main text, we remarked that the fidelities of the measurement-based CNOT experiments converge to a value slightly below 0.4, as is observed in Fig. 1(c). As discussed, this is due to the structure of the measurement-based circuit in Fig. 1(a). While the circuit also experiences infidelity on the top and bottom qubits due to the idle time and some CNOT gates, the only infidelity that actually scales with n is due to incorrect Z and X corrections on the top and bottom qubits, respectively.

We can model this noise as $\Gamma_{ZI}^{\lambda_{ZI}}\Gamma_{IX}^{\lambda_{IX}}$ in the limit of large $\lambda_{ZI}, \lambda_{IX}$, in which case ω_{ZI}, ω_{IX} approach $1/2$. We proceed as in Eq. (G8). Since these Pauli errors cannot cancel, the calculation is exact:

$$\mathcal{F}_{\text{proc}}(I, \Gamma_{ZI}^{\lambda_{ZI}}\Gamma_{IX}^{\lambda_{IX}}) = c_{I,I} = (1 - \omega_{ZI})(1 - \omega_{IX}) = 1/4. \tag{G10}$$

This converts to $\mathcal{F}_{\text{gate}}(I, \Gamma_{ZI}^{\lambda_{ZI}}\Gamma_{IX}^{\lambda_{IX}}) = [4\mathcal{F}_{\text{proc}}(I, \Gamma_{ZI}^{\lambda_{ZI}}\Gamma_{IX}^{\lambda_{IX}}) + 1]/(4 + 1) = 0.4$.

[1] J. Preskill, Quantum computing in the NISQ era and beyond, Quantum **2**, 79 (2018).

[2] H. Buhrman, M. Folkertsma, B. Loff, and N. M. P. Neumann, State preparation by shallow circuits using feed forward, ArXiv:2307.14840.

[3] B. M. Terhal, Quantum error correction for quantum memories, Rev. Mod. Phys. **87**, 307 (2015).

[4] P. Høyer and R. Špalek, in *STACS 2003*, edited by H. Alt and M. Habib (Springer Berlin Heidelberg, Berlin, Heidelberg, 2003), p. 234.

[5] P. Pham and K. M. Svore, A 2D nearest-neighbor quantum architecture for factoring in polylogarithmic depth, Quantum Inf. Comput. **13**, 937 (2013).

[6] D. M. Greenberger, M. A. Horne, and A. Zeilinger, in *Bell's Theorem, Quantum Theory and Conceptions of the Universe*, edited by M. Kafatos (Springer Netherlands, Dordrecht, 1989), p. 69.

[7] R. Raussendorf, S. Bravyi, and J. Harrington, Long-range quantum entanglement in noisy cluster states, Phys. Rev. A **71**, 062313 (2005).

[8] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, Topological quantum memory, J. Math. Phys. **43**, 4452 (2002).

[9] L. Piroli, G. Styliaris, and J. I. Cirac, Quantum circuits assisted by local operations and classical communication: Transformations and phases of matter, Phys. Rev. Lett. **127**, 220503 (2021).

[10] T.-C. Lu, L. A. Lessa, I. H. Kim, and T. H. Hsieh, Measurement as a shortcut to long-range entangled quantum matter, PRX Quantum **3**, 040337 (2022).

[11] R. Verresen, N. Tantivasadakarn, and A. Vishwanath, Efficiently preparing Schrödinger's cat, fractons and non-Abelian topological order in quantum devices, ArXiv:2112.03061.

[12] N. Tantivasadakarn, R. Thorngren, A. Vishwanath, and R. Verresen, Long-range entanglement from measuring symmetry-protected topological phases, ArXiv:2112.01519.

[13] N. Tantivasadakarn, R. Verresen, and A. Vishwanath, The shortest route to non-Abelian topological order on a quantum processor, ArXiv:2209.03964.

[14] N. Tantivasadakarn, A. Vishwanath, and R. Verresen, Hierarchy of topological order from finite-depth unitaries, measurement, and feedforward, PRX Quantum **4**, 020339 (2023).

[15] S. Bravyi, I. Kim, A. Kliesch, and R. Koenig, Adaptive constant-depth circuits for manipulating non-Abelian anyons, ArXiv:2205.01933.

[16] T.-C. Lu, Z. Zhang, S. Vijay, and T. H. Hsieh, Mixed-state long-range order and criticality from measurement and feedback, PRX Quantum **4**, 030318 (2023).

[17] E. H. Lieb and D. W. Robinson, The finite group velocity of quantum spin systems, Commun. Math. Phys. **28**, 251 (1972).

[18] S. Bravyi, M. B. Hastings, and F. Verstraete, Lieb-Robinson bounds and the generation of correlations and topological quantum order, Phys. Rev. Lett. **97**, 050401 (2006).

[19] R. Raussendorf, D. E. Browne, and H. J. Briegel, Measurement-based quantum computation on cluster states, Phys. Rev. A **68**, 022312 (2003).

[20] R. Jozsa, An introduction to measurement based quantum computation, ArXiv:0508124.

[21] M. Beverland, V. Kliuchnikov, and E. Schoute, Surface code compilation via edge-disjoint paths, PRX Quantum **3**, 020342 (2022).

[22] M. D. Barrett, J. Chiaverini, T. Schaetz, J. Britton, W. M. Itano, J. D. Jost, E. Knill, C. Langer, D. Leibfried, R. Ozeri, and D. J. Wineland, Deterministic quantum teleportation of atomic qubits, Nature **429**, 737 (2004).

[23] W. Pfaff, T. H. Taminiau, L. Robledo, H. Bernien, M. Markham, D. J. Twitchen, and R. Hanson, Demonstration of entanglement-by-measurement of solid-state qubits, Nat. Phys. **9**, 29 (2012).

[24] D. Ristè, M. Dukalski, C. A. Watson, G. de Lange, M. J. Tiggelman, Y. M. Blanter, K. W. Lehnert, R. N. Schouten, and L. DiCarlo, Deterministic entanglement of superconducting qubits by parity measurement and feedback, Nature **502**, 350 (2013).

[25] Y. Wan, D. Kienzler, S. D. Erickson, K. H. Mayer, T. R. Tan, J. J. Wu, H. M. Vasconcelos, S. Glancy, E. Knill, D. J. Wineland, A. C. Wilson, and D. Leibfried, Quantum gate teleportation between separated qubits in a trapped-ion processor, Science **364**, 875 (2019).

[26] A. D. Córcoles, M. Takita, K. Inoue, S. Lekuch, Z. K. Minev, J. M. Chow, and J. M. Gambetta, Exploiting dynamic quantum circuits in a quantum algorithm with superconducting qubits, Phys. Rev. Lett. **127**, 100501 (2021).

[27] M. Foss-Feig, A. Tikku, T.-C. Lu, K. Mayer, M. Iqbal, T. M. Gatterman, J. A. Gerber, K. Gilmore, D. Gresh, A. Hankin, N. Hewitt, C. V. Horst, M. Matheny, T. Mengle, B. Neyenhuis, H. Dreyer, D. Hayes, T. H. Hsieh, and I. H. Kim, Experimental demonstration of the advantage of adaptive quantum circuits, ArXiv:2302.03029.

[28] M. Iqbal, N. Tantivasadakarn, T. M. Gatterman, J. A. Gerber, K. Gilmore, D. Gresh, A. Hankin, N. Hewitt, C. V. Horst, M. Matheny, T. Mengle, B. Neyenhuis, A. Vishwanath, M. Foss-Feig, R. Verresen, and H. Dreyer, Topological order from measurements and feed-forward on a trapped ion quantum computer, ArXiv:2302.01917.

[29] S. Moses, C. Baldwin, M. Allman, R. Ancona, L. Ascarrunz, C. Barnes, J. Bartolotta, B. Bjork, P. Blanchard, M. Bohn, *et al.*, A race track trapped-ion quantum processor, ArXiv:2305.03828.

[30] E. Bäumer, V. Tripathi, A. Seif, D. Lidar, and D. S. Wang, Quantum Fourier transform using dynamic circuits, ArXiv:2403.09514.

[31] S. Cao, *et al.*, Generation of genuine entanglement up to 51 superconducting qubits, Nature **619**, 738 (2023).

[32] K. C. Smith, E. Crane, N. Wiebe, and S. Girvin, Deterministic constant-depth preparation of the AKLT state on a quantum processor using fusion measurements, PRX Quantum **4**, 020315 (2023).

[33] E. H. Chen, G.-Y. Zhu, R. Verresen, A. Seif, E. Bäumer, D. Layden, N. Tantivasadakarn, G. Zhu, S. Sheldon, A. Vishwanath, S. Trebst, and A. Kandala, Realizing the Nishimori transition across the error threshold for constant-depth quantum circuits, ArXiv:2309.02863.

[34] S. T. Flammia and Y.-K. Liu, Direct fidelity estimation from few Pauli measurements, Phys. Rev. Lett. **106**, 230501 (2011).

[35] M. P. da Silva, O. Landon-Cardinal, and D. Poulin, Practical characterization of quantum devices without tomography, Phys. Rev. Lett. **107**, 210404 (2011).

[36] D. Gottesman, The Heisenberg representation of quantum computers, ArXiv:9807006.

[37] P. Hoyer and R. Spalek, Quantum fan-out is powerful, Theory Comput. **1**, 81 (2005).

[38] W. Yang and P. Rall, Harnessing the power of long-range entanglement for Clifford circuit synthesis, ArXiv:2302.06537.

[39] P. D. Nation and M. Treinish, Suppressing quantum circuit errors due to system variability, PRX Quantum **4**, 010327 (2023).

[40] D. Leibfried, E. Knill, S. Seidelin, J. Britton, R. B. Blakestad, J. Chiaverini, D. B. Hume, W. M. Itano, J. D. Jost, C. Langer, R. Ozeri, R. Reichle, and D. J. Wineland, Creation of a six-atom 'Schrödinger cat' state, Nature **438**, 639 (2005).

[41] V. Tripathi, H. Chen, M. Khezri, K.-W. Yip, E. Levenson-Falk, and D. A. Lidar, Suppression of crosstalk in superconducting qubits using dynamical decoupling, Phys. Rev. Appl. **18**, 024068 (2022).

[42] L. Viola, E. Knill, and S. Lloyd, Dynamical decoupling of open quantum systems, Phys. Rev. Lett. **82**, 2417 (1999).

[43] P. Jurcevic, *et al.*, Demonstration of quantum volume 64 on a superconducting quantum computing system, Quantum Sci. Technol. **6**, 025020 (2021).

[44] G. J. Mooney, G. A. L. White, C. D. Hill, and L. C. L. Hollenberg, Generation and verification of 27-qubit Greenberger-Horne-Zeilinger states in a superconducting quantum computer, J. Phys. Commun. **5**, 095004 (2021).

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

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

[47] S. A. Kutin, D. Petrie Moulton, and L. M. Smithline, Computation at a distance, ArXiv:quant-ph/0701194.

[48] P. M. Q. Cruz and B. Murta, Shallow unitary decompositions of quantum Fredkin and Toffoli gates for connectivity-aware equivalent circuit averaging, ArXiv:2305.18128.

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

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