

Leveraging Qubit Loss Detection in Fault-Tolerant Quantum Algorithms

Gefen Baranes^{1,2,*}, Madelyn Cain^{2,*}, J. Pablo Bonilla Ataides^{2,*}, Dolev Bluvstein,² Josiah Sinclair¹,
Vladan Vuletić¹, Hengyun Zhou^{2,3} and Mikhail D. Lukin^{2,‡}

¹*Department of Physics and Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA*

²*Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA*

³*QuEra Computing Inc., Boston, Massachusetts 02135, USA*



(Received 10 April 2025; revised 25 September 2025; accepted 6 November 2025; published 2 January 2026)

Qubit loss errors constitute a dominant source of noise in many quantum hardware systems, particularly in neutral-atom quantum computers. We develop a theoretical framework to effectively detect and correct loss errors in logical algorithms and leverage such loss information in decoding. Considering general quantum error correction codes and logical circuits, we introduce a delayed-erasure decoder for experimentally motivated error models which leverages information from delayed loss detection to accurately correct loss errors, even when the precise moment of the error is unknown. Using this decoder, we identify strategies for detecting and correcting loss errors based on the logical circuit structure. For deep circuits prior to logical measurement, we explore methods to integrate loss detection into syndrome extraction with minimal overhead, identifying optimal strategies depending on the qubit loss fraction in the noise and hardware capabilities. In contrast, we find that many key algorithmic subroutines involve frequent gate teleportation, shortening the circuit depth before logical measurement and naturally replacing qubits with no additional experimental overhead. We simulate this setting using a toy model algorithm for small-angle synthesis and find a significant performance improvement as the loss fraction increases. These results provide a path forward for advancing large-scale fault-tolerant quantum computation in systems with loss error detection.

DOI: [10.1103/ycwc-3myc](https://doi.org/10.1103/ycwc-3myc)

Subject Areas: Atomic and Molecular Physics,
Quantum Physics, Quantum Information

I. INTRODUCTION

Quantum error correction (QEC) is believed to be essential for realizing large-scale quantum computation, as it enables suppression of errors [1–3]. However, its practical implementation remains challenging due to its substantial resource overhead. Recent experimental advancements have demonstrated remarkable progress in implementing quantum operations across multiple logical qubits [4–8] and operating below error thresholds [9,10]. These advances make it clear that practical QEC performance can be substantially improved by tailoring the error correction strategy to the particular experimental error

model [9,11,12], choice of logic gates [4], and the structure of the algorithm itself [13,14].

In particular, qubit loss and leakage from computational subspace are dominant noise sources in many hardware systems, and can have a dramatic effect on the practical performance of quantum processors [15,16]. In the absence of loss correction, all qubits within a QEC code can eventually disappear, destroying the encoded quantum information. Conversely, directly detecting loss provides valuable information about the error location, in contrast with Pauli errors, which are indirectly inferred from syndrome information. Recent work has shown that the rich information provided by so-called erasure detection can in fact improve QEC performance substantially [11,12,17–27]. Similarly, noise bias present in many platforms [4,9,10,15] can also be leveraged to improve performance [28–31]. These studies motivate developing strategies that incorporate loss detection and harness bias to improve practical QEC performance.

In this article, we explore the role of loss errors in error-corrected circuits across a range of quantum hardware platforms, including alkali and alkaline-earth-like neutral atoms, as well as superconducting qubits or trapped ions.

*These authors contributed equally to this work.

†Contact author: gbaranes@mit.edu

‡Contact author: lukin@physics.harvard.edu

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

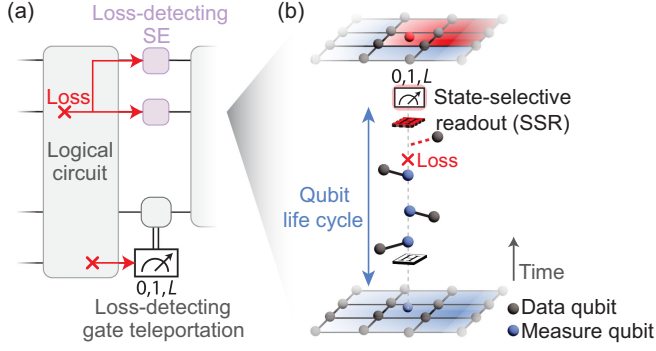


FIG. 1. Loss errors in logical circuits. (a) Depiction of a logical algorithm with loss-detecting SE and gate teleportation. Physical qubit losses (red crosses) can generate correlated errors within and between logical qubits. (b) Space-time diagram of a logical circuit, focusing on a measure qubit life cycle during syndrome extraction. Physical qubits progress through time, undergoing initialization, gate operations, idling, and measurement. A loss event causes future gates to be canceled, generating correlated errors between the qubits in the gate and flipping the corresponding stabilizers.

We focus on the realistic scenario in which detection and correction of loss are delayed for several gate operations, and develop techniques that leverage this delayed information with minimal experimental overhead. In particular, we develop a delayed-erasure decoder to accurately interpret logical measurement results from the measured syndromes and loss detections, despite uncertainty in the exact moment of the loss error.

Using this delayed-erasure decoder, we investigate the impact of loss errors in logical circuits. We demonstrate that the algorithmic structure significantly influences the optimal strategy to detect and correct loss, as summarized in Fig. 1(a). Concretely, for high-depth circuits prior to logical measurement, we develop hardware-efficient methods to detect and replace lost qubits during syndrome extraction (SE) with minimal additional overhead. We perform circuit-level simulations to compare modified circuit-based [3] and measurement-based [32] approaches which incorporate loss detection upon measurement, alongside other state-of-the-art approaches such as erasure conversion techniques [11,18]. We observe that QEC performance can be substantially improved by optimizing circuit design to leverage loss errors in cases where it constitutes a substantial fraction of the error budget. By analyzing biased errors we find that loss has a stronger impact on performance than bias, even with bias-preserving gates. We then study how loss-detecting SE rounds can be optimally interleaved between transversal gates in multiqubit deep logical Clifford circuits.

We next find that many key algorithmic subroutines have short circuit depth before logical measurement, including magic-state distillation [33], quantum arithmetic [34], and small-angle synthesis [35,36]. In such circuits, we find that

because gate teleportation naturally detects and replaces lost qubits, no loss-detecting SE is needed. We discover through numerical simulations of a teleportation-based algorithm that as the frequency of gate teleportation increases, the performance approaches that of frequent erasure conversion. These results provide a comparative analysis of experimental solutions for leveraging loss and highlight that algorithmic structure plays a central role in how these factors impact performance.

Before proceeding, we note that loss-to-erasure conversion has been explored at the memory level, using mid-circuit measurement in alkaline-earth-like atomic systems [11,12,17,18] and superconducting qubits [16,20–22,37–39], as well as using leakage-reduction units with extra qubits and gates [26,40–42]. More recently, loss-resolving readout was demonstrated experimentally in alkali atoms systems and dual-rail superconducting qubits, and its implications for logical qubit performance were explored in Refs. [8,10,23–25,43]. Here we present an end-to-end analysis of loss handling techniques, comparing all these approaches across a range of quantum computing platforms. Our study extends beyond single logical qubits to examine the role of loss at every level of computation, from decoding strategies and QEC architectures to full logical algorithms. We systematically compare different loss detection and correction methods, including their frequencies and associated space-time overheads. By focusing on logical algorithms, we provide new insights into how circuit structure influences the algorithmic performance and develop new, optimized strategies for circumventing loss errors.

Our paper is organized as follows. Section II discusses the effect of qubit loss and its detection using state-selective readout (SSR), and the delayed-erasure decoder developed in this work. In Section III, we tailor and compare different SE techniques designed to handle loss with minimal overhead, which are necessary in circuits with high depth before logical measurements. Section IV explores the optimal frequency of interleaving loss-detecting SE with transversal gates in deep logical Clifford circuits. In Sec. V, we highlight the prevalent use of teleportation in many algorithmic subroutines and evaluate its effect with and without the delayed-erasure decoder on a toy model for the small-angle synthesis algorithm. Finally, we present our conclusions in Sec. VI.

II. DETECTING AND DECODING DELAYED ERASURES

We consider general logical algorithms subject to loss errors, in which loss is periodically detected using hardware-specific mechanisms. In alkali neutral-atom systems or dual-rail superconducting qubits, this can be done via state-selective readout, a projective measurement that distinguishes between $|0\rangle$, $|1\rangle$, and loss. As it occurs as part of the qubit measurement, it can be realized through

various methods with minimal experimental overhead [8,10,12,25,43–48]. In contrast, midcircuit erasure conversion, demonstrated in alkaline-earth-like atoms and superconducting qubits [18,49,50], identifies loss or leakage events during circuit execution by monitoring auxiliary degrees of freedom without performing a full projective measurement. These detection mechanisms, combined with methods like teleportation-based protocols, feed into our syndrome extraction and decoding framework [see Fig. 1(a)].

We explicitly differentiate between two closely related error types. We define a loss error as one in which a qubit leaves the computational subspace and becomes unavailable until detection and replacement, encompassing both loss and leakage (depending on hardware platform). An erasure error, by contrast, is a loss that is detected at the gate layer where it occurs and is immediately reinitialized; both the time and qubit index are known to the decoder.

Loss events lead to correlated errors: In neutral-atom systems, a lost qubit cancels subsequent gates, while its intended interaction partner may experience an effective single-qubit error [15]; in superconducting platforms, a qubit interacting with a lost qubit suffers from depolarizing noise [21]. In both cases, correlated errors can propagate within or across logical qubits during entangling operations [see Figs. 1(a) and 1(b)], and must be handled carefully by the decoder. Here we adopt the gate-cancellation model for loss.

While in principle these errors can be corrected by detecting and replacing the lost qubits and treating them as erasure errors, unlike the ideal erasure channels, where the moment in time of the loss error is precisely known, in a natural QEC cycle, the loss event is detected when the SSR or other loss detection mechanism is performed and can correspond to a number of different potential loss locations [Fig. 2(a)], each resulting in a potentially different set of correlated errors. These loss detection events can be viewed as *delayed* erasure detections, and the information revealed by the detection mechanism can be leveraged by the appropriate decoder to improve logical performance.

As a result, loss events require two adjustments to the decoder. First, when a qubit is lost, its associated stabilizer checks are no longer valid. Therefore, they need to be replaced with a so-called “supercheck,” a product of multiple stabilizer checks into a single check which is independent of the lost qubit [51,52]. For example, consider a lost qubit q_4 that participates in neighboring stabilizers $S_1 = Z_1 Z_2 Z_3 Z_4$ and $S_2 = Z_4 Z_5 Z_6 Z_7$. The resulting supercheck operator is $S_1 S_2 = Z_1 Z_2 Z_3 Z_5 Z_6 Z_7$, which is independent of the lost qubit q_4 . Second, loss causes correlated errors, necessitating updates to the circuit error model to account for the likelihood of error propagation from the possible loss locations (see Appendix A).

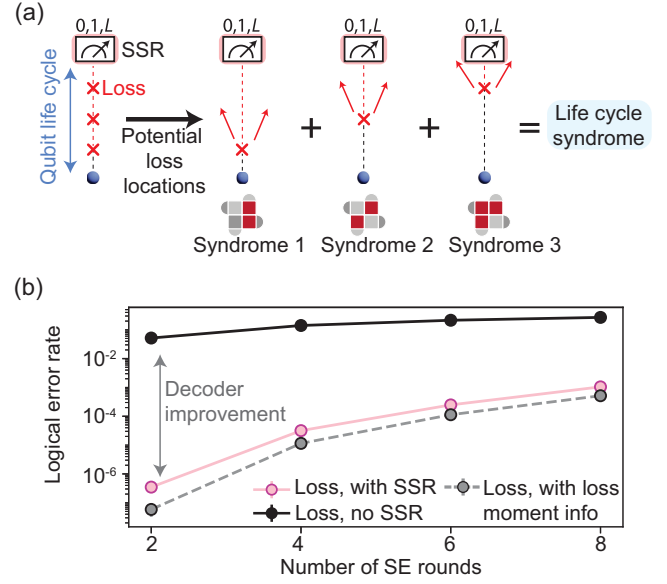


FIG. 2. Delayed-erasure decoder. (a) Illustration of a qubit life cycle and its usage in the delayed-erasure decoder. From initialization to measurement, each physical qubit can be lost at multiple possible time points, each occurring with a potentially different probability and corresponding syndrome. Upon detection of a qubit loss, the decoder accounts for these possibilities in order to improve the accuracy of the assigned correction. (b) Logical error rate for a logical memory as a function of the number of conventional SE rounds before logical measurement, here with distance $d = 5$ and loss errors with probability $p_{\text{loss}} = 1\%$ per entangling gate. The delayed-erasure decoder (pink) substantially outperforms a decoder which does not account for loss information (black). Even without loss moment information, the delayed-erasure decoder achieves comparable performance to a decoder with perfect loss time location (gray, dashed line).

We develop a delayed-erasure decoder which includes these two adjustments to effectively utilize imperfect information about the location of losses in time obtained from SSR. As a key quantity in predicting the performance of different loss detection and correction methods, we consider the qubit life cycle, defined as the number of circuit locations where a given qubit can potentially be lost, starting at initialization and ending at measurement. Formalizing our approach, our goal is to automatically construct a decoding hypergraph based on the observed loss, which captures how errors (hyperedges) trigger checks (vertices that compare consecutive stabilizer checks in time) and is the input to the decoder. Ideally, we would solve the most-likely-error (MLE) decoding problem, which identifies the most likely configuration of Pauli and loss errors consistent with both the observed error syndromes and loss events (see Appendix B 1). However, a full MLE solution would require considering all combinations of loss locations, since loss-induced errors can produce non-Pauli, correlated error patterns, making them nonadditive, as observed numerically.

Instead, we approximate the MLE solution by handling each loss event independently (see Appendix B 2 for full details). For each loss event, we trace back the qubit's life cycle, accounting for all potential loss events and their associated probabilities. Each potential loss event corresponds to a loss circuit, in which certain gates are canceled due to the loss. Using STIM [53], a Clifford circuit simulator, we construct the hyperedges and their probabilities for each loss circuit. Next, we integrate all hyperedges within a qubit's life cycle, reweighting them based on the probabilities of their corresponding potential loss locations. In cases of multiple losses, each loss event is calculated independently, and the results are combined to construct the final decoding hypergraph used by the decoder, calculated by

$$\sum_i D_i + D_{\text{Pauli}} + \omega D_{\text{first comb}}, \quad (1)$$

where D_i is the decoding hypergraph for a lossy life cycle i , D_{Pauli} is the decoding hypergraph for Pauli errors generated from the lossless circuit, and $D_{\text{first comb}}$ is the decoding hypergraph built from the earliest potential loss in each lossy life cycle i . The weighting parameter ω tunes the contribution of $D_{\text{first comb}}$, and is set to $\omega = 0$ in the main-text simulations (see Appendix B 2 for details). Empirically, this heuristic achieves performance comparable to methods that consider combinations of loss events, while significantly reducing computational overhead.

The updated decoding hypergraph can then be processed using a standard decoder appropriate for the QEC code in use. For example, one may use an MLE decoder, minimum-weight perfect matching (MWPM) for topological codes [9,49,54], belief propagation with ordered statistics decoding (BP + OSD) for sparse codes [55], or machine-learning-based decoders [9,56]. Unlike previous approaches to decoding loss errors [9,22,23], our decoder automatically adjusts the error model based on any general circuit and loss information, eliminating the need for hand-tuned models that may not easily generalize to complex logical algorithms or different SE methods. Here we primarily use the MLE decoder from Ref. [13], as it can decode generic stabilizer codes and logical algorithms. We also present results for the MWPM decoder in Appendix C.

To benchmark the decoder, we use a surface-code logical memory with repeated SE rounds in which loss is detected only by SSR: Losses on measure qubits are detected each round, whereas losses on data qubits are detected only at the final projective readout. Unless otherwise noted, this paper considers loss and Pauli errors occurring during entangling operations, as this represents a dominant source of error in many quantum computing architectures [4,9,15]. Details of the error model, along with results using alternative error models, can be found in Appendix F and Table I. As shown in Fig. 2(b), a delayed-erasure

decoder that uses only SSR information lowers the logical error rate by orders of magnitude relative to an MLE decoder that ignores loss information and remains comparable to a decoder with perfect loss time location, even for deeper circuits with longer life cycles.

III. TECHNIQUES FOR ADDRESSING QUBIT LOSS IN DEEP CIRCUITS

In deep logical circuits prior to logical measurement, qubit life cycles are extended, introducing many correlated errors and degrading performance. While the delayed-erasure decoder mitigates these effects, its effectiveness deteriorates when life cycles become too long due to error accumulation. To maintain high performance, loss detection and qubit replacement must be integrated into the QEC process. In this section, we explore practical syndrome extraction methods that enable such replenishment with minimal overhead. We primarily focus on (i) conventional SE, which resembles traditional circuit-based quantum computing with various modifications for delayed-erasure conversion utilizing SSR or erasure conversion capabilities, and (ii) teleportation-based SE, which resembles measurement-based quantum computing (MBQC) [57,58]. Additionally, a description of a modified Steane SE approach, which bridges conventional and teleportation-based SE but is not simulated here, is provided in Appendix D 5. These methods differ in their qubit overhead, gate operations, loss detection capabilities, and experimental requirements, as analyzed below.

We find that the underlying circuits and resulting performance of the methods are similar, but not identical: All approaches successfully remove qubit loss and have trade-offs in logical error rates and qubit overheads depending on the specific ratios between loss and Pauli errors. We begin by detailing each SE approach before analyzing numerical results in Sec. III D. In Sec. III E, we relate these performance differences to key metrics such as life cycle length and number of entangling gates to predict performance based on the noise model for each SE method.

A. Modified conventional SE

The conventional SE method involves repeated stabilizer checks using physical measure qubits. Data qubits are not directly measured or replaced when lost, and thus over time, logical performance can degrade in the presence of loss. We consider augmenting conventional SE by utilizing SSR and physical SWAP gates to detect losses on all qubits (SWAP SE), as proposed in Refs. [42,59] and further explored in Refs. [24,25]. At the end of each SE round, a SWAP gate and physical SWAP movement are performed between data and measure qubits. Conveniently, this approach does not require any additional entangling gates by applying gate cancellation identities [see Fig. 3(a)]. If the data qubit is not lost, the SWAP operations cancel each

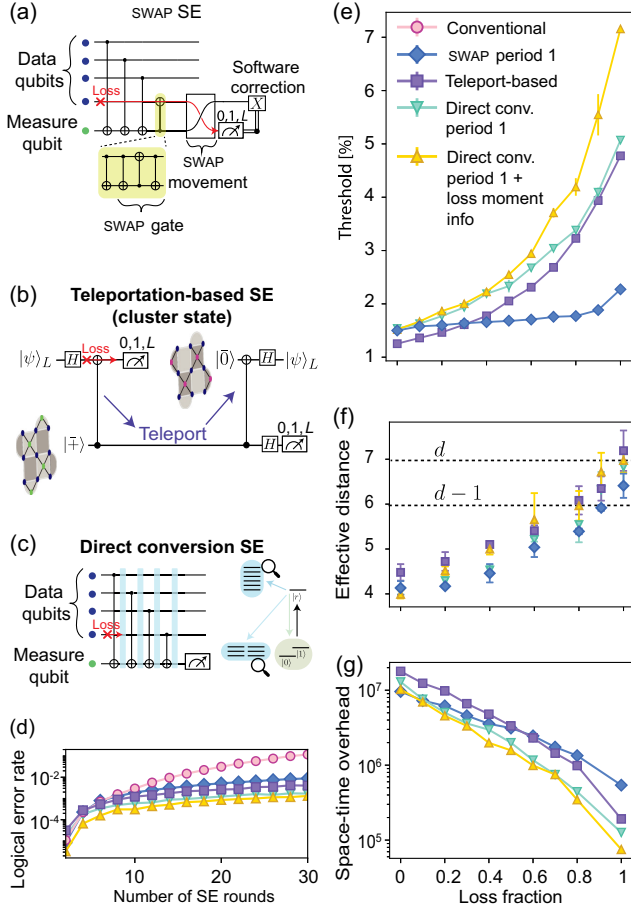


FIG. 3. Loss-detecting SE methods. (a) Modified SWAP SE: Data and measure qubits are swapped each round to detect loss via SSR; the SWAP's CNOT decomposition cancels with existing parity-extraction gates, with the remaining CNOT replaced by classical feedforward, so no additional entangling gates are required. (b) Teleportation-based SE: Logical qubits are teleported to fresh blocks prepared in alternating bases, detecting loss via SSR and shortening life cycles, at the cost of extra qubits. (c) Direct conversion SE: Midcircuit measurement and replacement convert loss to erasure using additional hardware [18]; detection or replacement frequency is varied in simulations. (d) Logical error rates vs SE rounds ($d = 7$, $p = 1\%$) using the delayed-erasure decoder. Comparable performance is observed across SE methods in regimes with short life cycles. (e) Error thresholds vs loss fraction, showcasing the improvement with loss for all SE methods. (f) Effective distance vs loss fraction ($d = 7$), improving with loss for all SE methods. (g) Space-time overhead to reach $P_L = 10^{-12}$ at $p = 0.5\%$ vs loss fraction. Legends for (d)–(g) appear in (e).

other, such that the resulting measure qubit stores the stabilizer outcome. Conversely, if the data qubit was lost, the loss is directly identified through the SSR measurement and automatically replaced. Thus, using this method, we correct losses of all types of qubits, each cycles through both roles of data and measure qubits, ensuring a uniform life cycle length of ~ 8 for all qubits in the bulk even in deep

circuits (see Appendix D 1). However, a data qubit loss detection comes at the cost of losing a stabilizer outcome information, potentially impacting performance, as observed numerically in the next section.

Each SWAP operation incurs a cost due to qubit movement, making it crucial to evaluate different SWAP periods to balance performance and experimental complexity. The SWAP period defines how often loss-detecting SE rounds are interspersed with conventional SE rounds, occurring at a fixed interval. We identify a trade-off between the life cycle lengths of data and measurement qubits for different SWAP periods, demonstrating that a period of 2 can be competitive with a period of 1 (see Figs. 14 and 15). A period of 2 reduces experimental complexity while maintaining a similar average life cycle length.

B. Teleportation-based SE

Teleportation-based SE uses repeated logical teleportation to fresh code blocks, inherently detecting loss via SSR at each step [17,23,57,58] [see Fig. 3(b)]. In our implementation, logical qubits are prepared in one initialization round and then entangled in the XZZX pattern, which is equivalent to the XZZX cluster state used in MBQC [31] (see Appendix D 4 for more details). While typically associated with MBQC, it can also be implemented in a circuit-based framework where gates are directly applied to qubits. Under our error model with noise only on entangling gates, this construction yields short life cycles of length four for all physical qubits in the bulk. Among the SSR-based approaches considered, it achieves the shortest life cycles without experimental operations beyond SSR, though at the cost of additional qubits. These features make it well suited for neutral-atom architectures and particularly effective in loss-dominant regimes, as discussed in Sec. III D.

C. Midcircuit erasure conversion SE approaches

Midcircuit erasure conversion offers an alternative to SSR-based detection by converting physical qubit loss into flagged erasure events during circuit execution [see Fig. 3(c)]. This has been demonstrated in alkaline-earth-like atoms [11,12,17,18] and superconducting qubits [16,20–22,37–39,50,60], using auxiliary fluorescence imaging, auxiliary qubits, or leakage monitors, followed by reset or transport. The resulting erasure is often delayed and requires active intervention, such as qubit replacement or reset, imposing experimental overhead in real-time implementations.

We refer to this as the direct conversion SE method, and explore multiple detection-replacement schedules: (i) detection and replacement after each gate layer or SE round (periods 0.25,1) and (ii) detection after each gate but replacement only at the end of the round (period 1 + loss moment information). Frequent detection improves decoder performance; frequent replacement reduces logical

error by halting loss propagation, but might extend circuit duration due to feedforward control. This flexibility is a practical advantage of erasure conversion.

D. Comparison of SE methods

We now compare the performance of these approaches in the context of a surface code memory, using circuit-level simulations with varying loss fractions $L = p_{\text{loss}}/(p_{\text{loss}} + p_{\text{Pauli}})$, where p_{loss} and p_{Pauli} are physical loss and Pauli error probabilities, respectively. We identify the optimal SE approach at different loss fractions and Pauli error biases (see Appendix C for detailed results with bias).

Figure 3(d) compares the logical error rates of conventional SE without loss detection, approaches with loss detection—such as SWAP SE, teleportation-based SE, and direct conversion SE—and direct conversion SE with perfect loss information, in a single logical-memory experiment at an experimentally motivated loss fraction of 0.5, comparable to that observed in recent neutral-atom experiments [10,11,15,61]. For a small number of SE rounds (≤ 10), all protocols behave similarly. However, as the number of SE rounds increases, loss detection and the replacement of lost qubits substantially enhance performance, providing a convenient approach to achieving high circuit depths without additional experimental gate overhead.

Figure 3(e) shows the threshold of each SE method as a function of the loss fraction. We determine the threshold by calculating the logical error rate for different code distances d , using a noiseless initialization, followed by $d - 1$ rounds of noisy stabilizer checks and a final noiseless transversal measurement. We find that the thresholds of all methods improve with increasing loss fraction. Notably, methods with shorter life cycles, such as teleportation-based SE, benefit more from increasing loss fraction. In the next section (Sec. III E), we provide further insight into the thresholds of each SE method by linking it to simple characterizations such as qubit life cycles.

Figure 3(f) presents the effective code distance of each SE method as a function of the loss fraction, for distance $d = 7$. The effective distance d_e refers to the number of errors required to cause a logical failure, as the logical error rate scales as $(p/p_{\text{th}})^{d_e}$ far below the threshold p_{th} , where p is the physical error rate. For Pauli noise, $d_e = (d + 1)/2$, and for erasure noise, $d_e = d$. To determine d_e for each loss fraction, we fit the logical error rate data far below the threshold to the function αp^β , where α and β are fitting parameters (see Fig. 24). All SE methods, while utilizing the delayed-erasure decoder, experience increased effective distance with an increased loss fraction. For loss errors only, teleportation-based SE and direct conversion SE methods (periods 0.25 and 1) achieve the optimal effective distance of $d_e \approx d$, while SWAP SE achieves $d_e \approx d - 1$, likely due to longer life cycles. Additional threshold and effective distance results

under an alternative error model from Ref. [23] are presented in Table I.

To account for the qubit overhead required in different SE methods, we evaluate the space-time overhead for performing d SE rounds of each method in Fig. 3(g). We define *space-time overhead* as the product of the total number of physical qubits and the circuit depth required for the computation. This metric provides a platform-independent measure of logical performance. However, it does not include experimental considerations such as real-time qubit transport, additional laser resources, or midcircuit measurement overhead. These factors can significantly affect implementation and are discussed qualitatively in Secs. III A–III C. For a given physical error rate of 0.5%, we determine the required code distance to achieve a logical error rate of 10^{-12} by fitting the logical error as a function of distance for $d = 3, 5, 7, 9$. We then present the space-time overhead of each approach as a function of the loss fraction, highlighting the overhead reduction as the loss fraction increases for all SE approaches. Our results highlight that although teleportation-based SE requires more qubits, its better error suppression when the loss fraction is high leads to a more favorable space-time volume compared to SWAP SE (see space-time overhead for each SE method in Table I).

At the experimentally relevant loss fraction of ~ 0.5 , seen in multiple different hardware systems such as neutral atoms and ions [10,11,15,61], the thresholds of the different methods range from 1.5% to 2.5%, and all SE methods achieve similar effective distances. However, their experimental feasibility differs. Direct conversion SE requires erasure conversion with midcircuit detection and real-time replacement of lost qubits, often demanding extra lasers and fast feedback. By contrast, SSR-based approaches such as SWAP or teleportation SE avoid these requirements. Thus, while all methods benefit from accurate loss detection, the optimal choice depends strongly on hardware-specific capabilities and constraints.

Another important aspect of neutral atom qubits is their intrinsic bias, as Z-type Pauli errors are much more common than X-type Pauli errors [15]. In Appendix C, we investigate the interplay between biased errors and delayed-erasure errors across multiple SE methods, employing the XZZX surface code [29] and the XZZX cluster state [31] for teleportation-based SE. Our analysis considers scenarios both with and without bias-preserving gates, as well as the presence of biased-erasure noise. Bias-preserving gates, such as native CX gates, preserve the bias of the errors, in contrast to CX gates decomposed into CZ and H gates. Biased erasure noise occurs when a qubit exits the computational subspace exclusively from the state $|1\rangle$. As a result, the replaced qubit follows a biased error channel instead of a depolarizing error channel. Additionally, we investigate scenarios where Pauli errors are biased, examining the two-dimensional space defined

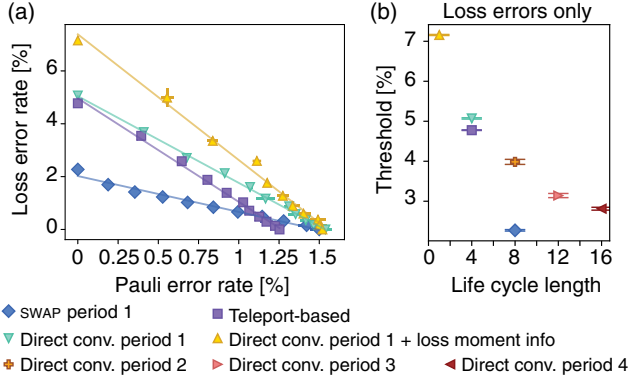


FIG. 4. Linking thresholds to key metrics. (a) Thresholds for different SE methods for different loss and Pauli error rates. The curves are linear fits to numerical finite-size data, with the region below each curve representing the correctable region. (b) Thresholds as a function of life cycle length for various SE methods, in the loss error only limit.

by the bias ratio and the loss fraction. Thresholds are calculated at each point in this space for various SE methods (see Figs. 11 and 12 in Appendix C). A key observation emerges across all SE methods: Increasing the loss fraction has a significantly greater impact on thresholds than increasing the bias ratio, even when using bias-preserving gates and biased erasure. This difference in impact between the loss fraction and the bias ratio becomes even more pronounced in the absence of bias-preserving gates.

E. Predicting performance by error counting

We now present a unified model for analyzing SE methods by linking performance to simple characterizations such as life cycle length and the number of entangling gates. Figure 4(a) shows the thresholds in the parameter space of loss and Pauli error rates for various SE methods. Thresholds are plotted in terms of the loss error rate and Pauli error rate, with the region below each curve representing the range of correctable errors. We numerically determine a good fit to a linear model (solid lines) based on finite-size data, given by $p_{\text{loss}} = p_{\text{loss,th}} - (p_{\text{loss,th}}/p_{\text{Pauli,th}}) \cdot p_{\text{Pauli}}$, where $p_{\text{loss,th}}$ and $p_{\text{Pauli,th}}$ are the respective thresholds for loss and Pauli errors only. The linear behavior suggests that the threshold depends on the loss fraction L according to the relationship $p_{\text{th}} = p_{\text{loss,th}} p_{\text{Pauli,th}} / [L(p_{\text{Pauli,th}} - p_{\text{loss,th}}) + p_{\text{loss,th}}]$, and is qualitatively similar to the behavior observed in Refs. [24,51,52,62].

The curve intersections with the axes provide key insights. The y-axis intersection, representing the threshold in the absence of Pauli errors, is related to the life cycle length. For the direct conversion SE period 1 with perfect loss information, the threshold is $p_{\text{loss,th}} \approx 7.2\%$. Increasing the life cycle reduces the threshold, which decays with the life cycle length, as shown in Fig. 4(b). Notably, the SWAP SE, with an average life cycle length of 8, has a lower

threshold of $p_{\text{loss,th}} \approx 2.3\%$, deviating from the heuristic. This lower performance is attributed to additional factors unique to SWAP SE, such as the loss of stabilizer information when the swapped in data qubit is lost and the fact that measurement qubit loss is detected only in the following round, after it propagates as a data qubit loss, as discussed in Sec. III A.

The x-axis intersection, representing Pauli error thresholds, correlates with the number of entangling gates, which impacts the number of Pauli errors in the final state. For example, over d SE rounds, teleportation-based SE uses 1.5 times more entangling gates than SWAP SE and direct conversion SE approaches, resulting in a proportional decrease in the Pauli error threshold, as similarly observed in Ref. [17].

IV. EFFECT OF LOSS ERRORS IN DEEP LOGICAL ALGORITHMS DESIGN

Loss errors in logical algorithms can have significantly different effects compared to standard memory benchmarks. We now study the effects of loss errors on the QEC design of multiqubit deep logical algorithms. Specifically, we analyze how physical loss errors detection influence the optimal frequency of SE rounds between transversal gates. Our analysis reveals that SWAP-based SE achieves comparable optimal SE frequencies in both the presence and absence of loss.

We focus on the effects of loss in random Clifford logical algorithms [Fig. 5(a)], where multiple logical qubits interact through transversal gates interspersed with periodic SE rounds at tunable frequencies. Circuit-level simulations are performed for logical circuits comprising 24 layers of logical CX and single-qubit logical gates (X , Y , Z), with SE rounds applied at varying intervals, as in Ref. [13] (for more details, see Appendix I). Two scenarios are analyzed: conventional SE without explicit loss detection, which detects only measurement qubit loss in each SE round [Fig. 5(b)], and SWAP-based SE [Fig. 5(c)]. The x axis represents the number of SE rounds per gate layer, ranging from no SE ($n_r = 1/24$) to multiple SE rounds after every gate layer ($n_r = 3$).

In the absence of SWAP operations (but still leveraging SSR detection and delayed-erasure decoding), we observe several distinct error correction regimes. For a small number of SE rounds per gate, loss improves the logical error rate due to short qubit life cycles. However, as the number of SE rounds per gate increases, performance degrades due to the cumulative effects of loss over longer life cycles. By contrast, incorporating SWAP-based SE, loss consistently enhances error correction performance. Loss detection via SWAP stabilizer checks restores the previous optimal SE rounds per CX observed in Ref. [13]. Consequently, when employing SWAP SE and SSR, the presence of loss does not alter the heuristic conclusions

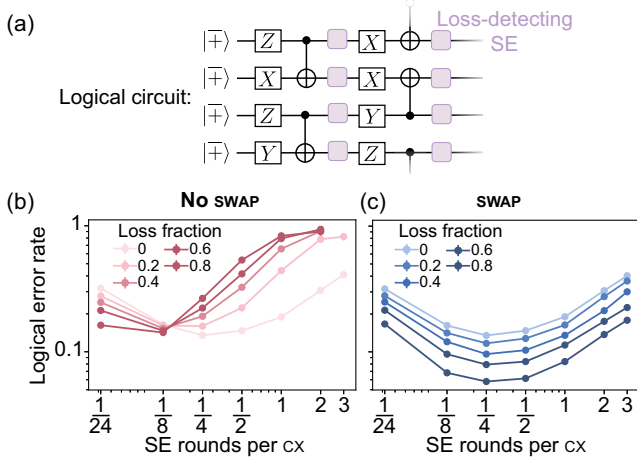


FIG. 5. Deep logical circuits with qubit loss. (a) A deep Clifford logical algorithm consisting of random logical single-qubit and transversal CX gate layers, with periodic SE rounds at varying frequencies. (b),(c) Circuit-level simulation results showing the logical error rate as a function of the number of SE rounds per transversal gate layer, for different loss fractions. ($p = 1\%$, $d = 5$, 24 layers). The SWAP SE method (c) effectively mitigates loss errors, restoring the optimal SE frequency observed in Pauli-dominated scenarios. In contrast, conventional SE (b) exhibits varying error correction regimes, where loss can either improve or degrade performance depending on the life cycle length.

made for logical algorithms dominated by Pauli errors and can, in fact, improve overall performance.

V. NATIVE LOSS DETECTION FROM LOGICAL TELEPORTATION

While we have techniques for handling algorithmic structures with long qubit life cycles, we now observe that in many cases qubit life cycles are short in realistic logical algorithms. In particular, teleportation is a powerful technique for loss detection and correction that avoids additional overhead, and it naturally emerges within logical algorithms through gate teleportation [see Fig. 6(a)]. Specifically, the SWAP-teleported gates exchange the logical data qubit with a teleported logical qubit that implements the desired gate, detecting loss using SSR and terminating the qubit life cycles without the need for additional SE rounds.

This approach is further enhanced by correlated decoding and algorithmic fault tolerance, which capitalize on the use of transversal gates and teleportation in universal quantum computation [13,14]. These ensure that life cycles are inherently short in many logical algorithms, simplifying loss management while enabling regular qubit replacement.

A key observation is that a wide range of known logical subroutines naturally employ teleportation, inherently keeping qubit life cycles short. Figure 6(b) illustrates the average number of SE rounds per physical qubit, from initialization to measurement, across various essential subroutines. The results show that most algorithms have

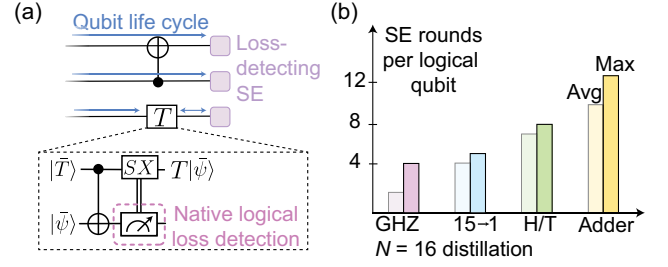


FIG. 6. Native loss detection in logical algorithms. (a) Logical qubit teleportation during gate teleportation, shortening life cycles and naturally detecting loss without loss-detecting SE rounds. (b) SE rounds per logical qubit before logical measurement for various key algorithmic subroutines. Each subroutine uses frequent gate teleportation, keeping life cycles short and detecting loss without additional experimental overhead. GHZ refers to Greenberger-Horne-Zeilinger state.

relatively brief life cycles. Detailed descriptions of these algorithms are provided in Appendix G. For this analysis, we conservatively assume one SE round per logical operation. However, as shown in Fig. 5(c), this number can be further reduced. Additionally, Figs. 3(d) and 22 reveal that, at an experimentally motivated loss fraction of 0.5, for fewer than 10 SE rounds, performance remains nearly unchanged without any active loss correction added to the SE rounds, relying solely on SSR and delayed-erasure decoding. These findings suggest that while memory benchmarks provide useful performance metrics, transversal logical algorithms are inherently well equipped to manage loss, with minimal to no loss detecting SE. Leveraging SSR and delayed-erasure decoding can significantly enhance the performance of the logical algorithm without incurring additional experimental complexity.

As a proof of concept, we consider the teleportation-based logical circuit shown in Fig. 7(a), which mirrors the structure of the small-angle synthesis algorithm [35,36]. This algorithm constructs small-angle rotations through sequences of H and T gates, with the latter often realized via teleportation. Consequently, loss detection and correction are integrated into the process of executing the logical algorithm. The logical error rates for the circuit in Fig. 7(a) are plotted in Fig. 7(b) as a function of the loss fraction using various decoders: a delayed-erasure decoder leveraging SSR, a regular MLE decoder, and an MLE decoder with perfect loss time-location information. The delayed-erasure decoder demonstrates significant performance improvements, with algorithmic logical error rates rapidly decreasing as the loss fraction increases, achieving the performance of the decoder with perfect loss information. Furthermore, for loss fractions less than 1, it achieves the lower bound set by the erasure channel—corresponding to loss detection and qubit replacement after every gate. Notably, this circuit is similar to the teleportation-based SE method investigated here in the memory setting, both utilizing SSR and replacing atoms frequently through

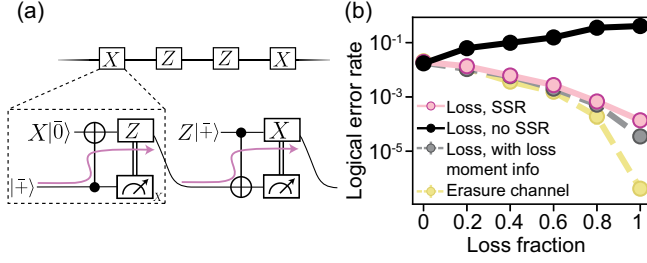


FIG. 7. Deep logical algorithms with teleported gates. (a) We study a deep circuit consisting of teleported X and Z logical gates, similar in structure to small-angle synthesis algorithms [35,36]. (b) Circuit-level simulation results for 11 layers of random teleported Z and X logical gates, with all logical qubits initialized in the presence of noise (physical error rate $p = 1\%$ and $d = 7$). The delayed-erasure decoder (pink) substantially outperforms an MLE decoder that does not account for loss (black) and improves with increasing loss fraction. It also matches the performance of a decoder with perfect loss information (gray, dashed line), and for loss fractions less than 1, approaches that of the erasure channel (khaki, dashed line). The erasure channel performance is achieved using direct conversion SE with period 0.25, corresponding to loss detection and qubit replacement after every gate, thus providing a lower bound.

teleportation without additional overhead. These results highlight the advantages of exploiting native loss detection within teleportation gadgets in logical circuits while also utilizing loss information to improve decoding. Finally, while our simulations focus on Clifford gate teleportation, they are expected to extend to non-Clifford gates, as the underlying loss detection principles remain consistent.

VI. CONCLUSION

These results provide a general framework for managing qubit loss or leakage errors across diverse quantum computing platforms, primarily neutral atoms but also including superconducting qubits and trapped ions, and highlight the impact of loss detection and bias on logical algorithm performance. Our analysis across a wide range of loss fractions and detection strategies reveals how loss detection and decoding can be leveraged in realistic regimes, and was already employed in the recent experimental studies of architectural mechanisms for universal fault-tolerant quantum computing using neutral-atom arrays [10].

Central to our approach is the use of the delayed-erasure decoder, which leverages loss information to approximate an MLE decoding solution, substantially improving the logical error rate. Using our decoder, we examine how the algorithmic structure influences the optimal strategy for detecting and correcting loss. For high-depth circuits involving a large number of gate layers prior to logical measurement, the performance depends on both loss decoding and frequent loss detection and replacement. Using a surface code logical memory, we explore different SE methods under various loss fractions. Our findings

indicate that, with appropriate decoding strategies, all methods enhance performance with increasing loss fraction. Notably, teleportation-based SE is a promising candidate for neutral-atom quantum computing, as it leverages SSR to achieve comparably high thresholds for high loss fractions, albeit with an additional space overhead. Determining the optimal SE method for a given system will depend on experimental validation and the specific noise characteristics of the hardware.

By applying these results to multiqubit deep logical algorithms, several key insights emerge. Loss errors, when managed using loss detection operations and SSR detection, are fully compatible with correlated decoding, allowing for an optimal number of approximately four logical operations per SE round using the delayed-erasure MLE decoder. This framework ensures effective error correction even in the presence of high loss rates.

By considering key subroutines involving extensive logical teleportation, we find that the use of transversal gates and correlated decoding keeps the number of SE rounds sufficiently small, such that loss is natively detected and managed by the gate teleportation intrinsic to universal processing. Examining a toy model of a small-angle synthesis algorithm, we observe that loss errors, decoded using our delayed-erasure decoder, significantly enhance performance compared to Pauli channels solely through logical teleportation. As such, while the logical-memory benchmark provides valuable insights into various strategies for loss management and detection to enhance performance, behavior can radically change when realizing logical algorithms.

This work opens up several new research directions. A major priority is reducing the decoding run-time, critical for scaling to more complex and larger algorithms. While the present work mostly employs an MLE inner decoder, its limited scalability for large-scale logical algorithms underscores the need for further refinement, possibly through alternative strategies, such as matching decoders [9,49,54] or machine-learning-based decoders [9,56].

An intriguing avenue for future exploration involves optimizing loss detection in specific algorithmic subroutines, tailoring the loss-detection frequency and SE method to align with the algorithmic gadget and system characteristics. Additionally, the delayed-erasure decoder can be naturally applied to the decoding of transversal non-Clifford circuits, as it is inherently compatible with and effective for correcting Clifford errors propagating through the circuit.

Finally, although our simulations focus on the surface code, the core techniques presented in this work are broadly applicable to a wide range of QEC codes and logical algorithms. In particular, our delayed-erasure decoder is agnostic to the underlying code and can be applied to any stabilizer code with a suitable decoding back end. Moreover, several loss-handling strategies explored here, such as erasure conversion and teleportation-based

syndrome extraction, are compatible with a wide class of codes, including high-rate qLDPC codes [63–71].

Note added.—Recently, we became aware of several related studies examining the effect of qubit loss on a single logical memory: teleportation SE [23], loss detection units [24], and SWAP SE [72]. While addressing similar challenges, these works differ from ours in focus and methodology. Notably, improved thresholds have been recently reported in updated versions of Refs. [23,72], based on the use of decoding strategies similar to the delayed-erasure decoder described in this work. These contributions also utilize error models slightly different from those presented here. For completeness, our Appendixes include results obtained using the error models utilized in Refs. [23,72].

ACKNOWLEDGMENTS

We thank Sasha Geim, Alex Kubica, Jeff Thompson, Liang Jiang, Isaac Chuang, Nazli Ugur Koyluoglu, Luke Stewart, Marcin Kalinowski, Edita Bytyqi, David C. Spierings, Katie Chang, Casey Duckering, and Shengtao Wang, Chen Zhao for insightful discussions, and Susanne Yelin for her guidance and valuable discussions. G. B. acknowledges support from the MIT Patrons of Physics Fellows Society. D. B. acknowledges support from the NSF GRFP Grant No. DGE1745303 and the Fannie and John Hertz Foundation. We acknowledge financial support from IARPA and the U.S. Army Research Office, under the Entangled Logical Qubits program (Cooperative Agreement No. W911NF-23-2-0219), the DARPA IMPAQT program (Grant No. HR0011-23-3-0012), and the DARPA MeasQuIT program (Grant No. HR0011-24-9-0359), QuEra Computing Inc. (Award No. A57912), the Center for Ultracold Atoms (an NSF Frontier Center), and the National Science Foundation (Award No. PHY-2012023). Support is also acknowledged from the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers, Quantum Systems Accelerator and from the NSF Quantum Leap Challenge Institute (Award No. 2016244).

DATA AVAILABILITY

The data that support the findings of this article are not publicly available upon publication because it is not technically feasible and/or the cost of preparing, depositing, and hosting the data would be prohibitive within the terms of this research project. The data are available from the authors upon reasonable request.

APPENDIX A: SYNDROME OF LOSS ERRORS

This appendix explains how heralded loss alters the stabilizer measurement structure and modifies the decoding hypergraph in QEC codes. Each loss mechanism induces a

modified circuit, which alters the set of stabilizer detectors and introduces new correlated error configurations. The resulting decoding hypergraph consists of a set of detectors (vertices) and hyperedges, where each hyperedge represents a possible error event, connects a subset of detectors it flips, and is assigned a corresponding probability. This hypergraph serves as the input to the decoder. Given a heralded loss at measurement (e.g., using SSR or erasure conversion), we reconstruct the life cycle of the affected qubit (defined as the sequence of operations from initialization to measurement) and enumerate its potential loss locations. Each potential loss point defines a distinct “loss circuit,” which produces a different set of hyperedges in the decoding hypergraph.

As discussed in the main text, loss requires two decoder-level adjustments: (1) modifying detectors to remove dependence on the lost measurement (via superchecks) and (2) accounting for error edges arising from truncated circuits after loss. Below, we elaborate on each adjustment and provide a specific example for clarity.

First, in each loss circuit, the measurement of the lost qubit is set to a random result. This results in assigning a probability of 0.5 to the relevant edge, which corresponds to a weight of 0 in the decoding hypergraph. This process generates a supercheck operator, defined as the product of neighboring checks, effectively eliminating the lost qubit’s contribution [51]. Since the loss is detected in all loss circuits for the potential loss locations, the supercheck operator appears in all decoding hypergraphs of the qubit’s life cycle. Consequently, the supercheck is incorporated into the final decoding hypergraph of the life cycle.

Second, when an atomic qubit is lost, subsequent gates involving this qubit act trivially. Therefore, each loss circuit omits certain gates due to the lost qubit. As a result, the decoding hypergraph for the circuit includes new edges that correspond to these missing gates, capturing the errors associated with the specific loss under consideration. These edges appear in the final decoding hypergraph of the life cycle, with probabilities reflecting the likelihood of the associated events.

Figure 8 illustrates a specific example using a $d = 3$ surface code over four time steps: an initialization round ($t = 1$), two syndrome extraction rounds ($t = 2, 3$), and a final transversal measurement ($t = 4$). In this example, we present the syndrome patterns caused by losses at different locations and times. The upper panels show the circuits with canceled gates and heralding, while the lower panels display the decoding hypergraph, listing which detectors are activated and their correlations. This example uses a conventional SE scheme with measure qubits read out every round and data qubits only at the end, but our decoding approach applies to any SE method, as discussed in the main text.

Detectors D_0 – D_4 correspond to parity checks between stabilizer outcomes across rounds: D_0 and D_1 compare the

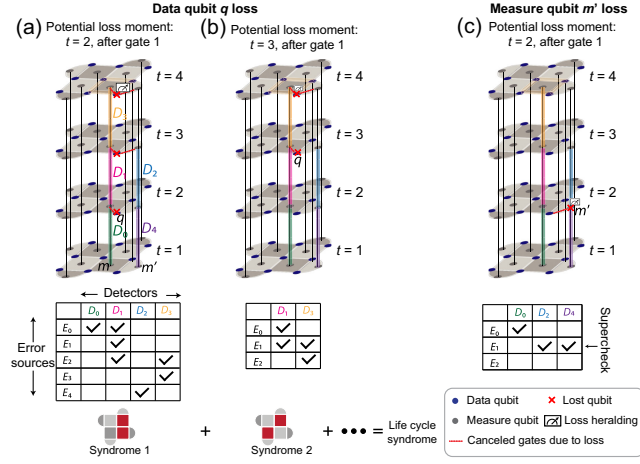


FIG. 8. Example of detector activations generated by different loss events in a $d=3$ surface code over four time steps ($t=1-4$). Panels (a) and (b) show two potential loss locations for the same data qubit q , in round $t=2$ and $t=3$, respectively, each after the first gate. These correspond to different loss circuits within the same life cycle and are combined into a single decoding hypergraph for that life cycle (labeled “life cycle syndrome”). Panel (c) shows a loss of a measure qubit m' in round $t=2$, which belongs to a separate life cycle and contributes a separate decoding hypergraph. Detectors D_0-D_4 are highlighted in color. Each row (E_i) in the decoding hypergraph tables corresponds to one error mechanism (hyperedge), i.e., a correlated set of detector activations triggered by a single loss event. Superchecks that are activated in all loss circuits of a given life cycle also appear in the final decoding hypergraph, such as the shared activation of D_3 in both (a) and (b).

same stabilizer across $t=1 \rightarrow 2$ and $t=2 \rightarrow 3$, respectively, D_2 monitors the neighboring stabilizer of q across $t=2 \rightarrow 3$, D_3 is the final plaquette stabilizer containing q given by the final transversal measurement and compared with previous round, and D_4 compares measure qubit m' across $t=1 \rightarrow 2$.

Each row in the decoding hypergraph (Fig. 8) corresponds to a distinct detector activation pattern produced by propagating a specific loss through the truncated circuit. For example, when data qubit q is lost at $t=2$, all subsequent CZ gates on q are canceled and the final readout is randomized. The missing CZ gates make stabilizers involving q inconsistent across rounds, producing timely detector flips such as $\{D_0, D_1\}$ or $\{D_1\}$. In some branches, the loss heralding itself introduces a plaquette mismatch (D_3), or the neighboring stabilizer of q registers a timely inconsistency (D_2). Thus the multiple rows for a single loss represent different deterministic ways a single heralded loss can violate stabilizer checks, depending on timing and Pauli by-product branches. Now we explain each row explicitly.

(a) Loss of data qubit q at $t=2$. All subsequent gates on q are canceled, and the final measurement is probabilistic. The missing CZ gates disrupt stabilizers involving q . The

decoding hypergraph contains five rows: $\{D_0, D_1\}$ (mismatches in both $t=1 \rightarrow 2$ and $t=2 \rightarrow 3$), $\{D_1\}$ (only $t=2 \rightarrow 3$), $\{D_1, D_3\}$ (round $t=2 \rightarrow 3$ plus the plaquette check containing q), $\{D_2\}$ (neighbor stabilizer at $t=2 \rightarrow 3$), and $\{D_3\}$ (plaquette $t=3 \rightarrow 4$ inconsistency).

(b) Loss of data qubit q at $t=3$. Since stabilizers remain intact through $t=2$, only the final checks are affected: $\{D_1\}$ (round $t=2 \rightarrow 3$), $\{D_1, D_3\}$ (round $t=2 \rightarrow 3$ plus plaquette $t=3 \rightarrow 4$), and $\{D_3\}$ (plaquette $t=3 \rightarrow 4$ inconsistency).

Since both (a) and (b) correspond to different loss locations within the same data qubit life cycle, any detectors that are activated in all loss circuits become part of the supercheck structure for that life cycle. In this case, detector D_3 , which corresponds to the final plaquette stabilizer involving q , is activated in both loss circuits and thus appears in both decoding hypergraphs. As a result, D_3 is retained in the final decoding hypergraph for the life cycle and serves as a supercheck.

(c) Loss of measure qubit m' at $t=2$. The round-2 measurement of m' is missing, invalidating detectors D_2 and D_4 . Their product forms a valid supercheck, $D_2 D_4 = M_{m',t=1} M_{m',t=3}$, correlating rounds 1 and 3 directly. Additionally, the canceled CZ introduces an error on the paired data qubit, activating D_0 .

These examples show how a single loss modifies the circuit and produces specific correlated detector patterns.

APPENDIX B: DELAYED-ERASURE DECODER

This section presents our decoding strategy for handling heralded loss in fault-tolerant circuits. Loss acts in a much more complex fashion compared to normal Pauli errors, and therefore requires more care in terms of their formulation. Here, we show how to formulate the full MLE problem, and then construct a scalable decoder that approximates this solution.

1. Exact MLE decoding in the presence of loss

This section formulates the MLE problem in the presence of heralded qubit loss. While exact MLE decoding is computationally challenging, it serves as a theoretical benchmark that guides the development of efficient approximate methods.

We generalize the framework of Ref. [13] by incorporating both Pauli and loss errors. Let $\vec{E} = (E_1, E_2, \dots)$ represent binary variables for Pauli error events and $\vec{L} = (L_1, L_2, \dots)$ represent binary variables for loss events. We assume Pauli and loss errors occur independently across circuit locations, although correlated models can be incorporated.

Measurement outcomes provide two kinds of information: (1) flags $\vec{F}$, indicating when qubit loss is detected, and (2) detectors $\vec{D}$, constructed as products of stabilizer measurement outcomes that are expected to be $+1$ in the

absence of error. Because loss renders some measurements invalid, the set of active detectors must be dynamically redefined based on the observed flags.

The most likely error problem seeks to find the assignment of Pauli errors $\vec{E}$ and loss errors $\vec{L}$ that maximizes the following quantity:

$$P(\vec{E}, \vec{L} | \vec{F}, \vec{D}). \quad (\text{B1})$$

Namely, given the observed flags, and the resulting detectors that we define conditional on seeing those flags, what is the most likely error?

We can rewrite this expression using Bayes's rule:

$$P(\vec{E}, \vec{L} | \vec{F}, \vec{D}) = \frac{P(\vec{E}, \vec{L}, \vec{F}, \vec{D})}{P(\vec{F}, \vec{D})} = \frac{P(\vec{F}, \vec{D} | \vec{E}, \vec{L}) P(\vec{E}, \vec{L})}{P(\vec{F}, \vec{D})}. \quad (\text{B2})$$

Since we are given a fixed observation $\vec{F}, \vec{D}$, the denominator is fixed. We therefore seek to maximize the product of the probability of the error configuration $P(\vec{E}, \vec{L})$ and the conditional probability of observing these syndromes. In the presence of loss, an important distinction from the usual case is that the latter factor is no longer always 0 or 1.

The error probability $P(\vec{E}, \vec{L})$ can be calculated based on the log-likelihood ratios of individual events. However, since only one loss event can occur per qubit life cycle, the effective probability of loss at location j becomes $p_j \prod_{k < j} (1 - p_k)$. This correction is relevant when the loss probabilities vary significantly across locations.

The more challenging factor to evaluate is $P(\vec{F}, \vec{D} | \vec{E}, \vec{L})$. First, given a loss pattern, detectors must be updated by replacing invalidated measurements with combinations of adjacent measurements (e.g., superchecks). While the exact choice is irrelevant in full MLE decoding due to stabilizer equivalence, it can affect the performance of heuristic decoders. After forming the detectors, we need to evaluate whether a given loss and error pattern is consistent with the observed loss flags and detectors. The flag check is straightforward: At most, one loss event should be active per qubit life cycle. Evaluating the detector consistency is more subtle. In the presence of loss, certain detector activations become probabilistic, depending on both the location and timing of the loss event. Given a specific loss pattern, the circuit becomes fixed, and the resulting conditional detector distribution can in principle be computed by propagating the error configuration through the modified circuit. However, this approach is computationally expensive and difficult to express as a simple optimization problem, motivating approximate methods.

One such approach is edge reweighting, where known correlations between loss and downstream Pauli errors are captured by modifying decoder edge weights. For instance, Ref. [42] models leaked qubits as inducing depolarizing

noise on all subsequent interactions, assigning higher error weights to gates further in time. While this captures leading-order effects, it neglects correlations introduced by the fact that leaked qubits persist across time steps, resulting in temporally correlated error patterns. Accounting for this structure can improve heuristic approximations to MLE decoding.

Ultimately, the goal is to identify practical heuristic methods that approximate MLE decoding while capturing the effects of loss. This can be pursued by analyzing which errors are commonly misclassified by a given weighting scheme or by benchmarking against exact MLE decoding on small instances. Neural network decoders offer a complementary strategy by directly learning mappings from syndrome and flag data to logical corrections, implicitly capturing both error and loss correlations [5]. However, for near-term systems where clarity and integration with standard decoders are important, improving heuristic approximations to MLE may offer greater practical benefits.

2. Approximate MLE decoding

In this section, we describe our practical decoding strategy that approximates the MLE decoding introduced above. The goal is to capture the key features of loss-induced error correlations, particularly the way loss truncates gate sequences and alters detector structure, without solving the full combinatorial problem.

As stated in the main text, our goal is to automatically construct a decoding hypergraph based on the observed loss, which serves as the input to any general decoder for the QEC code analyzed. Our general approach relies on precomputing the decoding hypergraph for each possible loss location within a qubit's life cycle. Given a heralded loss upon measurement, we infer that the loss occurred at some earlier location in the life cycle, and assign a probability to each possible location. Given a specific experiment with a given loss pattern, we can sum over the relevant precomputed life cycle decoding hypergraphs, as described below. Figure 9 illustrates this procedure, showing how the loss locations in each life cycle contribute to the overall decoding hypergraph for a given shot.

We now formalize our approach. Let C_i denote the life cycle of qubit i , and let $\{L_{ij}\}$ denote the set of possible loss locations for that qubit, each associated with a probability p_{ij} . For each possible loss event L_{ij} , we simulate the loss circuit, a modified circuit generated by removing all gates after the loss and setting the qubit's final measurement to a uniformly random outcome. Using STIM's gauge detector infrastructure [53], we extract the decoding hypergraph denoted DEM_{ij} , which consists of a set of hyperedges (error events) and their associated detector configurations.

Each decoding hypergraph DEM_{ij} is a matrix, a collection of error probabilities p_n and corresponding detector configurations D_n :

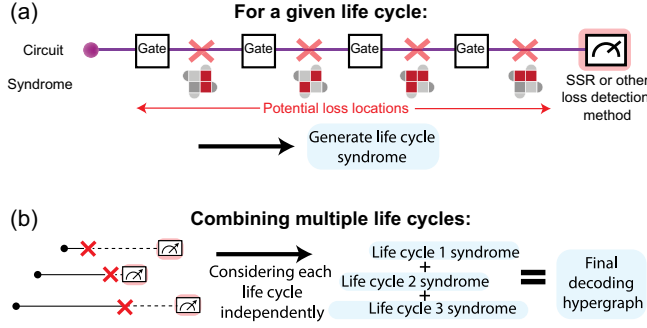


FIG. 9. Overview of the approximate MLE decoding procedure. (a) For a given qubit life cycle, we identify potential loss locations (red crosses) and simulate modified circuits for each, generating a set of decoding hypergraphs (DEM_{ij}). These include updated stabilizer syndromes (detectors) and correlated error edges resulting from the truncated loss circuits. (b) When multiple qubit losses are present in a given shot, we compute the decoding hypergraph for each life cycle independently (left), optionally include a combined loss decoding hypergraph (right), and sum with the standard Pauli decoding hypergraph to generate the final decoding hypergraph ($\text{DEM}_{\text{final}}$), according to Eq. (B5). This forms the input to the decoder.

$$\text{DEM}_{ij} = \{(D_n, p_n)\}, \quad (\text{B3})$$

where the columns $D_n = [D_n(0), D_n(1), \dots]$ represent the list of detectors triggered by the error event in each row. Repeating this process for all potential loss events L_{ij} within a life cycle C_i results in a set of decoding hypergraphs DEM_{ij} and their associated probabilities p_{ij} . This is done in a preprocessing stage.

To compute the final decoding hypergraphs for a life cycle C_i of a lost qubit i , we sum over all decoding hypergraphs in the life cycle:

$$\text{DEM}_i = \sum_j \tilde{p}_{ij} \cdot \text{DEM}_{ij}, \quad (\text{B4})$$

with normalized weights $\tilde{p}_{ij} = \Pr(L_{ij} | \text{flag in } C_i)$; this is an approximation that assumes at most one loss per life cycle and neglects cross-life-cycle combinations.

When multiple qubit losses occur in different life cycles, loss errors can cause correlated Clifford errors. Combining syndromes from distinct life cycles may interfere non-linearly, resulting in syndromes that cannot be derived from independent consideration of loss events. While the optimal decoder would consider all combinations of potential loss locations for all qubits, this approach scales poorly with multiple lossy life cycles and potential loss events, making it impractical for small code distances ($d < 7$).

To address this, we explore a decoder that evaluates each life cycle C_i independently and averages over them (DEM_i). Additionally, we introduce the option to include a single combination of losses across different life cycles. Specifically, we consider the first potential loss locations

L_{i1} for all lossy life cycles L_i and generate a corresponding decoding hypergraph, $\text{DEM}_{\text{first comb}}$.

The decoding hypergraphs for individual potential loss locations L_{ij} are generated in a preprocessing step, independent of any specific error model or probabilities of Pauli and loss errors. During real-time decoding, these hypergraphs are summed using the specific error model probabilities $\tilde{p}_{ij}$. Note that we separately account for loss errors and Pauli errors. Lossless circuits (without losses) are used to efficiently generate the decoding hypergraph for Pauli errors, denoted $\text{DEM}_{\text{Pauli}}$.

For a specific shot with a heralded loss pattern using SSR and a given error model, the final decoding hypergraph is computed by summing (1) lossy life cycle decoding hypergraphs DEM_i , (2) the Pauli decoding hypergraph $\text{DEM}_{\text{Pauli}}$, and (3) the first potential loss combination decoding hypergraph $\text{DEM}_{\text{first comb}}$.

The final decoding hypergraph is given by

$$\text{DEM}_{\text{final}} = \sum_i \text{DEM}_i + \text{DEM}_{\text{Pauli}} + \omega \cdot \text{DEM}_{\text{first comb}}, \quad (\text{B5})$$

where ω is a combination weight determined based on the analysis in Appendix B 2 a.

The summation of probabilities is calculated using the following equation:

$$\sum_{i=1}^l p_i \prod_{j \neq i} (1 - p_j) + \sum_{i=1}^l \sum_{j=i+1}^l \sum_{k=j+1}^l p_i p_j p_k \prod_{m \neq i,j,k} (1 - p_m). \quad (\text{B6})$$

This ensures accuracy up to $O(p^3)$.

Once $\text{DEM}_{\text{final}}$ is constructed for a given shot, any decoder that accepts a hypergraph-based detector error model may be used—such as MLE, MWPM, belief propagation, or others, depending on the QEC code in use.

a. Adjusting loss combination weight

In this section, we explore the influence of the weight of the combination ω on the efficacy of the delayed erasure decoder, which considers both independent error events and the first combination of potential loss events. As described earlier, the decoder's decision-making process involves summing different decoding hypergraphs [see Eq. (B5)]. To incorporate the impact of individual loss events, the first loss combination event is scaled by ω along with the standard Pauli error decoding hypergraph.

To benchmark the performance of our decoder and determine the optimal value of ω , we performed circuit-level simulations on a memory logical qubit with multiple rounds of conventional syndrome extraction. We varied ω between 0 and 1. For $\omega = 0$, each life cycle's decoding hypergraph is considered independently, without including an additional decoding hypergraph for the first combination.

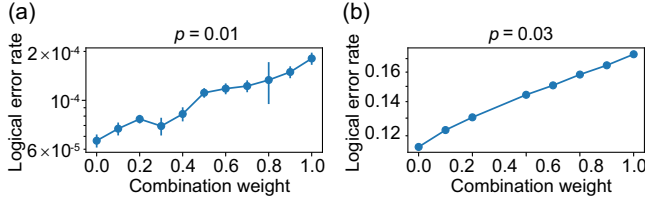


FIG. 10. Logical memory simulations with d rounds of the conventional SE method. Variation of the logical error rate as a function of the combination weight ω for distance 9, across different physical error rates ($p = 0.01$ and $p = 0.03$). The results indicate that nonzero values of ω deteriorate the decoder's performance.

The results for the conventional SE method are presented in Fig. 10. Nonzero values of ω degrade the decoder's performance. However, it is important to note that these results pertain to the conventional SE method, and different SE methods employing distinct gate sets may behave differently. For simplicity, the numerical simulations presented in this paper assume $\omega = 0$.

APPENDIX C: INTERPLAY BETWEEN LOSS, ERASURE, AND BIASED ERRORS

1. Interplay between erasure and biased errors

We examine the combined influence of bias and erasure on logical-memory performance. Using the direct conversion SE method with period 0.25 (the erasure channel) and the first error model described in Appendix F, we perform circuit-level simulations of the XZZX surface code for various bias values under both bias-preserving gates and biased-erasure conditions. The data are decoded using our delayed-erasure decoder together with a minimum-weight perfect-matching decoder. The complete results are shown in Fig. 11. We analyze three scenarios: (a) without bias-preserving gates and unbiased erasure, (b) with bias-preserving gates and biased erasure, and (c) with bias-preserving gates and unbiased erasure. Numerical threshold values are listed in the tables next to each plot. As shown, increasing the erasure fraction influences the thresholds far more strongly than increasing the bias fraction, even in the presence of bias-preserving gates. Furthermore, when bias-preserving gates are absent, the effect of bias is reduced, yielding approximately a twofold improvement in the threshold in the limit of infinite bias (only Z errors).

2. Interplay between loss and biased errors

We now turn to simulations that include loss in the presence of biased-Pauli errors. Using the first error model described in Appendix F, we vary both the bias and loss fractions and calculate the threshold for each parameter set. For the teleportation-based SE method, we employ the XZZX cluster state shown in Fig. 17 (see also Ref. [31]), without bias-preserving gates. For the direct conversion SE

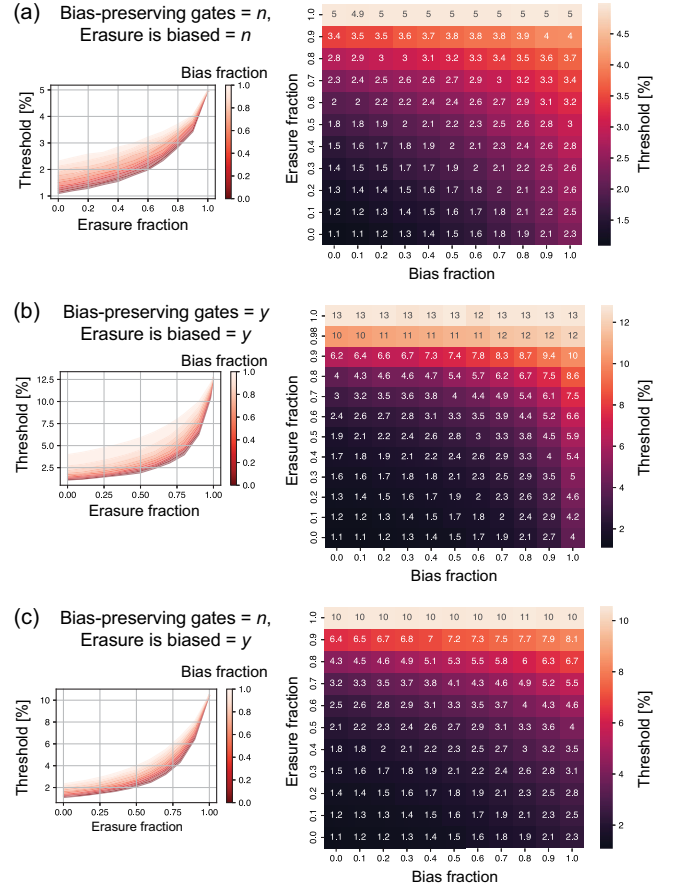


FIG. 11. Thresholds for various values of erasure and bias fractions, in cases (a) without bias-preserving gates and erasure is not biased, (b) with bias-preserving gate and erasure is biased, and (c) without biased-preserving gates and erasure is biased. The heat maps on the right provide numerical values for the thresholds illustrated on the left-hand plots.

method with period 1, we use the surface code, likewise without bias-preserving gates. In this regime, we confirm that the regular surface code and the XZZX surface code exhibit equivalent performance under biased noise, as expected. Circuit-level simulations are performed to estimate the thresholds of each SE method under simultaneous

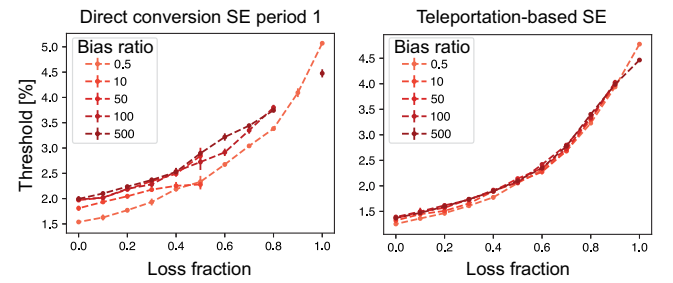


FIG. 12. Thresholds for various values of loss fractions and Pauli bias fractions, for methods (a) direct conversion SE with period 1 and (b) teleportation-based SE. We do not use biased-preserving gates.

loss and biased-Pauli errors. We decode the data using our delayed-erasure decoder combined with an MLE decoder. The results are shown in Fig. 12. Consistent with the erasure-and-bias results above, bias increases the threshold by roughly a factor of 2 in the limit of infinite bias. Moreover, increasing the loss fraction has a considerably stronger effect on the threshold than increasing the bias fraction.

APPENDIX D: SYNDROME EXTRACTION METHODS WITH LOSS DETECTION

As explained in the main text, there are multiple approaches to obtaining syndrome measurements and performing QEC, each with its own advantages and disadvantages. Specifically, each method provides different life cycle lengths and requires varying levels of overhead. Additionally, each method demands different experimental capabilities. Here, we provide further information on each method and the assumptions made in this paper to simulate and compare all methods.

1. Loss detection using physical SWAP SE

This section examines the physical SWAP method for loss detection and correction during rounds of stabilizer measurements. The method integrates stabilizer measurements with loss detection by leveraging the ability to swap both the positions and the quantum information of data and measure qubits in each round. Using SSR, it exploits all three measurement outcomes ($|0\rangle$, $|1\rangle$, or loss) to infer both the stabilizer value and the loss status of the data qubits.

The SWAP method operates by pairing each data qubit with a measure qubit for every stabilizer-check round, as shown in Fig. 13. The measure qubit interacts with its neighboring data qubits through four gates (in the surface-code case). After completing the stabilizer check, the measure and data qubits swap their quantum information and physical locations. If the data qubit remains present, the measurement yields $|0\rangle$ or $|1\rangle$, reflecting the stabilizer outcome. If instead the data qubit is missing, the SWAP operation fails, and the measure qubit heralds loss through SSR. This enables immediate detection of data-qubit loss and replacement with a fresh measure qubit.

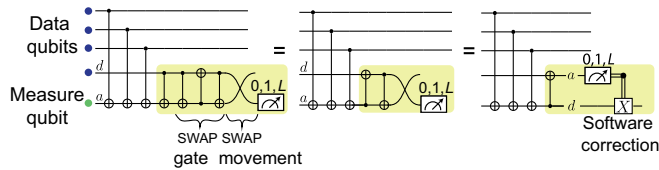


FIG. 13. SWAP SE circuit. Left: the original circuit with the four gates for a parity check, plus a SWAP gate and a SWAP movement. Right: an equivalent circuit that uses only four gates and therefore adds no extra gate overhead compared with the conventional SE approach.

In a conventional SE round (without the additional SWAP operations), SSR detects loss only on measure qubits, whereas during a SWAP SE round, SSR detects loss only on data qubits. Loss of a measure qubit before the SWAP remains undetected, resulting in replacement of a valid data qubit with a lost one; however, this loss is detected in the subsequent SWAP SE round.

a. Optimizing SWAP period

Each SWAP operation incurs a cost, such as idling errors due to movement. Evaluating different SWAP periods is therefore crucial for identifying the frequency that minimizes SWAP operations in practical settings. Generally, increasing the SWAP period extends the operational lifetime of each qubit. Interestingly, periods of 1 and 2 yield the same average qubit life cycle, as illustrated in Fig. 14. Briefly, with a period of 1, both data and measure qubits have an average life cycle of ~ 8 , since they are replaced every round. With a period of 2, data qubits have an average life cycle of ~ 12 because they are not subject to loss detection in every round, while during conventional SE rounds the measure qubits have

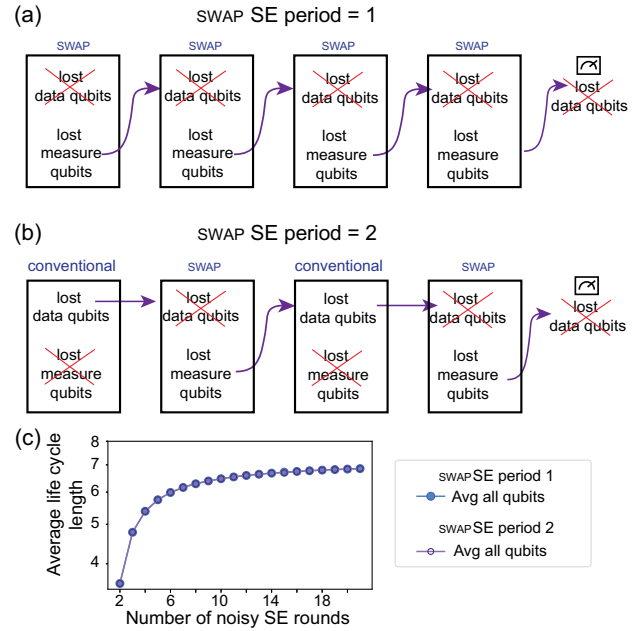


FIG. 14. Life cycle analysis of qubits under SWAP SE with different periods. Panels (a) and (b) depict detected losses (red $\times$ marks) via SSR. (a) Continuous SWAP operations (period = 1): Each loss of a measure qubit is converted into a data-qubit loss in the next cycle. (b) SWAP SE with period 2: Rounds alternate between SWAP and conventional SE. Edge cases (not shown) typically exhibit even longer life cycles due to the absence of SWAP partners. (c) Average qubit life cycles for both methods at $d = 9$. Both periods exhibit the same average life cycle length, which approaches 8 in the large-distance limit.

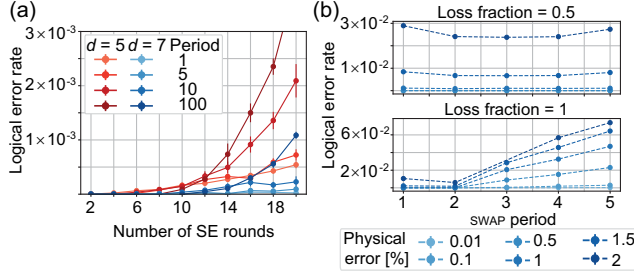


FIG. 15. Optimizing the SWAP period in the presence of movement errors. (a) Logical error vs number of SE rounds for SWAP SE with movement errors at a physical error rate of 0.5% and distances $d \in \{5, 7\}$. Blue (red) curves show $d = 7$ ($d = 5$), with a gradient indicating different SWAP periods. A period of 100 (effectively no SWAP loss detection, meaning conventional SE) yields the worst errors. While small periods are competitive, period 1 is not always optimal. (b) Logical error vs SWAP period for different physical error rates, evaluated after 20 SE rounds, at loss fractions $L \in \{0.5, 1\}$. The optimal period depends on both the loss fraction and the physical error rate.

an average life cycle of only ~ 4 , as they are not converted into data qubits. This makes a period of 2 competitive with a period of 1 in terms of loss detection while reducing complexity.

We incorporate movement-error costs into SWAP logical-memory circuit-level simulations to determine the optimal SWAP period across different loss fractions and physical error rates. According to the data in Fig. 15, the optimal period depends on the specific error model, enabling reduced experimental demands by selecting the most efficient configuration.

b. SWAP SE technical challenges

The strength of the physical SWAP SE method lies in its ability to detect data-qubit loss and replace it with a fresh measure qubit. However, several practical considerations arise.

- (i) **SWAP pairs.** SWAP pairs are organized according to the final gate executed in each round, leaving $O(d)$ qubits along the lattice edge temporarily unpaired. To avoid introducing additional qubits and to ensure that every qubit is paired across rounds, we alternate the gate order between even and odd SWAP rounds. Each SWAP SE round type (even or odd) is characterized by a different ordering of the parity check gates, thus different SWAP pairs.
- (ii) **Movement errors.** During the execution of SWAP rounds, the physical movement of qubits introduces an opportunity for error. These errors, termed “movement errors,” arise due to idle errors that occur during the qubit’s idle time. Let p_{idle} denote the per-slot Pauli error probability during idling (aggregated over axes), τ the duration of one idling slot, and T the total movement (or effective idle)

time. The resulting movement-induced error is modeled as

$$p_{\text{move}} = 1 - (1 - p_{\text{idle}})^{T/\tau}. \quad (\text{D1})$$

This expression captures the accumulated effect of independent idle errors across T/τ slots and is applied uniformly to all qubits that incur the same motion-induced idle time.

c. Comparing SWAP SE and conventional SE

SWAP-based SE enables loss detection during each round, but it has disadvantages discussed above—most notably the sacrifice of some syndrome information and the fact that loss can be detected on only one qubit type (data qubit or measure qubit) at a time.

For short circuits (i.e., small total life cycle lengths), it can be advantageous to avoid SWAP operations altogether. Figure 16(a) compares average life cycles for data and measure qubits under the two methods. In conventional SE, without the SWAP elements at the end of round, data-qubit life cycles grow rapidly with the number of noisy SE

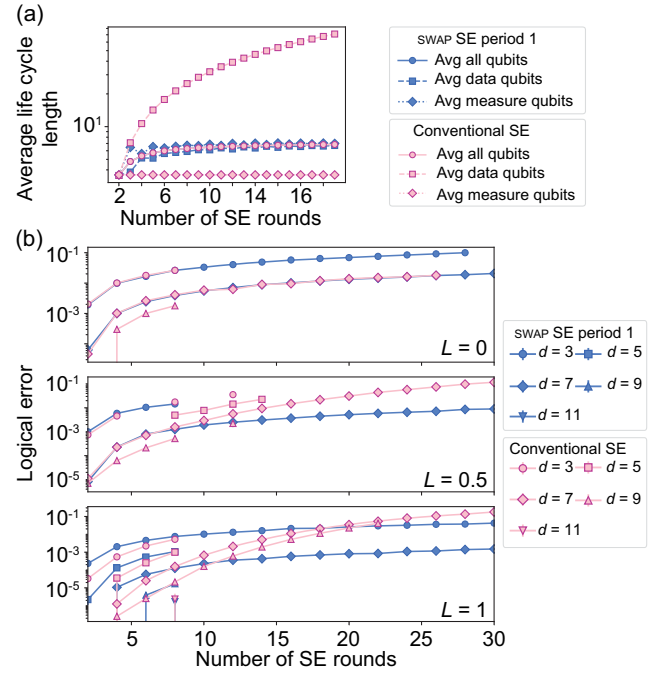


FIG. 16. (a) Life cycle analysis for SWAP SE and conventional SE at distance $d = 9$. The average life cycle over all qubits is equal for both methods. As the number of noisy SE rounds increases, data-qubit life cycles in conventional SE grow, while in SWAP SE they remain constant. (b) Logical error vs number of SE rounds for SWAP SE and conventional SE at physical error rate $p = 1\%$ and various distances. Top to bottom: loss fractions $L = 0, 0.5, 1$. For $L = 1$ (loss only), SWAP SE becomes beneficial over the conventional SE after approximately six SE rounds.

rounds, while in SWAP-based SE they remain constant. Conventional SE maintains short life cycles for measure qubits at the expense of longer life cycles for data qubits; SWAP SE balances both.

We also compare logical performance using circuit-level simulations [Fig. 16(b)]. Both methods use the delayed-erasure decoder developed in this work, and we assume a noiseless first SE round and SSR measurement. As expected, for a small number of rounds, conventional SE outperforms SWAP-based loss detection. The crossover occurs at approximately $k \sim 6$ noisy SE rounds and is roughly distance independent in the data presented here.

These observations can be understood by linking the numerical trends to the life cycle behavior: When only a few noisy SE rounds are executed, the longer data-qubit life cycles induced by SWAP SE do not yet pay off, whereas for deeper circuits, SWAP SE's constant life cycles provide an advantage. This is consistent with the perspective in Ref. [73], where errors on data and measure qubits are effectively categorized into distinct channels and impact thresholds differently.

2. Loss detection using direct conversion SE

Another method considered in this work is direct loss-to-erasure conversion, applied either after every gate, after multiple gates, or after each SE round. In this approach, the experiment detects loss events without requiring additional gates or qubits, instead relying on other hardware capabilities demonstrated in Yb-atom and superconducting-qubit systems [9,11,18]. This technique requires midcircuit measurement and replacement of lost qubits, with overheads that depend on the physical implementation.

Since we cannot simulate these experimental overheads, we assume that once a direct conversion occurs, all losses are converted into delayed erasures, corresponding to an experimental platform that successfully detects and replaces each lost qubit. These simulations therefore provide an upper bound on the threshold and effective distance achievable with midcircuit erasure conversion techniques. We vary the detection and replacement period from 0.25 (detection and replacement after every gate) to $k \geq 1$ (detection and replacement after every k SE rounds). We also consider a distinct scenario in which detection is performed more frequently than replacement, since the two processes rely on different experimental capabilities and may require different timescales. Specifically, we simulate the case of detection after every gate (period 0.25) but replacement only at the end of each round (period 1). We name this approach a direct conversion SE period 1 with loss moment information.

3. Loss detection using teleportation-based SE

In this section, we describe the teleportation-based SE method examined in this work, which employs

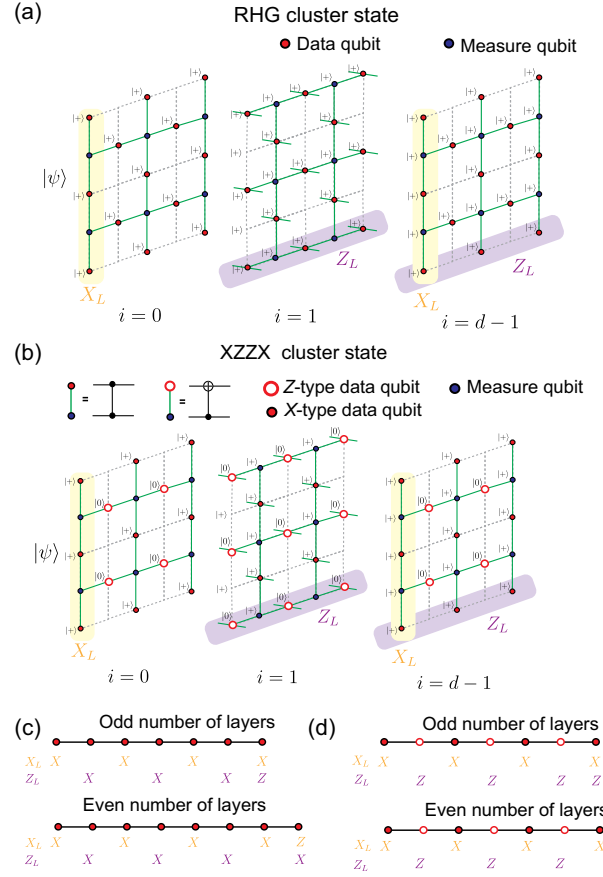


FIG. 17. Implementing teleportation-based SE. Construction of the RHG cluster state (a),(c) and the XZZX cluster state (b),(d) for teleportation-based SE with the surface code.

teleportation of all qubits in each step to both measure stabilizers and detect loss.

We implement the teleportation-based SE architecture using the foliated surface code, also known as the RHG cluster state [32,57,58]. All entangling gates are CZ gates, and qubits are initialized in $|+\rangle$, except for the first layer, which encodes the logical qubit according to the standard surface-code encoding. The main-text results correspond to the XZZX cluster state, which, in the absence of bias-preserving gates, reproduces the RHG cluster state.

Figure 17 shows the full construction, including an explicit example of a distance $d=3$ surface code. The figure also illustrates how the logical operators X_L and Z_L propagate through the cluster state, demonstrating both even and odd number of layers cases. Because each layer detects only X or Z errors, we simulate $2d$ layers for all threshold and effective distance plots presented in this work. This approach introduces a space-time overhead corresponding to a $d \times d \times d$ lattice containing $d \cdot (2d^2 + d^2 - 1) = d \cdot (3d^2 - 1)$ physical qubits, with a cluster state generated by 6 time steps. As in other SE methods simulations, the first and last layers are considered noiseless.

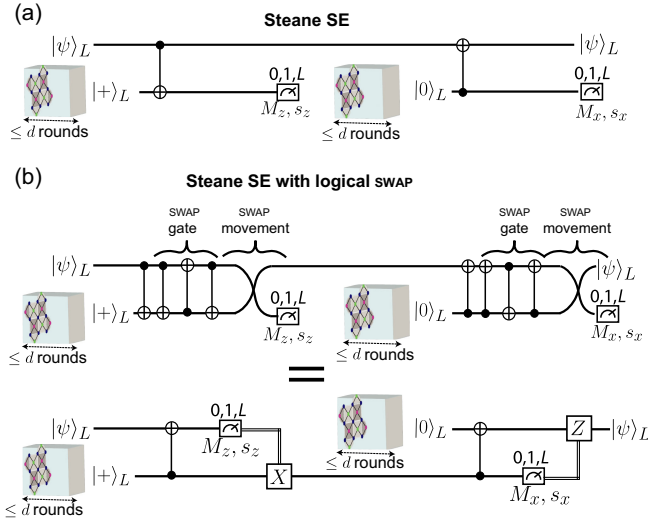


FIG. 18. Adjusting the Steane SE method to utilize loss errors. (a) Regular Steane SE circuit and logical-qubit connections. (b) Adjusted Steane SE with a logical SWAP, yielding a Knill-style SE variant that corrects both Pauli and loss errors. Each SSR measurement is used to detect both error types while teleporting quantum information to another logical qubit.

4. Modified Steane SE

a. Loss detection using modified Steane SE

To provide a comprehensive perspective on SE techniques, we include a modified version of Steane SE adjusted for loss errors, although we do not numerically analyze it in the main text. We highlight its conceptual advantages as a promising direction for future studies. Steane SE uses transversal CX gates applied to fault-tolerant logical measure qubits to extract syndromes and detect Pauli errors [74]. Transversal gates, applying the same operation across corresponding physical qubits in each code block, ensure that errors propagate predictably from logical data qubits to logical measure qubits. If the measure qubits are prepared fault tolerantly with d SE rounds, one round of Steane SE suffices to capture the syndromes in a given basis [74]. However, conventional Steane SE lacks a native mechanism for detecting loss: Losses on data qubits are not observed by the logical measure qubits and remain undetected until the end.

We modify conventional Steane SE by incorporating logical SWAP operations inspired by the physical SWAP operations used in SWAP SE, as illustrated in Fig. 18, thereby enabling loss detection through SSR at each logical measurement. These logical SWAP operations exchange the information of the logical data and measure qubits at each transversal CX gate. This effectively teleports the logical data qubits at every syndrome extraction step, similar in spirit to Ref. [75], and ensures that no physical qubit retains a long life cycle. This modified Steane SE scheme combines the benefits of teleportation-based techniques with the simplicity of transversal

operations, minimizing qubit life cycles while integrating loss detection into the SE process.

A key advantage of Steane SE relative to the other methods studied here is the ability to preselect on the quality of the logical measure-qubit blocks. To gauge the quality of the prepared measure-qubit blocks, one can employ the logical-gap method for preselection, potentially increasing fidelity without significant overhead [76–78]. This method uses the syndromes collected during preparation to compute the gap between the probability of the proposed error and that probability conditioned on applying the logical operator. A large (small) gap indicates that the decoder is confident (not confident) in its correction. One can then set a cutoff on the gap and keep only measure-qubit blocks whose gap exceeds the threshold.

b. Connecting Steane SE with logical SWAP to teleportation-based SE

The modified Steane SE can be viewed as an interpolation between modified conventional SE and teleportation-based SE, similar to the ideas in Ref. [79]. In the limit of many logical measure qubits, each prepared with a single SE round, the modified Steane SE becomes equivalent to teleportation-based SE (see Fig. 19). Conversely, multiple SE rounds can be used to prepare higher-quality logical measure qubits, providing flexibility to match different error models and resource constraints. Thus, Steane SE supports any combination between these two options, with $n \in \{1, \dots, d\}$ preparation rounds per logical measure qubit and $k \in \{1, \dots, d\}$ rounds of logical teleportation. Because modified Steane SE interpolates between modified conventional SE and teleportation-based SE, we expect its performance to qualitatively reflect this interpolation.

APPENDIX E: COMPARING SE METHODS LIFE CYCLES AND SPACE-TIME OVERHEADS

This paper considers multiple approaches for detecting loss and converting to delayed-erasure errors. As detailed above, we consider methods based on midcircuit loss detection, known as erasure conversion [11,18] in different frequencies (every gate, few gates, rounds), named direct conversion SE. We also analyze methods that rely solely on SSR, which require adjusting the SE approach to utilize teleportation, such as SWAP SE, and teleportation-based SE. A further method that can be modified to incorporate teleportation is Steane SE, which we refer to as the modified Steane SE. Table I summarizes the average life cycles, error thresholds, and space-time overheads for each SE method considered numerically in this paper. The results are presented for multiple error models, including both independent and correlated loss channels between qubits participating in entangling gates.

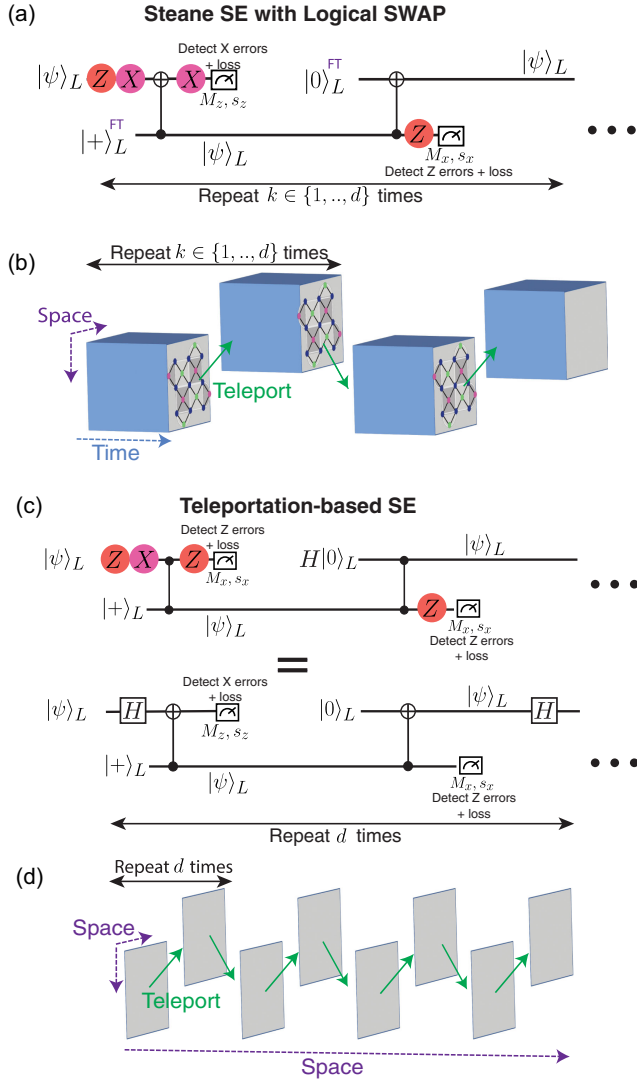


FIG. 19. Connection between Steane SE with logical SWAP and teleportation-based SE, revealing an intermediate scheme. (a) Steane SE circuit (without Pauli corrections), showing logical data and measure qubits, error propagation, and detection. This circuit can repeat $k \in \{1, \dots, d\}$ times; $k = 1$ is standard Steane SE, and $k > 1$ uses logical measure qubits that are prepared with fewer rounds and can be preselected. (b) Teleportation illustration of the modified Steane QEC. (c) Teleportation-based SE circuit (without Pauli corrections), equivalent to (a) for $k = d$. (d) Teleportation process in teleportation-based SE. The red and pink circles with X and Z in (a) and (c) indicate errors and their propagation or detection.

APPENDIX F: NUMERICAL SIMULATIONS METHODS

We begin by outlining the metaparameters considered in this work, which are determined by the experimental platform.

- (i) Bias-preserving gates: Boolean parameter indicating whether the system's two-qubit gates preserve error bias (i.e., Z errors remain Z). Bosonic systems have

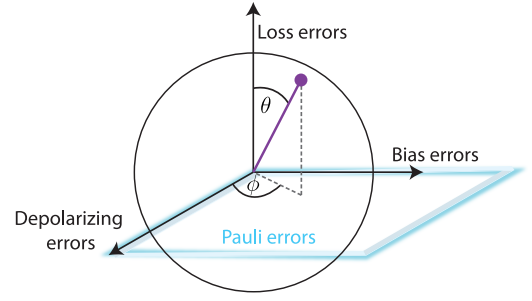


FIG. 20. Representation of the error model on a sphere.

demonstrated such gates, whereas atom-based platforms that implement only CZ gates, without direct CX execution, generally do not.

- (ii) Loss is biased: Boolean parameter specifying whether loss events are biased toward a particular state. For example, in neutral-atom systems where losses predominantly occur from $|1\rangle$, loss is considered biased. Replacing a lost qubit with $|1\rangle$ primarily induces Z errors, reflecting a bias in the effective error channel.
- (iii) Architecture: Distinguishes between measurement-based quantum computation (used in teleportation-based SE) and circuit-based computation (used in other SE methods).
- (iv) Period of loss detection rounds: The number is set by the chosen loss detection approach.
- (v) Cost of loss detection rounds: Depends on the selected strategy and experimental overhead.
- (vi) QEC code: Any CSS code can be used; here we focus on the standard and XZZX surface codes.
- (vii) State-selective readout: Indicates whether the measurement apparatus can resolve loss events to produce three outcomes: $|0\rangle$, $|1\rangle$, and $|L\rangle$ (loss).

Next, we describe the error models that incorporate both loss and biased-Pauli errors. Throughout this work, we consider errors occur primarily during two-qubit gates, each with probability p_{CZ} (the physical error parameter reported in all numerical plots as the physical error rate, specifically, reported in Figs. 2, 3, 4, 5, 7 in the main text and in Figs. 11, 12, 15, 16, 22, 23, 24, 25, 26 in the Appendix.). We simulate several error models that include correlations between loss and Pauli errors, or between the two qubits involved in a gate.

Error model A: Loss-biased Pauli-channel correlations on each qubit. Each qubit experiences an independent error channel with probability $p = 1 - \sqrt{1 - p_{CZ}}$, ensuring channel normalization. Two parameters define the model: bias and loss fraction. The relationship can be represented on a unit sphere, as shown in Fig. 20. The Z axis corresponds to loss errors, which remove the qubit from the computational subspace, while the XY plane corresponds to Pauli errors, which remain within it. Pauli errors range from uniform (depolarizing) to single-channel

TABLE I. Comparison of SE methods under various error models. Each column represents a method, and each row describes a property such as average qubit life cycles, thresholds, effective distance and overheads. The final column estimates performance under very frequent detection and replacement, providing the erasure channel bound. Channel definitions for $L = 1$ are as follows. Channel 1: $p'\{L \otimes I, I \otimes L\}$. Channel 2: $(p'/2)\{L \otimes Z, Z \otimes L, L \otimes I, I \otimes L\}$. $p' = 1 - \sqrt{1 - p} \sim p/2$ where p is the entangling gate physical error rate.

Category	Conventional SE	SWAP SE period 1	Direct conversion SE period 2	Teleportation-based SE	Direct conversion SE period 1	Direct conversion SE period 1 plus exact loss information	Direct conversion SE period 0.25 plus exact loss information
Average life cycle length of data qubits	$4d$	8	8	4	4	1 (detection), 4 (replacement)	1 (detection), 1 (replacement)
Average life cycle length of measure qubits	4	8	4	4	4	1 (detection), 4 (replacement)	1 (detection), 1 (replacement)
Space-time overhead	$\sim 8d^3 - 4d$	$\sim 8d^3 - 4d$	$\sim 8d^3 - 4d$	$\sim 18d^3 - 6d$	$\sim 8d^3 - 4d$	$\sim 8d^3 - 4d$	$\sim 8d^3 - 4d$
Extra experimental requirements	None	SSR	Erasure conversion [11,18], detection and replacement every 2 SE rounds	SSR	Erasure conversion [11,18], detection and replacement every SE round	Erasure conversion [11,18], detection every gate, and replacement every SE round	Erasure conversion [11,18], detection and replacement every gate
Threshold for $L = 1$ [%] (Channel 1)	Not applicable	2.27 ± 0.02	3.98 ± 0.06	4.78 ± 0.01	5.07 ± 0.02	7.16 ± 0.01	9.51 ± 0.04
Effective distance for $d = 7$ (Channel 1)	Not applicable	6.41 ± 0.27	6.77 ± 0.2	7.19 ± 0.45	6.83 ± 0.37	6.97 ± 0.24	7.46 ± 0.36
Threshold for $L = 1$ [%] (Channel 2)	Not applicable	1.94 ± 0.01	2.75 ± 0.07	3.5 ± 0.05	3.76 ± 0.01	5.21 ± 0.01	6.5 ± 0.01
Effective distance for $d = 7$ (Channel 2)	Not applicable	6.13 ± 0.26	5.5 ± 0.23	6.5 ± 0.21	6.32 ± 0.87	6.45 ± 0.47	7.47 ± 0.49

(highly biased) noise, parametrized by angle ϕ . The angle θ links the probabilities of loss and Pauli errors and determines the loss fraction. Errors are normalized analogously to a Bloch-sphere vector.

Given the total error probability p , the likelihood of a loss event is controlled by

$$L = \cos^2 \theta = \frac{p_{\text{loss}}}{p_{\text{loss}} + p_{\text{Pauli}}}.$$

The biased-Pauli channel is described by

$$p_x = p_y = \frac{p}{2(1+\eta)}, \quad p_z = \eta(p_x + p_y) = \frac{\eta p}{1+\eta}, \quad (\text{F1})$$

where η is the bias parameter related to ϕ .

Normalization check:

$$P(\text{error}) = \underbrace{Lp}_{\text{loss channel}} + \underbrace{(1-L)(p_x + p_y + p_z)}_{\text{biased-Pauli channel}} = p. \quad (\text{F2})$$

Unless stated otherwise, this is the default model used in our numerical simulations. Results based on this model appear in the main text, Table I, and Figs. 12, 15, 16, 22, 23, and 24.

Error model B: Loss-Pauli correlations between qubits. For $L = 1$ (loss-only errors), a loss of one qubit induces a Z-error channel on its partner within the two-qubit gate. The channel is

$$\frac{p}{2} \{Z \otimes L, L \otimes I, Z \otimes L, L \otimes I\},$$

with $p \approx p_{\text{CZ}}/2$. When a qubit is lost, its neighbor experiences a $\frac{1}{2}\{Z, I\}$ error channel. This is the error model used in Ref. [23]; our results appear in Table I and Figs. 25 and 26.

Error model C: Erasure-biased Pauli-channel correlations on each qubit. This model is identical to error model A except that loss channels are replaced by perfect erasure channels—representing ideal loss detection and qubit replacement after every gate (direct conversion SE with period 0.25). This corresponds to the model used in Ref. [17]; results appear in the main text for direct conversion SE with period 0.25, Table I, Figs. 11, 23, and 24.

APPENDIX G: ALGORITHMIC PROCEDURES

Here we describe the approach used to calculate the average and maximum life cycle length for each algorithmic procedure shown in Fig. 21. For each algorithm, we count the number of noisy SE rounds per logical qubit, from initialization to the final SSR measurement. Non-Clifford gates are assumed to be implemented through teleported-gate circuits, which increase the life cycle but also allow natural loss detection and termination of the life

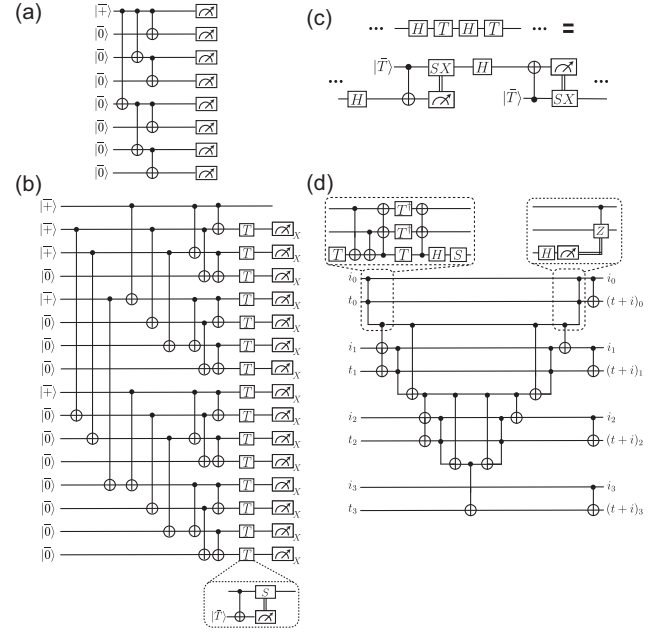


FIG. 21. Common logical algorithmic procedures. (a) GHZ state preparation. (b) Magic-state distillation circuit. (c) HT decomposition algorithm. (d) Adder circuit.

cycle for one of the logical data qubits in the teleported-gate circuit.

Inspired by the results in Refs. [13,14] and our numerical results presented in the main text, we consider a single SE round for state preparation and a single SE round after every gate. Accordingly, we count the number of logical CX gates per logical qubit as follows.

- (1) GHZ: We use the logarithmic-depth implementation shown in Fig. 21(a). For n qubits, the average number of CX gates per qubit is approximately $(\log n)/2$, and the maximum is $\lfloor \log n \rfloor$.
- (2) Magic-state distillation: Shown in Fig. 21(b). For each input qubit, there are three or four entangling gates before the logical T gate, which adds another entangling gate and ends the life cycle of that input qubit. The T gate is performed via teleportation, which measures out the input qubits, giving a total of four or five entangling gates from initialization to measurement. The output qubit passes through three CX gates before being teleported and measured in the next distillation layer. Thus, the average number of CX gates per qubit is 4, with a maximum of 5, independent of the number of layers.
- (3) HT decomposition algorithm (small-angle synthesis) [35,36]: As shown in Fig. 21(c), each qubit is measured shortly after initialization, giving a short life cycle. An input magic state undergoes only two transversal entangling gates before measurement. Assuming it originates from a magic-state distillation circuit, the average life cycle length is 7, and the maximum is 8.

- (4) Adder circuit [34]: Shown in Fig. 21(d). The counting differs across qubit rows. Counting entangling gates based on T -gate teleportation without SWAP yields an average of 9 in the limit of many logical qubits, with a maximum of 13. Specifically, for the top three qubits, the life cycle lengths are 6, 6, and 11; for the bottom two, 1 and 2; and for intermediate qubits (in the large-system limit), 8, 7, and 13.

APPENDIX H: ADDITIONAL NUMERICAL RESULTS

Here we present additional circuit-level simulation results that complement those discussed in the main text. All results here are decoded using the delayed-erasure decoder developed in this work, combined with an MLE decoder.

1. Logical error rate vs loss fraction for various SE methods

Figure 22 shows the logical error rate as a function of the number of SE rounds for several SE methods. Each subplot corresponds to a different loss fraction L : 0, 0.5, and 1, respectively. For $L = 0$, all methods yield comparable results, with teleportation-based SE exhibiting a slightly higher logical error due to additional gate overhead. For $L = 0.5$, all methods that incorporate loss detection—SWAP SE, teleportation-based SE, and direct conversion SE—show similar performance and support deeper circuits. For $L = 1$ (loss-only errors), SWAP SE yields higher logical error rates than the other methods because of its longer qubit life cycles, as explained above.

2. Thresholds as a function of life cycle length

In Fig. 4(b) of the main text, we present the thresholds of multiple SE methods as a function of life cycle length. Fitting these data points (excluding SWAP SE) to both exponential and polynomial models shows that the decay

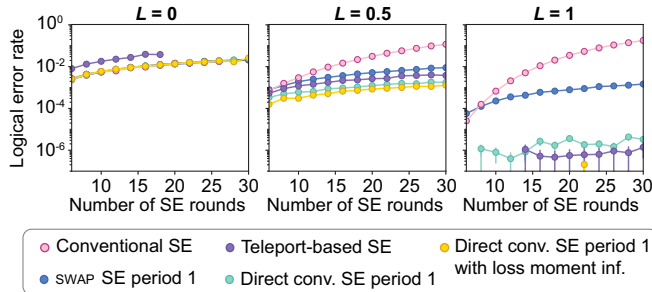


FIG. 22. Comparison of different SE methods for various loss fractions L . Logical-memory circuit-level simulations showing the logical error rate as a function of the number of SE rounds. (a) $L = 0$, (b) $L = 0.5$, (c) $L = 1$. The physical error rate is 1% and the code distance is $d = 7$.

follows a polynomial scaling. Numerically, we observe a decay proportional to $7/(\text{life cycle length})^{1/3}$.

3. Thresholds and effective distances for various error models and SE methods

Figures 23–26 show the logical error rate as a function of the physical error rate for various SE methods at a loss fraction of 1. We consider two different error models: the loss-only model (error model A in Appendix F, Figs. 23 and 24) and the correlated loss-Z model (error model B in Appendix F, Figs. 25 and 26). These results are used to extract both the threshold and effective distance for each error model and SE method.

a. Comparison with published errors models

As a validation step, we benchmark our simulations against other error models, presented in Refs. [17,18,23]. First, we perform circuit-level simulations for the case of erasure errors (detected immediately after leaving the qubit subspace), corresponding to direct conversion SE with period 0.25 and perfect operations. We then compare our results with those reported in Refs. [17,18]. Using the XZZX surface code at various erasure fractions and the two-qubit error model of Ref. [17], and decoding with our delayed-erasure + MWPM decoder, we reproduce the thresholds reported in Table 1 of Ref. [17], consistent with the data shown in Fig. 11.

We further benchmark the teleportation-based SE using the loss-Z correlated error model used in Ref. [23]. Employing our delayed-erasure decoder, we obtain comparable thresholds of approximately 3.5% for a loss

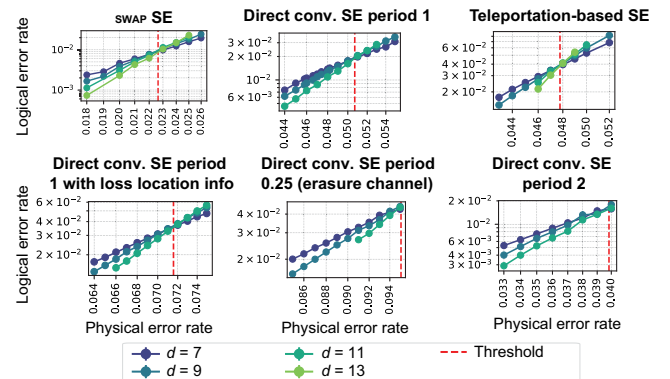


FIG. 23. Threshold data: Logical error rate vs physical error rate for different SE methods under loss-only errors ($L = 1$). SE methods considered here: SWAP SE, direct conversion SE (period 1), teleportation-based SE, direct conversion SE (period 1) with detection period 0.25 (perfect loss moment information), direct conversion SE (period 0.25) with detection and replacement after every gate (erasure channel), direct conversion SE (period 2). These data are used to determine the threshold for each SE method.

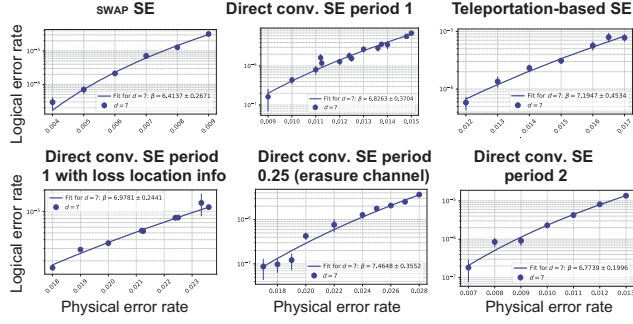


FIG. 24. Effective distance data for loss-only errors ($L = 1$). SE methods considered here: SWAP SE, direct conversion SE (period 1), teleportation-based SE, direct conversion SE (period 1) with detection period 0.25 (perfect loss moment information), direct conversion SE (period 0.25) with detection and replacement after every gate (erasure channel), direct conversion SE (period 2). These data are used to determine the threshold for each SE method.

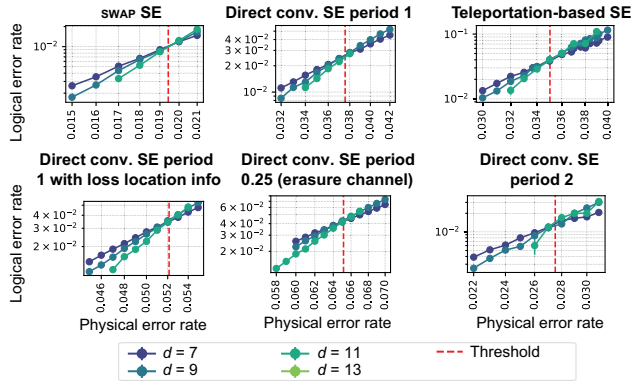


FIG. 25. Same as Fig. 23, but for a correlated Z-loss channel (error model B), as described in Appendix F.

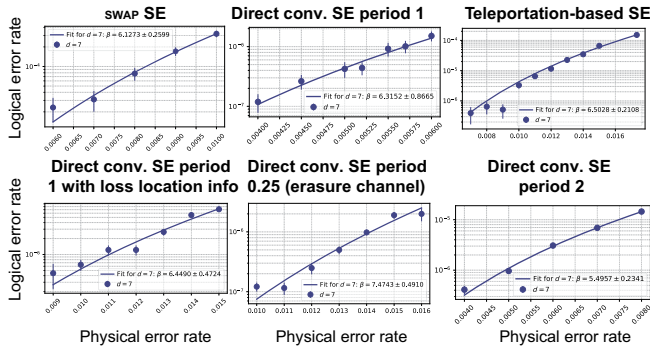


FIG. 26. Same as Fig. 24, but for a correlated Z-loss error channel (error model B), as described in Appendix F.

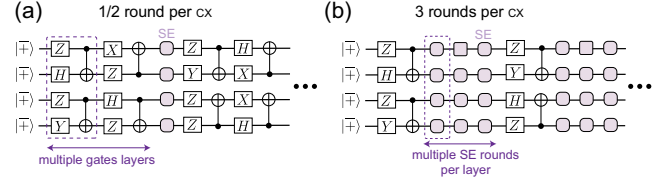


FIG. 27. Illustration of deep random Clifford circuits with multiple layers of transversal CX and single-qubit logical gates, including periodic SE rounds. (a) One SE round every two CX layers ($n_r = 0.5$). (b) Three SE rounds per CX layer ($n_r = 3$).

fraction of 1. These results are included in Table I and Fig. 25.

APPENDIX I: RANDOM DEEP LOGICAL TRANSVERSAL CLIFFORD CIRCUITS

We study deep logical Clifford circuits composed of multiple layers of transversal gates, similarly to Ref. [13]. Each layer consists of single-qubit logical gates from the set $\{X_L, Y_L, Z_L\}$ and transversal CNOT gates applied between random pairs of qubits in a random order. After each layer, we perform n_r rounds of SE. For $n_r < 1$, a single QEC round is performed after every $1/n_r$ layers, while for $n_r \geq 1$, n_r SE rounds are executed after each layer, see illustration in Fig. 27. We use the circuit-level biased-noise error model A described in Appendix F. Unless otherwise noted, these noise channels are applied throughout the entire circuit, except during state preparation, the final stabilizer measurements of the last transversal layer, and the final logical stabilizer measurements.

The figures in the main text show the algorithmic logical error probability P_L . The maximum error $P_{L,\max} = 1 - (1/2^N)$ corresponds to a maximally mixed logical state with N logical qubits.

- [1] P. W. Shor, *Scheme for reducing decoherence in quantum computer memory*, *Phys. Rev. A* **52**, R2493 (1995).
- [2] A. R. Calderbank and P. W. Shor, *Good quantum error-correcting codes exist*, *Phys. Rev. A* **54**, 1098 (1996).
- [3] P. W. Shor, *Fault-tolerant quantum computation*, in *Proceedings of 37th Conference on Foundations of Computer Science* (IEEE, New York, 1996), pp. 56–65.
- [4] D. Bluvstein, S. J. Evered, A. A. Geim, S. H. Li, H. Zhou, T. Manovitz, S. Ebadi, M. Cain, M. Kalinowski, D. Hangleiter *et al.*, *Logical quantum processor based on reconfigurable atom arrays*, *Nature (London)* **626**, 58 (2024).
- [5] Google Quantum AI, *Suppressing quantum errors by scaling a surface code logical qubit*, *Nature (London)* **614**, 676 (2023).
- [6] H. Putterman, K. Noh, C. T. Hann, G. S. MacCabe, S. Aghaeimeibodi, R. N. Patel, M. Lee, W. M. Jones, H. Moradinejad, R. Rodriguez *et al.*, *Hardware-efficient quantum error correction using concatenated bosonic qubits*, *arXiv:2409.13025*.

- [7] M. Da Silva, C. Ryan-Anderson, J. Bello-Rivas, A. Chernoguzov, J. Dreiling, C. Foltz, F. Frachon, J. Gaebler, T. Gatterman, L. Grans-Samuelsson *et al.*, *Demonstration of logical qubits and repeated error correction with better-than-physical error rates*, [arXiv:2404.02280](#).
- [8] B. W. Reichardt, A. Paetznick, D. Aasen, I. Basov, J. M. Bello-Rivas, P. Bonderson, R. Chao, W. van Dam, M. B. Hastings, A. Paz *et al.*, *Logical computation demonstrated with a neutral atom quantum processor*, [arXiv:2411.11822](#).
- [9] R. Acharya, L. Aghababaie-Beni, I. Aleiner, T. I. Andersen, M. Ansmann, F. Arute, K. Arya, A. Asfaw, N. Astrakhantsev, J. Atalaya *et al.*, *Quantum error correction below the surface code threshold*, *Nature (London)* **638**, 920 (2025).
- [10] D. Bluvstein, A. A. Geim, S. H. Li, S. J. Evered, J. Ataides, G. Baranes, A. Gu, T. Manovitz, M. Xu, M. Kalinowski *et al.*, *Architectural mechanisms of a universal fault-tolerant quantum computer*, *Nature (London)* **649**, 39 (2026).
- [11] S. Ma, G. Liu, P. Peng, B. Zhang, S. Jandura, J. Claes, A. P. Burgers, G. Pupillo, S. Puri, and J. D. Thompson, *High-fidelity gates and mid-circuit erasure conversion in an atomic qubit*, *Nature (London)* **622**, 279 (2023).
- [12] P. Scholl, A. L. Shaw, R. B.-S. Tsai, R. Finkelstein, J. Choi, and M. Endres, *Erasure conversion in a high-fidelity Rydberg quantum simulator*, *Nature (London)* **622**, 273 (2023).
- [13] M. Cain, C. Zhao, H. Zhou, N. Meister, J. P. B. Ataides, A. Jaffe, D. Bluvstein, and M. D. Lukin, *Correlated decoding of logical algorithms with transversal gates*, *Phys. Rev. Lett.* **133**, 240602 (2024).
- [14] H. Zhou, C. Zhao, M. Cain, D. Bluvstein, C. Duckering, H.-Y. Hu, S.-T. Wang, A. Kubica, and M. D. Lukin, *Algorithmic fault tolerance for fast quantum computing*, [arXiv:2406.17653](#).
- [15] S. J. Evered, D. Bluvstein, M. Kalinowski, S. Ebadi, T. Manovitz, H. Zhou, S. H. Li, A. A. Geim, T. T. Wang, N. Maskara *et al.*, *High-fidelity parallel entangling gates on a neutral-atom quantum computer*, *Nature (London)* **622**, 268 (2023).
- [16] K. C. Miao, M. McEwen, J. Atalaya, D. Kafri, L. P. Pryadko, A. Bengtsson, A. Opremcak, K. J. Satzinger, Z. Chen, P. V. Klimov *et al.*, *Overcoming leakage in quantum error correction*, *Nat. Phys.* **19**, 1780 (2023).
- [17] K. Sahay, J. Jin, J. Claes, J. D. Thompson, and S. Puri, *High-threshold codes for neutral-atom qubits with biased erasure errors*, *Phys. Rev. X* **13**, 041013 (2023).
- [18] Y. Wu, S. Kolkowitz, S. Puri, and J. D. Thompson, *Erasure conversion for fault-tolerant quantum computing in alkaline earth Rydberg atom arrays*, *Nat. Commun.* **13**, 4657 (2022).
- [19] M. Kang, W. C. Campbell, and K. R. Brown, *Quantum error correction with metastable states of trapped ions using erasure conversion*, *PRX Quantum* **4**, 020358 (2023).
- [20] A. Kubica, A. Haim, Y. Vakhin, H. Levine, F. Brandão, and A. Retzker, *Erasure qubits: Overcoming the T_1 limit in superconducting circuits*, *Phys. Rev. X* **13**, 041022 (2023).
- [21] K. Chang, S. Singh, J. Claes, K. Sahay, J. Teoh, and S. Puri, *Surface code with imperfect erasure checks*, [arXiv:2408.00842](#).
- [22] S. Gu, A. Retzker, and A. Kubica, *Fault-tolerant quantum architectures based on erasure qubits*, [arXiv:2312.14060](#).
- [23] C.-C. Yu, Z.-H. Chen, Y.-H. Deng, M.-C. Chen, C.-Y. Lu, and J.-W. Pan, *Tracking and decoding Rydberg leakage error with MBQC*, [arXiv:2411.04664](#).
- [24] H. Perrin, S. Jandura, and G. Pupillo, *Quantum error correction resilient against atom loss*, [arXiv:2412.07841](#).
- [25] M. Chow, V. Buchemavari, S. Omanakuttan, B. Little, S. Pandey, I. Deutsch, and Y. Jau, *Circuit-based leakage-to-erasure conversion in a neutral atom quantum processor*, [arXiv:2405.10434](#).
- [26] I. Cong, H. Levine, A. Keesling, D. Bluvstein, S.-T. Wang, and M. D. Lukin, *Hardware-efficient, fault-tolerant quantum computation with Rydberg atoms*, *Phys. Rev. X* **12**, 021049 (2022).
- [27] A. B. Ziad, A. Zalawadiya, C. Topal, J. Camps, G. P. Gehér, M. P. Stafford, and M. L. Turner, *Local clustering decoder: A fast and adaptive hardware decoder for the surface code*, [arXiv:2411.10343](#).
- [28] J. S. Nelson and A. D. Baczewski, *Assessment of quantum phase estimation protocols for early fault-tolerant quantum computers*, *Phys. Rev. A* **110**, 042420 (2024).
- [29] J. P. Bonilla Ataides, D. K. Tuckett, S. D. Bartlett, S. T. Flammia, and B. J. Brown, *The XXXX surface code*, *Nat. Commun.* **12**, 2172 (2021).
- [30] A. S. Darmawan, B. J. Brown, A. L. Grimsmo, D. K. Tuckett, and S. Puri, *Practical quantum error correction with the XXXX code and Kerr-cat qubits*, *PRX Quantum* **2**, 030345 (2021).
- [31] J. Claes, J. E. Bourassa, and S. Puri, *Tailored cluster states with high threshold under biased noise*, *npj Quantum Inf.* **9**, 9 (2023).
- [32] R. Raussendorf, D. E. Browne, and H. J. Briegel, *Measurement-based quantum computation on cluster states*, *Phys. Rev. A* **68**, 022312 (2003).
- [33] S. Bravyi and A. Kitaev, *Universal quantum computation with ideal Clifford gates and noisy ancillas*, *Phys. Rev. A* **71**, 022316 (2005).
- [34] C. Gidney, *Halving the cost of quantum addition*, *Quantum* **2**, 74 (2018).
- [35] A. Y. Kitaev, *Quantum computations: Algorithms and error correction*, *Russ. Math. Surv.* **52**, 1191 (1997).
- [36] A. Y. Kitaev, A. Shen, and M. N. Vyalii, *Classical and Quantum Computation* (American Mathematical Society, Providence, RI, 2002), p. 47.
- [37] H. Levine, A. Haim, J. S. Hung, N. Alidoust, M. Kalaei, L. DeLorenzo, E. A. Wollack, P. Arrangoiz-Arriola, A. Khalajhedayati, R. Sanil *et al.*, *Demonstrating a long-coherence dual-rail erasure qubit using tunable transmons*, *Phys. Rev. X* **14**, 011051 (2024).
- [38] S. J. de Graaf, S. H. Xue, B. J. Chapman, J. D. Teoh, T. Tsunoda, P. Winkel, J. W. Garmon, K. M. Chang, L. Frunzio, S. Puri *et al.*, *A mid-circuit erasure check on a dual-rail cavity qubit using the joint-photon number-splitting regime of circuit QED*, *npj Quantum Inf.* **11**, 1 (2025).
- [39] K. S. Chou, T. Shemma, H. McCarrick, T.-C. Chien, J. D. Teoh, P. Winkel, A. Anderson, J. Chen, J. Curtis, S. J. de Graaf *et al.*, *Demonstrating a superconducting dual-rail cavity qubit with erasure-detected logical measurements*, [arXiv:2307.03169](#).

- [40] A. G. Fowler, *Coping with qubit leakage in topological codes*, *Phys. Rev. A* **88**, 042308 (2013).
- [41] P. Aliferis and B. M. Terhal, *Fault-tolerant quantum computation for local leakage faults*, *Quantum Inf. Comput.* **7**, 139 (2007).
- [42] M. Suchara, A. W. Cross, and J. M. Gambetta, *Leakage suppression in the toric code*, in *Proceeding of the 2015 IEEE International Symposium on Information Theory (ISIT)* (IEEE, New York, 2015), pp. 1119–1123.
- [43] K. S. Chou, T. Shemma, H. McCarrick, T.-C. Chien, J. D. Teoh, P. Winkel, A. Anderson, J. Chen, J. C. Curtis, S. J. de Graaf *et al.*, *A superconducting dual-rail cavity qubit with erasure-detected logical measurements*, *Nat. Phys.* **20**, 1454 (2024).
- [44] E. Deist, Y.-H. Lu, J. Ho, M. K. Pasha, J. Zeiher, Z. Yan, and D. M. Stamper-Kurn, *Mid-circuit cavity measurement in a neutral atom array*, *Phys. Rev. Lett.* **129**, 203602 (2022).
- [45] B. Hu, J. Sinclair, E. Bytyqi, M. Chong, A. Rudelis, J. Ramette, Z. Vendeiro, and V. Vuletić, *Site-selective cavity readout and classical error correction of a 5-bit atomic register*, [arXiv:2408.15329](https://arxiv.org/abs/2408.15329).
- [46] J. W. Lis, A. Senoo, W. F. McGrew, F. Rönchen, A. Jenkins, and A. M. Kaufman, *Midcircuit operations using the omg architecture in neutral atom arrays*, *Phys. Rev. X* **13**, 041035 (2023).
- [47] T. M. Graham, L. Phuttitarn, R. Chinnarasu, Y. Song, C. Poole, K. Jooya, J. Scott, A. Scott, P. Eichler, and M. Saffman, *Midcircuit measurements on a single-species neutral alkali atom quantum processor*, *Phys. Rev. X* **13**, 041051 (2023).
- [48] W. Huie, L. Li, N. Chen, X. Hu, Z. Jia, Won Kyu Calvin Sun, and J. P. Covey, *Repetitive readout and real-time control of nuclear spin qubits in ^{171}Yb atoms*, *PRX Quantum* **4**, 030337 (2023).
- [49] O. Higgott and C. Gidney, *Sparse blossom: Correcting a million errors per core second with minimum-weight matching*, [arXiv:2303.15933](https://arxiv.org/abs/2303.15933).
- [50] A. Koottandavida, I. Tsioutsios, A. Kargioti, C. R. Smith, V. R. Joshi, W. Dai, J. D. Teoh, J. C. Curtis, L. Frunzio, R. J. Schoelkopf *et al.*, *Erasure detection of a dual-rail qubit encoded in a double-post superconducting cavity*, *Phys. Rev. Lett.* **132**, 180601 (2024).
- [51] T. M. Stace, S. D. Barrett, and A. C. Doherty, *Thresholds for topological codes in the presence of loss*, *Phys. Rev. Lett.* **102**, 200501 (2009).
- [52] S. D. Barrett and T. M. Stace, *Fault tolerant quantum computation with very high threshold for loss errors*, *Phys. Rev. Lett.* **105**, 200502 (2010).
- [53] C. Gidney, *Stim: A fast stabilizer circuit simulator*, *Quantum* **5**, 497 (2021).
- [54] Y. Wu and L. Zhong, *Fusion blossom: Fast MWPM decoders for QEC*, in *Proceedings of the 2023 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, New York, 2023), Vol. 1, pp. 928–938.
- [55] P. Panteleev and G. Kalachev, *Degenerate quantum LDPC codes with good finite length performance*, *Quantum* **5**, 585 (2021).
- [56] J. Bausch, A. W. Senior, F. J. Heras, T. Edlich, A. Davies, M. Newman, C. Jones, K. Satzinger, M. Y. Niu, S. Blackwell *et al.*, *Learning high-accuracy error decoding for quantum processors*, *Nature (London)* **635**, 834 (2024).
- [57] R. Raussendorf and H. J. Briegel, *A one-way quantum computer*, *Phys. Rev. Lett.* **86**, 5188 (2001).
- [58] D. B. Robert Raussendorf and H. Briegel, *The one-way quantum computer—A non-network model of quantum computation*, *J. Mod. Opt.* **49**, 1299 (2002).
- [59] J. Ghosh, A. G. Fowler, J. M. Martinis, and M. R. Geller, *Understanding the effects of leakage in superconducting quantum-error-detection circuits*, *Phys. Rev. A* **88**, 062329 (2013).
- [60] N. Mehta, J. D. Teoh, T. Noh, A. Agrawal, A. Anderson, B. Birdsall, A. Brahmabhatt, W. Byrd, M. Cacioppo, A. Cabrera *et al.*, *Bias-preserving and error-detectable entangling operations in a superconducting dual-rail system*, [arXiv:2503.10935](https://arxiv.org/abs/2503.10935).
- [61] A. Quinn, G. Gregory, I. Moore, S. Brudney, J. Metzner, E. Ritchie, J. O'Reilly, D. Wineland, and D. Allcock, *High-fidelity entanglement of metastable trapped-ion qubits with integrated erasure conversion*, [arXiv:2411.12727](https://arxiv.org/abs/2411.12727).
- [62] T. M. Stace and S. D. Barrett, *Error correction and degeneracy in surface codes suffering loss*, *Phys. Rev. A* **81**, 022317 (2010).
- [63] J.-P. Tillich and G. Zémor, *Quantum LDPC codes with positive rate and minimum distance proportional to the square root of the blocklength*, *IEEE Trans. Inf. Theory* **60**, 1193 (2014).
- [64] P. Panteleev and G. Kalachev, *Degenerate quantum LDPC codes with good finite length performance*, *Quantum* **5**, 585 (2021).
- [65] A. Leverrier and G. Zémor, *Quantum Tanner codes*, [arXiv:2202.13641](https://arxiv.org/abs/2202.13641).
- [66] P. Panteleev and G. Kalachev, *Quantum LDPC codes with almost linear minimum distance*, *IEEE Trans. Inf. Theory* **68**, 213 (2022).
- [67] P. Panteleev and G. Kalachev, *Asymptotically good quantum and locally testable classical LDPC codes*, in *Proceedings of the 54th Annual ACM SIGACT Symposium Theory Computing*, STOC 2022 (Association for Computing Machinery, New York, 2022), pp. 375–388.
- [68] N. P. Breuckmann and J. N. Eberhardt, *Quantum low-density parity-check codes*, *PRX Quantum* **2**, 040101 (2021).
- [69] Q. Xu, J. P. Bonilla Ataides, C. A. Pattison, N. Raveendran, D. Bluvstein, J. Wurtz, B. Vasić, M. D. Lukin, L. Jiang, and H. Zhou, *Constant-overhead fault-tolerant quantum computation with reconfigurable atom arrays*, *Nat. Phys.* **20**, 1084 (2024).
- [70] Q. Xu, H. Zhou, G. Zheng, D. Bluvstein, J. Ataides, M. D. Lukin, and L. Jiang, *Fast and parallelizable logical computation with homological product codes*, [arXiv:2407.18490](https://arxiv.org/abs/2407.18490).
- [71] J. Ataides, H. Zhou, Q. Xu, G. Baranes, B. Li, M. D. Lukin, and L. Jiang, *Constant-overhead fault-tolerant Bell-pair distillation using high-rate codes*, [arXiv:2502.09542](https://arxiv.org/abs/2502.09542).
- [72] C.-C. Yu, Y.-H. Deng, M.-C. Chen, C.-Y. Lu, and J.-W. Pan, *Locating Rydberg decay error in SWAP-LRU*, [arXiv:2503.01649](https://arxiv.org/abs/2503.01649).
- [73] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, *Surface codes: Towards practical large-scale quantum computation*, *Phys. Rev. A* **86**, 032324 (2012).

- [74] A. M. Steane, *Active stabilization, quantum computation, and quantum state synthesis*, [Phys. Rev. Lett. **78**, 2252 \(1997\)](#).
- [75] E. Knill, *Quantum computing with realistically noisy devices*, [Nature \(London\) **434**, 39 \(2005\)](#).
- [76] K. Sahay and B. J. Brown, *Decoder for the triangular color code by matching on a Möbius strip*, [PRX Quantum **3**, 010310 \(2022\)](#).
- [77] C. Gidney, M. Newman, P. Brooks, and C. Jones, *Yoked surface codes*, [arXiv:2312.04522](#).
- [78] S. C. Smith, B. J. Brown, and S. D. Bartlett, *Mitigating errors in logical qubits*, [Commun. Phys. **7**, 386 \(2024\)](#).
- [79] S. Huang and K. R. Brown, *Between Shor and Steane: A unifying construction for measuring error syndromes*, [Phys. Rev. Lett. **127**, 090505 \(2021\)](#).