

Surface Code with Imperfect Erasure Checks

Kathleen Chang^{1,2,*}, Shraddha Singh^{1,3,2}, Jahan Claes^{1,3,2}, Kaavya Sahay^{1,3,2}, James Teoh^{1,4}, and Shruti Puri^{3,2}

¹*Yale University, Department of Physics, New Haven, Connecticut 06520, USA*

²*Yale Quantum Institute, Yale University, New Haven, Connecticut 06511, USA*

³*Yale University, Department of Applied Physics, New Haven, Connecticut 06520, USA*

⁴*Quantum Circuits, Inc., 25 Science Park, New Haven, Connecticut 06511, USA*



(Received 16 September 2024; revised 29 August 2025; accepted 8 October 2025; published 4 December 2025)

Recently, significant effort has been devoted towards designing erasure qubits in which dominant physical noise excites detectable leakage states that can be returned to the qubit subspace. Interest in these erasure qubits has been driven by studies showing that the requirements for fault-tolerant quantum error correction are significantly relaxed when noise in every gate operation is dominated by erasures. However, these studies assume perfectly accurate erasure checks after every gate operation which generally come with undesirable time and hardware overhead costs. In this work, we investigate the consequences of using an imperfect but overhead-efficient erasure check for fault-tolerant quantum error correction with the surface code. We show that, under physically reasonable assumptions on the imperfect erasure checks, the threshold error rate is still at least over twice that for Pauli noise. We also study the impact of imperfect erasure checks on the effective error distance and find that it degrades the effective distance under a general error model in which a qubit suffers from depolarizing noise when interacting with a leaked qubit. We then identify a more restrictive but realistic noise model for a qubit that interacts with a leaked qubit, under which the effective error distance is twice that for Pauli noise. We apply our analysis to recently proposed superconducting dual-rail erasure qubits and show that achieving good-performance surface code quantum memories with relaxed system requirements is possible.

DOI: [10.1103/d1v7-nctj](https://doi.org/10.1103/d1v7-nctj)

I. INTRODUCTION

Fault-tolerant quantum error-correcting codes promise scalable quantum computation as long as the underlying hardware is not too noisy. The amount of error suppression with a given code depends on the nature of noise in the underlying hardware [1–19]. For example, consider a *Pauli noise channel* and an *erasure channel*. Under Pauli noise, a qubit may undergo a bit-flip or phase-flip unbeknownst to the user. In contrast, under erasures, the user receives a flag when a specific qubit suffers an error. It has been shown that quantum error-correction (QEC) protocols in which every erroneous gate is flagged have larger effective distance and threshold compared to when the gates are subject to Pauli noise [15–19]. Consequently, theoretical and experimental efforts are being directed towards designing

erasure qubits in which the dominant hardware noise in the system is converted into erasures [17–27].

In the leading proposals for erasure qubits, quantum operations are designed so that the dominant noise takes the qubit to states out of the computational subspace or *leakage states*. The population in these states is detected via its effect on the state of an auxiliary mode. This detection of the leaked population, referred to as an *erasure check*, heralds the erroneous qubit, converting the dominant noise to erasures. A particularly interesting solid-state erasure qubit platform is the recently proposed superconducting dual-rail (DR) qubit, which has the advantage of relatively fast microwave controls with the existing circuit-QED toolbox [19,20,24–27]. While our results are more generally applicable, we tailor our investigation to regimes relevant to the DR qubit in this work. In this platform, the qubit is encoded in the single-excitation manifold of two microwave cavities or two transmons. Additional transmons are required to implement erasure checks. In the case of cavity-based dual-rails, the transmons are also used to implement two-qubit gate operations [20,28].

Unfortunately, in practice, the benefits of erasure qubits are limited due to the time cost of erasure checks. For

*Contact author: katie.chang@yale.edu

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

example, dual-rail erasure checks are expected to take as long to implement as two-qubit gates. Thus, by checking for erasures after every noisy gate, the time cost of one error-correction cycle will, at the very least, double, which is undesirable. Furthermore, new erasures are introduced during the time of these erasure checks, increasing the error rate per surface code cycle, thereby reducing the efficacy of the code.

An important benefit of erasure qubits emerges from the fact that the *effective error distance*, d_{eff} , of codes implemented with qubits suffering from independently and identically distributed (i.i.d.) erasure errors is double that than for qubits that suffer from i.i.d. Pauli noise. Here, d_{eff} is the subthreshold logical error rate scaling with physical error rate p , and the doubling of d_{eff} results in a substantial reduction in the number of erasure qubits required to reach a target logical error rate compared to qubits with Pauli noise. However, in practice, this benefit may also be limited by the hardware cost of erasure qubits. For example, consider the transmon-based dual-rail erasure qubit in which each qubit is composed of two transmons and requires an additional transmon for erasure checks, and contrast it with the conventional transmon qubit, which suffers from Pauli noise. In this case, d_{eff} for a code increases only by 15% when it is implemented using dual-rail transmons compared to when it is implemented by conventional transmons. When also considering

the number of transmons used to implement tunable couplers, there is only a 55% increase in d_{eff} , compared to the naive expectation of a 100% increase one would expect by introducing erasures (see Appendix D). The overhead advantage in terms of performance is dramatically reduced when using them as dual-rail erasure qubits.

From the above discussion, we see that it is beneficial to reduce the complexity of erasure checks for dual-rail qubits for practical advantage—for example, by designing faster erasure checks or using the recently proposed two-qubit gate operations with built-in erasure checks for cavity dual-rails [20,28]. The latter approach is attractive because it uses fewer erasure checks and thus fewer transmons, and does not significantly increase the code cycle length, thereby minimizing the accumulation of noise. However, these alternative strategies come at the cost of a decrease in the quality of erasure checks. For example, the false negative rates may increase or the erasure spatial resolution may decrease, i.e., the exact location of the leaked qubit becomes unknown. We describe these built-in checks in detail, including how these entangling gates are operated, their erasure-detection properties, and hardware requirements in Appendices A, B, and C. Consequently, in this work, we study the effect of these two imperfections in the erasure checks on the performance of the surface code when used as a quantum memory. We find that, under reasonable assumptions about noise in the

TABLE I. We evaluate thresholds and effective error distances for four models of leakage-induced Pauli errors and imperfect erasure checks. Details of these noise models are described in Sec. III. We estimate these quantities at experimentally motivated parameters for the erasure fraction R_e , and temporal resolution η . The quantity η quantifies the fraction of two-qubit gate erasures which are detected immediately after the gate before they propagate to other qubits via subsequent gate operations. Recent superconducting dual-rail experiments have observed R_e ranging from 0.96 to 0.98 and η from 0.95 to 0.99 [24,25]. In the last two columns, we connect these erasure check models to existing superconducting dual-rail architectures. To provide a sense of the hardware costs, we quote the estimated number of transmonlike elements and read-out lines required to build a 3×3 rotated code. This transmon count includes data qubits, ancilla qubits, and their tunable couplers (see Appendix D). The low transmon-element count for cavity-based built-in erasure checks is because, in principle, it only requires one extra transmon per ancilla dual-rail qubit to carry out gates and erasure checks.

Temporal resolution $\eta = 0.986$						
Spatial resolution	Leakage-induced Pauli	R_e	Threshold (%)	d_{eff} (3×3)	Dual-rail architecture and No. of transmons (3×3)	
					Transmon based	Cavity based
Imperfect	General	0.98	4.16 ± 0.03	1.99 ± 0.02		
		0.99	4.49 ± 0.02	2.03 ± 0.03		
Perfect	General	0.98	4.17 ± 0.03	2.00 ± 0.02		
		0.99	4.55 ± 0.02	2.00 ± 0.02		
Imperfect	Tailored	0.98	4.23 ± 0.03	2.38 ± 0.03		49 transmons & 8 readout lines for QND checks built into gates [20,28]
		0.99	4.47 ± 0.03	2.60 ± 0.05		
Perfect	Tailored	0.98	5.47 ± 0.03	2.36 ± 0.04	75 (58) transmons if three (two) transmons per dual-rail & 17 readout lines for explicit QND checks with false negatives [19,25]	58 transmons & 17 readout lines for explicit QND checks with false negatives [20,24,27]
		0.99	6.01 ± 0.03	2.54 ± 0.02		

superconducting dual-rail platform, it is still possible to achieve QEC advantage with imperfect erasure checks.

Our paper is organized as follows. In Sec. III, we introduce the noise model for imperfect erasure detection. For different erasure detection accuracies and noise models, we quote thresholds and effective error distances for physically relevant parameters in superconducting erasure platforms in Table I and summarize our main results. In Sec. IV, we describe how the decoder is calibrated for checks with imperfect spatial and temporal resolution. Our main results are presented in Sec. V A, which examines the impact of imperfect erasure checks on the effective error distance, and Sec. V B, which discusses the impact on the threshold. Note that the numerical analysis in this work is performed with unrotated surface codes. However, the threshold results should also be valid for the rotated codes, and we also present theoretical analysis for the logical error rate scaling of the rotated code (see Appendix F).

II. PRELIMINARIES

A. Dual-rail qubits

The dual-rail qubit, first proposed in the context of linear quantum optics [29], encodes qubits in two bosonic modes. Explicitly, the computational states of the dual rail are the single-excitation manifold spanned by $\{|01\rangle, |10\rangle\}$. These modes may be encoded in orthogonal optical polarization states [30], two superconducting cavities [20,31,32], or nonlinear modes such as two transmons [25,33]. Under loss errors, this encoding goes to the $|00\rangle$ state whose population can be monitored in the circuit QED platforms by measuring the joint parity of excitations in both modes [20,24–26]. With this detection available, the dual-rail qubit can operate as an erasure qubit.

B. Surface codes

The surface code is well established as a leading choice of encoding to build a scalable, fault-tolerant quantum computer. This is due to its relatively relaxed hardware requirements of only nearest-neighbor connectivity, fast, accurate decoders [34,35], and high thresholds to circuit-level Pauli noise and erasure noise [36].

In the main text we consider surface codes defined on a square lattice (Fig. 1), where physical qubits are placed on the vertices of the diamonds and triangles. The surface code encodes one logical qubit, where the X - (Z -)logical operator is a product of Pauli- X (Pauli- Z) operators that act on the qubits on the top and right boundaries of the code, respectively. In fact, the distance to an X_L or Z_L is d , the length of one side of the lattice. Each diamond or triangle is a stabilizer that is a weight-4 or weight-3 product of Pauli operators with support on the vertices of each face. These stabilizers are a product of X 's (light faces) or Z 's (dark faces). Each stabilizer has an ancilla qubit in the center

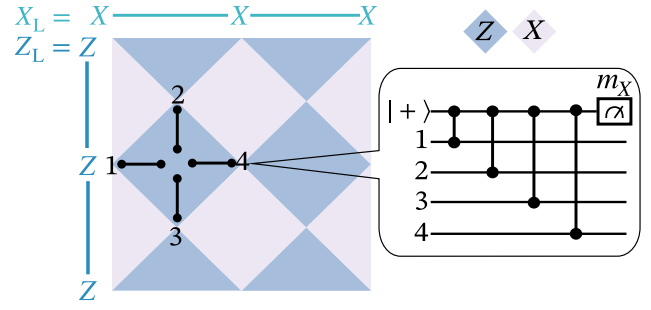


FIG. 1. The CSS surface code. Dark (light) plaquettes are Z (X) stabilizers with support on data qubits at the vertices. The stabilizer is measured by applying two qubit gates on a $|+\rangle$ -initialized ancilla (not shown) with neighboring data qubits 1–4 in order. This order is shown for a representative Z stabilizer and is tiled on every stabilizer.

to mediate the stabilizer measurement on the four or three adjacent data qubits. An example of a circuit for measuring a stabilizer is given on the right of Fig. 1, where the ancilla qubit is on the top rail and measures Z on four adjacent physical qubits 1–4. The same circuit applies for measuring X stabilizers, but instead uses controlled- X (CX) gates with the target on the data qubits. This ordering of entangling gates ensures that the correct X and Z stabilizers can be measured simultaneously in four timesteps while ensuring that Pauli errors from the ancilla do not propagate and reduce the distance of the code [37]. In Appendix F we also describe the rotated surface code and the appropriate way to measure stabilizers there [37].

The threshold, defined as the physical error rate below which distance scaling presents a QEC advantage, depends on assumptions about the underlying noise model. The surface code has a threshold for pure Pauli noise of about 1% when depolarizing noise occurs after every entangling gate. Ideal assumptions of heralded noise with only dual-rail loss errors and perfect erasure detection after every noisy gate lead to even larger thresholds ($5.56 \pm 0.03\%$) compared to unheralded, Pauli noise.

1. Previous studies of erasure thresholds and advantages

Previous studies on the advantages of erasure qubits on the threshold differ in a couple of key assumptions, including that detection occurs in the circuit often enough, and that erasure detection always resets any detectable leakage back into the computational space [19,25,38]. We note that the former assumption is studied in Ref. [39]. However, of particular note is that the assumption of unconditional reset back into the computational space after every check, such as through engineered dissipation, has yet to be proposed or demonstrated in erasure experimental works across multiple erasure platforms [17,20,24–26]. As a result, a critical concern in the realistic context of imperfect erasure detection is detection with false negatives. This implies that

leakage is (temporarily) undetected while interacting with other qubits, leading to the spread of correlated errors on these qubits. We study whether leakage-induced correlated errors will eclipse the threshold advantage of introducing erasure qubits, that is, whether the threshold remains above the 1% Pauli error baseline for surface codes.

Another factor not previously as emphasized in the comparison between erasure versus Pauli noise is the additional hardware complexity to realize an erasure architecture (Table I). For example, a dual-rail transmon architecture requires twice the number of transmons compared to a direct transmon encoding. Even more transmons are actually required to realize erasure checks on each erasure qubit [25]. In the context of this punishing space and time overhead implementing an erasure platform, we find that we can trade off accuracy for efficiency in our erasure checks, cementing a convincing argument for the maintenance of overhead advantage in erasure qubits.

III. IMPERFECT ERASURE DETECTION FRAMEWORK AND SUMMARY OF MAIN RESULTS

In this section, we introduce the imperfections we consider in erasure checks. We consider four models of imperfect erasure checks listed in Table I in a $d \times d$ CSS surface code as depicted in Fig. 1. Consider the standard surface code cycle [40] in which each stabilizer is measured by interacting an ancilla initialized in $|+\rangle$ with the data qubits supported by the stabilizer through CX or CZ gates following a specific order as shown in Fig. 1(a). Subsequently, the ancilla qubit is measured on the X basis. To simplify the analysis, we only consider erroneous two-qubit gates. We do not consider state preparation, measurement, and idling errors because, in principle, they can be combined into the two-qubit gate errors. Given probability p of an error (leakage and Pauli) at a two-qubit gate, we set the probability of leakage during a two-qubit gate as $p_e = R_e p$ and the mutually exclusive probability of a Pauli error as $p_p = (1 - R_e)p$. Here $0 \leq R_e \leq 1$ is the fraction of leakage errors, or, equivalently, $1 - R_e$ is the fraction of Pauli errors in the gate. The Pauli error, drawn uniformly at random from $\mathcal{E}_p = \{I, X, Y, Z\}^{\otimes 2} / \{I \otimes I\}$ with probability $p_p/15$, is applied after a two-qubit gate in simulations. We assume that after every two-qubit gate, an erasure check provides a *flag* to the user indicating if a qubit has leaked. Importantly, we assume that the flag itself is imperfect, which we describe shortly. This model of (imperfectly) detectable leakage is different from prior studies in which leakage cannot be directly detected and is kept in check with leakage-reduction units or is periodically brought back to the computational subspace and converted to Pauli errors [41–48]

Before we define these imperfect flags, we first consider what happens when a leaked qubit interacts with another

unleaked qubit before it is detected and reset. We assume that a leaked qubit can at most induce a Pauli error on the qubit that it is interacting with [42,44,45,49,50]. The case in which this leakage-induced Pauli error is drawn from $\mathcal{E}_{\text{gen.}} = \{I, X, Y, Z\}$ with equal probability will be referred to as the *general Pauli model*. We also analyze a *tailored Pauli model* in which the Pauli error induced by the leaked control qubit on the target qubit in a CZ (CX) gate is drawn from $\mathcal{E}_{\text{tail.}} = \{I, Z\}$ ($\mathcal{E}_{\text{tail.}} = \{I, X\}$) with equal probability. In the tailored Pauli model, when the target qubit is leaked, the leakage-induced Pauli error on the control qubit is a Pauli error drawn from $\{I, Z\}$ with equal probability irrespective of whether a CZ or CX gate was applied. The tailored Pauli model is different from a biased noise model studied in Refs. [1,2,4,6,7,9,12,14,51–57] in the sense that both Z and X errors can be induced instead of just Z errors. They are conceptually similar in that decoding for biased noise and the tailored Pauli model use information about the likely errors from the noise model to achieve higher thresholds compared to unbiased noise and the general Pauli model, respectively.

The tailored Pauli model was previously studied in the context of leakage reduction units for trapped ions [45]. Furthermore, all of the dual-rail gates proposed in Ref. [20], including the gate with built-in erasure checks [28], are described by the tailored Pauli model. This is because in the CZ gate, if either one of the qubits undergoes a leakage error, the other qubit may undergo a Z -rotation. This noise channel reduces to a dephasing channel under Pauli twirl approximation (PTA) achieved in practice by inserting random single-qubit Pauli gates before and after the CZ gate. The CX gate is implemented by conjugating the target of the CZ gate with Hadamard gates, and, hence, its leakage-induced noise channel on the target is the bit-flip channel.

Next, we consider the following properties for the erasure checks.

Check with perfect spatial resolution, in which a flag received by the user after a two-qubit gate indicates exactly which qubit involved in the gate was leaked. This model is appropriate when each qubit has its own independent erasure check. In this case, the user resets the flagged qubit in a computational basis state, after which the error-correction cycle continues from where it was left off. This type of noise on the flagged qubit is accounted for in simulations as a Pauli error drawn from $\mathcal{E}_L = \{I, X, Y, Z\}$ with equal probability [17,58].

Check with imperfect spatial resolution, in which the flag only indicates that there was leakage during the gate but does not indicate which of the two qubits leaked. Imperfect spatial resolution is a property of gates with built-in checks. In this case, the user resets both the qubits participating in the gate and the resulting noise is accounted for in simulations as a Pauli error drawn from $\{I, X, Y, Z\}^{\otimes 2}$ with equal probability [17,58]. This noise

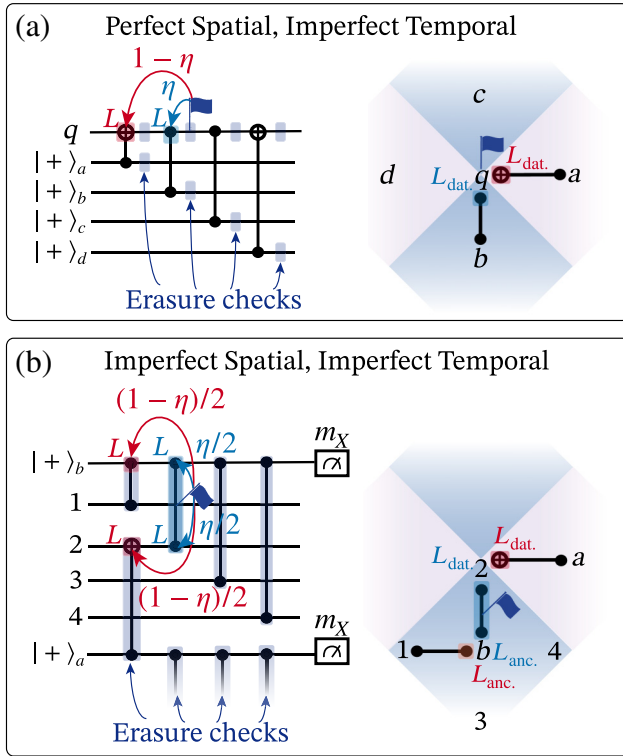


FIG. 2. Given checks with (a) perfect spatial and imperfect temporal resolution or (b) imperfect spatial and imperfect temporal resolution, we show an example of a flag received by the user. Given this flag, we show possible locations of leakage, shown on the left for each figure, with probabilities dependent on η . Blue denotes leakage events detected on time, while red denotes leakage events detected one time step later. The same potential leakage locations are labeled on the surface code plaquettes, shown on the right in each figure, where the subscripts on L denote if the leakage is on a data or ancilla qubit.

model is motivated by the noise properties of entangling gates with built-in erasure checks proposed for superconducting dual-rail qubits [20,28] (see Appendices A and B).

Check with imperfect temporal resolution, in which a flag received by the user after a gate could have resulted from a leakage in that gate with probability η or from a leakage in an immediately preceding gate on the qubits with probability $1 - \eta$. We illustrate an example in Fig. 2(a), in which we assume that checks have perfect spatial resolution and imperfect temporal resolution. Here, receiving a flag on the data qubit q after the second gate in the circuit means that leakage could have happened on the data qubit during the second gate (in blue) with probability η , or on the data qubit at the first gate (in red) with probability $1 - \eta$. To clarify, the right-hand side of Fig. 2(a) shows the same scenario as the circuit, but mapped onto the code, making it visually evident that we know the spatial location but not the time of leakage.

In Fig. 2(b), we consider a more complicated scenario with imperfect spatial resolution and imperfect temporal resolution. These checks cannot resolve which qubit has leaked at each gate, so the user assumes that the flag could have resulted from the ancilla or the data qubit with equal probability. As a result, both the ancilla qubit b and data qubit 2 have probability $\eta/2$ of having leaked during the flagged gate, and probability $(1 - \eta)/2$ of having leaked during the previous gate. The right-hand side of Fig. 2(b) depicts this same scenario on the code to emphasize that both the spatial location and the time of the leakage are now unknown.

Checks with imperfect temporal resolution are motivated by the noise properties of superconducting dual-rail qubits. For example, the QND built-in checks proposed in Refs. [20,28] fail to detect a certain fraction, say $1 - \eta$, of leakage, but this missed leakage is caught by the check of the next gate. This, however, means that if a gate is flagged then, with probability $1 - \eta$, the leakage actually occurred in the qubits during the gates that acted on them immediately before the flagged gate. For the built-in checks, η can be as high as 0.986 for reasonable system parameters (see Appendix A). As another example, consider false negatives in erasure checks occurring with probability $1 - \eta$. Since the erasure checks are QND and the false negative rate is small, both physically realistic assumptions, a missed leakage at one check is almost certainly caught by the next one. This implies that, with probability $1 - \eta$, the flag raised at one check came due to leakage missed by the preceding check. Our model neglects the possibility of two consecutive false negatives, which is negligible for physically reasonable false negative rates. In a recent transmon-based dual-rail experimental demonstration, an erasure probability of 2.9% was accompanied by a 1.5% false negative probability, giving $\eta = 98.6\%$ and negligible double false negatives 2.2×10^{-4} [25]. Similarly, a cavity-based implementation quotes an erasure probability of 2.92% with a false negative probability of 3.7% [27]. These two implementations achieve false negative rates of the same order of magnitude. We calculate experimentally motivated η values for the built-in checks used in Table I in Appendix A. Additionally, we would like to remark that the physics of the built-in check in cavity-based dual-rail qubits motivates studying leakage that is detected at most one time step late (Appendices B and C). However, we also investigate cases where detection is not guaranteed to happen after a fixed number of gates in Sec. VB below. Specifically, we consider a noise model in which each check has a finite probability p_m of giving a false negative result.

False positive erasure checks. Since qubits are reset after an erasure flag, false positive erasure checks simply increase the effective erasure rate. Specifically, given a false positive rate $p_{f.p.}$, $p \rightarrow p_e + p_{f.p.} + p_p$, $R_e \rightarrow (p_e + p_{f.p.})/(p_e + p_{f.p.} + p_p)$, and $\eta \rightarrow \eta \cdot p_e/(p_e + p_{f.p.})$.

As a result, all subsequent results in Sec. VB below can be reinterpreted with an appropriate false positive rate. Importantly, false positives do not introduce correlated errors, and therefore we do not expect the effective distance nor the threshold to change with an increasing false positive rate. This is not a given for false negatives, which we capture through temporally imperfect checks. The most detrimental effect of false negatives is when an attempted erasure check does not unconditionally reset the erased qubit back into the computational space, which was not reflected in previous studies [19,38].

Remark on the low- η regime. Although the regime $\eta \approx 1$ is motivated by low false negative rates and the built-in check for cavity-based dual-rail qubits, it is also interesting to consider low η values for other physically realistic settings. For example, a scenario where the threshold at $\eta = 0$ is relevant is when the latency of classical real-time processing and feedforward for an erasure detection measurement is comparable to the duration of a two-qubit gate. In this scenario, one may decide to proceed with the next two-qubit gate before knowing the outcome of the erasure check to reduce the time overhead of the syndrome extraction circuit. The lack of immediate erasure information means that one cannot prematurely reset the leaked qubit before the next gate, and thus the leaked qubit propagates for at least two time steps before it is reset to the computational space. Furthermore, although the built-in check for cavity-based dual-rail qubits motivates a particular fraction of η , it is possible that new proposals for hardware-efficient erasure checks motivates the inspection of lower values of η .

Our numerical results on thresholds and the effective error distance for practical noise models and values of η , R_e are summarized in Table I. Although, in general, temporally imperfect erasure checks ($\eta < 1$) may reduce the effective distance of the code, we find that, for the tailored Pauli model, which is realistic for superconducting dual-rail qubits, the distance is not reduced. Additionally, for realistic parameters in Table I, the threshold for built-in erasure checks is $4.23 \pm 0.02\%$, which is well above the Pauli threshold $\sim 1\%$ and only slightly worse than the perfect erasure detection case, $5.56 \pm 0.03\%$. Therefore, with a tailored Pauli model, we find that it is possible to achieve a higher threshold and higher effective distance for a $d \times d$ surface code compared to qubits with conventional Pauli noise. Importantly, both built-in spatially imperfect checks and explicit spatially perfect checks achieve these improvements. Recall that built-in checks with cavity-based DRs do not lead to an increase in the length of the error-correction cycle and require fewer transmons. Specifically, the number of transmons required to achieve the effective distance of d_{eff} with built-in checks is 29% of the number of transmons required to reach the same effective distance if they are used as conventional qubits with Pauli noise (Appendix D). This is because

the built-in checks only require one transmon per ancilla dual-rail qubit (Appendix A). Thus, we find that the cavity-based DRs with built-in checks are highly desirable for practical hardware-efficient QEC.

In the following section, we describe in more detail how the noise model affects the performance of the code and how we implement the weighted union-find (UF) decoder.

IV. NOISE AND DECODING FOR IMPERFECT ERASURES

Irrespective of the spatial resolution, the main impact of imperfect temporal resolution is that a single leakage error can spread to three qubits. To see this, consider the examples illustrated in Figs. 3(a), 3(b). In these examples, a qubit labeled A leaks during a two-qubit gate with another qubit labeled B at time step T_k with probability p_e . Then a Pauli error is introduced on qubit B according to the general $\mathcal{E}_{\text{gen.}}$ or tailored $\mathcal{E}_{\text{tail.}}$ noise models. With probability η , this leakage gets detected by the erasure check corresponding to the gate between A and B at T_k , but, with probability $1 - \eta$, it will get detected by the erasure check corresponding to the gate applied to A at time T_{k+1} . Let us label the qubit that A interacts with at T_{k+1} as C . Since C has now interacted with a leaked qubit, it also suffers from a Pauli error. Thus, a single error on A has led to an error on two more qubits B and C . This can lead to a reduction in the effective distance. This effective distance reduction occurs when B and C are along a minimum-weight logical operator and both receive an error of the same type as the action of the logical operator on qubits B and C .

A. Noise and the effective distance

We define the effective error distance d_{eff} to be the exponent of the physical error rate p in the approximate expression for the logical error rate $p_L \propto p^{d_{\text{eff}}}$ deep in the subthreshold regime of physical error rates. The effective distance can be roughly understood to be the log probability of the most likely nontrivial logical operator [10], which is often, in practice, the minimum number of errors needed to create a logical error. As shown in Fig. 3(a), delayed detection can result in two Z errors on B and C , which lie in the support of a Z -type logical operator. Since one imperfect erasure corrupts two data qubits with a Z error along the boundary, it takes roughly half the number of physical errors to create the most probable Z_L error. In fact, we expect the Z_L effective distance to reduce by half and approach $(d + 1)/2$ for odd codes, which is the same as the effective distance for pure Pauli noise.

In contrast to the general Pauli model, the effective distance to a Z_L error for the tailored Pauli model should remain at d . An illustrative example is given in Fig. 3(b), in which data qubits B and C have errors chosen from $\mathcal{E}_{\text{tail.}} = \{I, X\}$, which importantly does not contain Z . Therefore, it

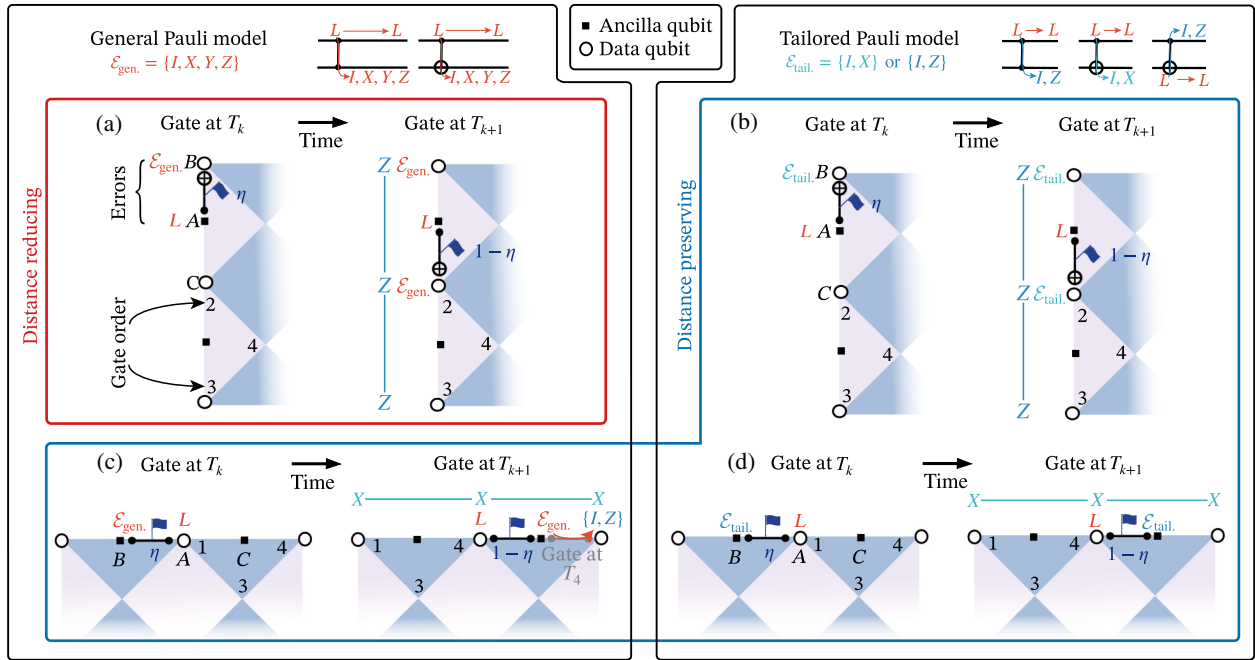


FIG. 3. (a) General Pauli model. Data qubits (open circles) and ancilla qubits (squares) along Z_L . The numbers indicate the order of two-qubit gates, which apply to every stabilizer. Here, we examine a leakage error on the ancilla qubit A during a CX gate with data qubit B at time T_k . With probability $1 - \eta$, this leakage is flagged at time T_{k+1} . Due to this delayed detection, two data qubits along Z_L , B and C , have errors chosen from $\mathcal{E}_{\text{gen.}} = \{I, X, Y, Z\}$, thereby reducing the effective distance to a Z_L error. In (b), we consider the same leakage location, but now under the tailored Pauli model. Here, B and C receive faults from $\mathcal{E}_{\text{tail.}} = \{I, X\}$. These errors do not decrease the effective distance to a Z_L error. (c) Despite a general Pauli model, the X_L error distance is preserved due to the scheduling of gates, in contrast to the reduced Z_L error distance. Although we show one example of data qubit leakage in (c), no other leakage location or time will lead to effective X_L distance reduction. In (d), we show that the effective distance to a X_L error is also not reduced for the tailored Pauli model.

is impossible for this singular leakage error to lead to multiple Z errors along Z_L . While we show a specific example of ancilla qubit leakage in Fig. 3(b), one may repeat this exercise with leakage at other circuit locations and qubits and find that no possible leakage event results in effective distance reduction to a Z_L error for the tailored Pauli model.

It turns out that effective distance reduction does not depend on whether checks are spatially perfect or imperfect. To elaborate, spatially perfect checks only reinitialize qubit A upon detection, whereas spatially imperfect checks reinitialize both qubits A and C . This means that if checks are spatially imperfect, qubit C will have errors from $\mathcal{E}_L = \{I, X, Y, Z\}$, but errors on B are unchanged regardless of spatial resolution. This fact, and the requirement that the effective distance needs errors on both B and C to have support on the same logical operator, means that the effective distance is not dependent on spatial resolution.

Finally, in Figs. 3(c), 3(d), we illustrate that due to the scheduling of gates along the X_L boundary, one delayed detection does not result in two data qubit errors along X_L . Therefore, the effective error distance to an X_L error is not reduced no matter the noise model. For example, let

us consider the general Pauli model in Fig. 3(c). Leakage on a data qubit A depolarizes two ancilla qubits, B and C . As per the usual syndrome measurement procedure, ancilla qubit B is reset before the error can propagate, so let us focus on qubit C . Depolarizing noise on qubit C may propagate to a Z error on the right data qubit; however, this Z error cannot reduce the effective distance to an X_L error. One may go through a similar exercise for leakage on the ancilla qubits, and find that X errors will never appear on two qubits along the X_L boundary.

We also study the effective distance of the rotated surface code in Appendix F. Unlike the unrotated code, the rotated code's effective distance is reduced along both X_L and Z_L under the general Pauli model. However, under the tailored Pauli model, the effective error distance to both an X_L and Z_L error is preserved because the pattern of gates in the measurement cycle of the rotated code ensures that the leakage-induced correlated X (Z) errors occur on qubits which do not lie along X_L (Z_L).

Next, we consider the correlated errors that arise due to imperfect temporal resolution of checks with both perfect and imperfect spatial resolution. We also describe how the decoder is calibrated for these correlated errors.

B. Noise and decoding: perfect spatial, imperfect temporal resolution

For all numerical results in this work, we use Monte Carlo simulations of stochastic circuit-level noise and a weighted UF decoder, which has been shown to be effective when both Pauli and erasure noise are present [17,18,59–61]. The decoder relies on constructing a graph with a vertex at each spacetime syndrome location. A weighted edge is placed between vertices that can be excited by a single circuit fault, and the weight reflects the probability of such a fault. The higher the probability of the fault, the lower the edge weight.

As we saw previously, due to imperfect temporal resolution, a single fault can cause errors on up to three qubits whose locations are exactly known when spatial resolution is perfect. This information is used to calibrate the decoder. Specifically, if a qubit is flagged at time T_{k+1} then this qubit is reset. Irrespective of when the leakage happened, the qubit that interacted with the leaked one at T_{k+1} suffers leakage-induced Pauli noise. Thus, the faults in the gate at T_{k+1} correspond to errors on the leaked qubit drawn from $\mathcal{E}_L$ and those on the qubit that it interacted with are drawn from $\mathcal{E}_{\text{gen.}}$ or $\mathcal{E}_{\text{tail.}}$. These circuit faults are used to define the corresponding edges of the decoding graph in the standard way. Importantly, the decoder needs to account for the fact that, with probability $1 - \eta$, the leakage in A could have occurred in the gate that was applied to it at T_k . Thus, there could have been additional possible faults corresponding to errors on the leaked qubit drawn from $\mathcal{E}_L$ with probability $(1 - \eta)/4$ and those on the qubit that A interacted with at T_k , drawn from $\mathcal{E}_{\text{gen.}}$ or $\mathcal{E}_{\text{tail.}}$ with probability $(1 - \eta)/4$ or $(1 - \eta)/2$, respectively. These probabilities are then used to define the edge weights of the decoding graph in the standard way. If no leakage is detected at T_{k+1} then the possible Pauli faults drawn from $\mathcal{E}_p$ with equal probability $p_p/15$ are accounted for by defining the edges of the decoding graph in the standard way. For details on these weights, see Fig. 14 in Appendix E.

C. Imperfect spatial, imperfect temporal resolution

In the case of imperfect spatial resolution, the flag only indicates an erasure during the gate and thus both qubits, say A and C in Fig. 3, participating in the gate are reset if flagged. If a gate is flagged at T_{k+1} then both qubits participating in the gate are reset, and the gate faults correspond to errors on each qubit from $\mathcal{E}_L$ with equal probability $1/16$. These circuit faults are used to define the corresponding edges of the decoding graph in the standard way. Additionally, the decoder must account for the fact that, with probability $1 - \eta$, A or C could have leaked during a gate that was previously applied at T_k . Thus, there could have been additional possible faults corresponding to errors on A and C , drawn from $\mathcal{E}_L$ with probability $(1 - \eta)/4$ and those on the qubits that they interacted with

at T_k drawn from $\mathcal{E}_{\text{gen.}}$ or $\mathcal{E}_{\text{tail.}}$ with probability $(1 - \eta)/4$ or $(1 - \eta)/2$, respectively. These probability distributions are then used to define the edge weights of the decoding graph in the standard way (Appendix E). Finally, if no leakage is detected at T_{k+1} then the possible Pauli faults drawn from $\mathcal{E}_p$ with equal probability $p_p/15$ are accounted for by defining the edges of the decoding graph in the standard way. For details on these weights, we refer the reader to Fig. 14 in Appendix E.

V. RESULTS

A. Impact on the effective error distance

We numerically confirm that the general and the tailored Pauli models have different consequences on the effective error distance, d_{eff} . In Fig. 4, we plot the effective error distance to a logical error by fixing the code size (3×3) and varying the temporal resolution η , the Pauli model, and the spatial resolution. We assume that physical errors are exclusively erasures ($R_e = 1$); thus, the effective error distance should be 3 when erasure detection is perfect. We extract d_{eff} by plotting the logical error rate for varying physical error rates that are significantly below the threshold ($p \sim 0.1p_{\text{th}}$). At this regime of low p , we approximate $p_L \propto p^{d_{\text{eff}}}$ and plot $\log(p_L)$ versus $\log(p)$. The slope of this line is the effective error distance.

The light blue diamonds in Fig. 4 are numerically extracted effective distances to an X_L error under the general Pauli model with spatially imperfect checks. The dotted lines connecting the diamonds are a guide to the eye. As expected, we note that the effective distance to an X_L error

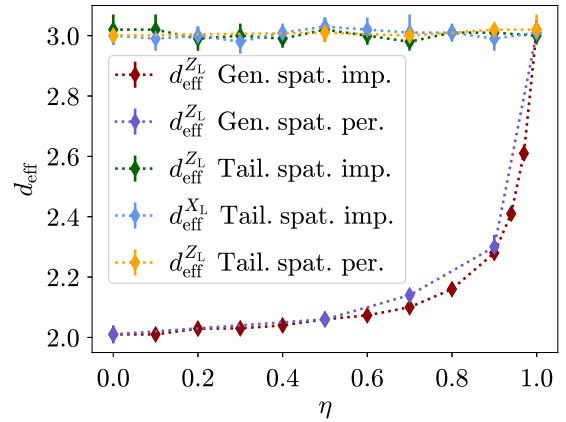


FIG. 4. Effective error distance d_{eff} at $R_e = 1$ with varying temporal resolution of checks (η). Red (and purple) diamonds show d_{eff} to a Z_L error for the general Pauli model with spatially imperfect (and perfect) checks. The dotted lines connecting the diamonds are a guide to the eye. Here d_{eff} to a Z_L error is shown for the tailored Pauli model with spatially perfect checks (orange) and imperfect checks (green). Light blue shows d_{eff} to an X_L error for the tailored Pauli model with imperfect checks.

is not reduced and remains at 3 regardless of the temporal resolution of checks. Since the remaining noise models are more restrictive than the general Pauli model with spatially imperfect checks (e.g., tailored, spatially resolved checks), we do not expect the effective distance to an X_L error in these models to be worse than 3. Furthermore, by virtue of examining a distance 3 code, the effective distance will never be greater than 3.

The remaining points in the figure illustrate the effective distance to a Z_L error. Now, let us examine how the effective error distance changes with different leakage-induced Pauli models. The red curve in Fig. 4 examines the effective error distance to a Z_L error under the general Pauli model with spatially imperfect checks. As predicted in Sec. IV A, the effective distance quickly approaches 2 as the temporal detection accuracy gets worse ($\eta \rightarrow 0$). This is the Pauli noise scaling for a $d = 3$ code ($p_L \propto p^{(d+1)/2} = p^2$). Therefore, with delayed erasure detection and the general Pauli model, the Z_L error rate no longer scales advantageously compared to Pauli errors. We confirm that the Pauli scaling $p^{(d+1)/2}$ is reproduced for larger lattice sizes in Fig. 16 in Appendix G.

Conversely, we observe that in the tailored Pauli case, with spatially imperfect checks (green) and spatially perfect checks (orange), the effective error distance remains at 3 (within error bars) regardless of the spatial resolution of checks or how temporally imperfect the detection is. This means that the logical error rate of both X_L and Z_L remains the scaling for perfect erasures, p^d , for the tailored Pauli model, as predicted in Sec. IV A.

In general, we see that the effective error distance is unaffected by the spatial resolution of checks. As we just observed, the tailored Pauli model has distance d regardless of the spatial resolution of checks. Furthermore, when we examine the effective distance to a Z_L error for the general Pauli model with spatially perfect checks (purple), there is no difference in d_{eff} to spatially imperfect checks (red) within error bars. This means that in terms of the effective error distance, there is no advantage to engineering erasure checks which are spatially resolving. These conclusions hold for our physically motivated noise model and spatial imperfections; however, there may be other models of spatial imperfection that may be harmful to the effective distance. Finally, we numerically show that, for the experimentally relevant tailored Pauli model, the effective distance to an X_L (light blue) or Z_L (green) error does not drop due to delayed erasure detection.

B. Impact on the threshold

Next, we numerically examine the consequences of imperfect detection resolution on the threshold. All thresholds are obtained with lattice sizes $d = 9, 11, 13, 15$, and fitted by using the quadratic expansion method used in Ref. [62].

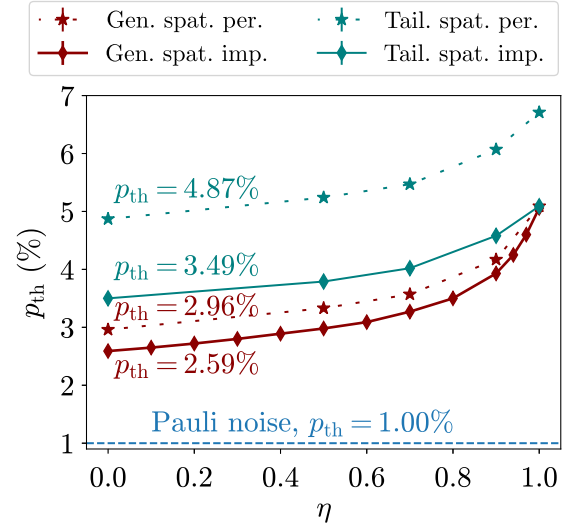


FIG. 5. Threshold for varying temporal resolution (η) at $R_e = 1$ for four noise models: the general Pauli model with perfect (red stars) and imperfect (red diamonds) spatial resolution, and the tailored Pauli model with perfect (teal stars) and imperfect (teal diamonds) spatial resolution. The connecting lines between points are a guide to the eye. The threshold for Pauli noise after every two-qubit gate is marked at the dotted blue line.

We continue to consider the case of only erasure noise ($R_e = 1$) in Fig. 5. For comparison purposes, we plot the threshold for pure Pauli noise, $1.00 \pm 0.01\%$, marked with the dotted blue line. We plot the threshold against η for the general and tailored Pauli models with and without spatially perfect checks. For all four noise models, the threshold drops the quickest in the regime of near-perfect temporal detection ($\eta \sim 1$). Thus, if the detection mechanism is only marginally temporally inaccurate, slightly improving the temporal detection accuracy pays off in threshold.

We observe that the thresholds for the tailored and spatially imperfect, general and spatially perfect, and general and spatially imperfect noise models all approach the threshold for perfect erasures when $\eta = 1$. This is an expected consequence since at exactly $\eta = 1$ all three noise models are equivalent. As an example, for spatially imperfect checks (general and tailored), both qubits participating in the gate are reset when the gate is flagged, and if detection is on time, these are the only faults on the qubits. Furthermore, in the general Pauli model for spatially imperfect checks at $\eta = 1$, the leaked qubit participating in the flagged gate has faults drawn from $\mathcal{E}_{\text{gen.}}$, which contain the same errors as $\mathcal{E}_L$. Therefore, we expect the thresholds to converge to the same value when all detection events are on time.

On the other hand, only the tailored Pauli model with spatially perfect checks does not converge to the same threshold value at $\eta = 1$ as the rest of the noise models.

Given a flagged gate with on-time detection, the flagged qubit has faults chosen from $\mathcal{E}_L$ and the leaked qubit has an error from the set $\mathcal{E}_{\text{tail}}$. Due to the restricted error set of $\mathcal{E}_{\text{tail}}$, we can know if the leaked qubit has a Z or X error depending on the type of gate applied (CX or CZ). This results in an erasure with extra information, where we know not only the location of the error but also what type (X or Z) it is. On the decoding graph, this results in fewer edges that are modified or set to a lower weight. In the language of percolation theory and its relationship to the erasure error threshold, when fewer edges are removed with probability p , the threshold error rate is higher [63,64].

The convergence in thresholds for temporally accurate checks results in interesting consequences. For example, in the general Pauli model, the difference in threshold between spatially imperfect (red diamonds) and perfect (red stars) checks is small in the regime of $\eta \rightarrow 1$. This means that the cost of engineering checks that can resolve which qubit is leaked at every two-qubit gate is hardly going to increase the threshold. In contrast, in the tailored Pauli model case, the difference in threshold between perfect (teal stars) and imperfect (teal diamonds) spatial resolution in this regime is over 1%.

Finally, as late detection is more frequent ($\eta \rightarrow 0$), the threshold approaches $2.59 \pm 0.01\%$, $2.96 \pm 0.02\%$, $3.49 \pm 0.02\%$, and $4.87 \pm 0.05\%$ for the general Pauli model with spatially imperfect and perfect checks, and the tailored Pauli model with spatially imperfect and perfect checks, respectively. In all four noise models, this is well above the Pauli noise threshold of $1.00 \pm 0.01\%$. This means that even when all detection is temporally delayed by one time step and each erasure results in correlated errors, delayed

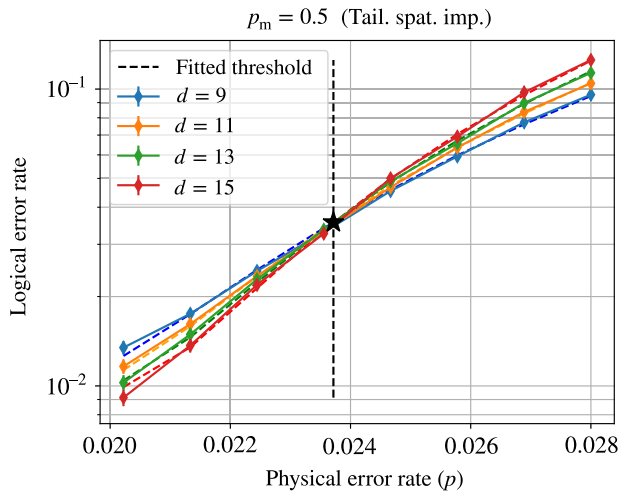


FIG. 6. Tailored Pauli model with spatially imperfect checks in the setting where each gate following an erasure has probability $p_m = 1/2$ to miss detection. Fitting $d = 9, 11, 13, 15$ gives a threshold of $2.37 \pm 0.02\%$.

detection has a higher threshold than Pauli noise. This suggests that the induced correlated errors are bounded enough in spacetime to still yield threshold advantages.

Unbounded propagation of leakage. The physics of the erasure detection in cavity-based dual-rail qubits motivates a model where erasure is detected after at most one time step, as explored in Appendix B. More generally, however, we also analyze a worst-case scenario in which every subsequent check after erasure has a fixed probability p_m to miss erasure. We refer to this model as unbounded propagation of leakage. This means that the average number of time steps until erasure detection is given by $\bar{T}(p_m) = 1/[\sum_g g \cdot p_m(1 - p_m)^{g-1}]$. Therefore, an appropriate p_m analogous to setting $\eta = 0$ in Fig. 5 is $p_m = 1/2$, such that, on average, leakage is also detected after two time steps, $\bar{T}(p_m = 1/2) = 2$. Crucially, in this model, there is a nonzero probability that erasure may

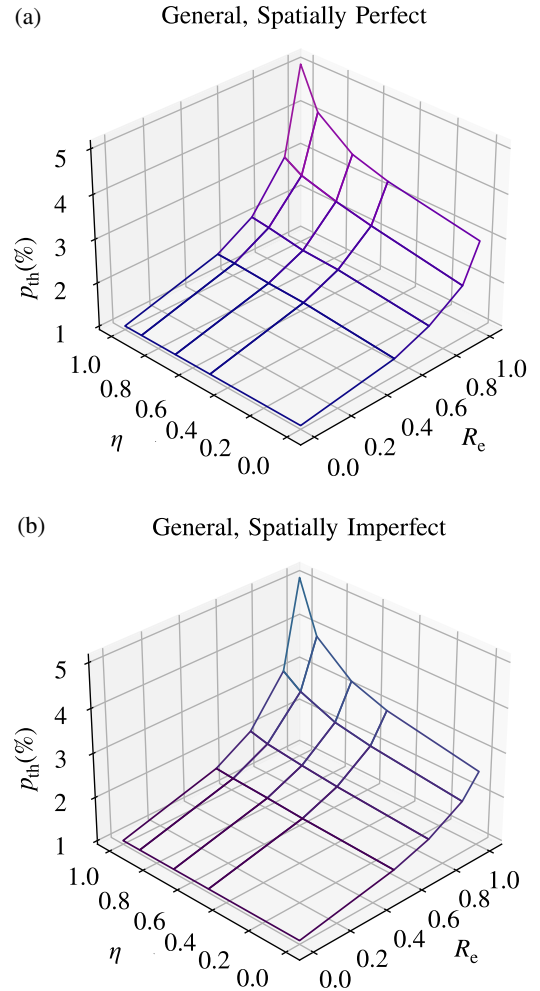


FIG. 7. (a) For spatially perfect checks and the general Pauli model, we plot the threshold (Z axis) parameterized over η and R_e . (b) Similarly, we repeat this for spatially imperfect checks with the same leakage-induced Pauli model. The threshold values in these plots are listed in table form in Appendix H.

remain undetected for the rest of the memory experiment, leading to unbounded temporal-correlated errors. In Fig. 6, we investigate this setting of $p_m = 1/2$ with a tailored Pauli model and spatially imperfect checks. We see that even with unbounded propagation of leakage, we observe a pseudothreshold of $2.37 \pm 0.02\%$ for distances $d \in \{9, 11, 13, 15\}$, which is still more than 2 times higher than the Pauli noise threshold. We note that $2.37 \pm 0.02\%$ is lower than the thresholds for $\eta = 0$ in Fig. 5, indicating that unbounded versus bounded propagation of leakage affects the threshold. This motivates the engineering of erasure architectures such that erasures are detected after a bounded number of time steps. Interestingly, even though we have used the imperfect erasure decoder for bounded propagation of leakage (Sec. IV), we still achieve a high pseudothreshold that remains above the Pauli noise threshold. It is possible that more specialized decoders [65,66]

that account for larger temporal correlations are likely to achieve even better thresholds. Furthermore, more study should be conducted to determine whether a true threshold exists.

We have so far discussed what happens when the errors are only erasures and we neglect any amount of Pauli errors. Now, we show a three-dimensional (3D) plot of the threshold parameterized over two variables—the temporal resolution η and erasure fraction R_e —for the general Pauli model in Figs. 7(a), 7(b) and the tailored Pauli model in Figs. 8(a), 8(b).

The thresholds all converge to $1.00 \pm 0.01\%$ when there is only Pauli noise ($R_e = 0$). We see the largest variation in thresholds among the four noise models when the erasure fraction is high because the thresholds are dominated by erasures and are susceptible to variations in the accuracy of erasure detection.

VI. DISCUSSION

In this work, we have analyzed the surface code parameters of threshold and effective error distance for erasure qubits with imperfect erasure checks. We find that erasure qubits outperform qubits with dominant Pauli noise on these metrics even when the erasure checks are imperfect, particularly if the leakage-induced noise has an inherent structure like that in the tailored Pauli model. Thus, our work firstly highlights the continued advantages of erasure qubits, and further emphasizes the importance of measuring the detailed noise model in experiments which can be nontrivial and may require the development of new techniques [67].

Our work motivates further analysis of the trade-offs between simplifying the erasure checks and code parameters. One simple approach to reduce the time cost of erasure checks is to perform them more infrequently, such as once per stabilizer measurement. However, this would result in a single leakage event persisting for the whole stabilizer measurement cycle which could degrade the threshold and effective distance [39]. It would be fruitful to identify practical gates and leakage-induced Pauli models which do not reduce the effective distance even with infrequent erasure checks.

Note added. A related, independent work [39] examines alternative erasure check schedules to relax check requirements on the performance of the surface code.

ACKNOWLEDGMENTS

This research was supported by the U.S. Army Research Office (ARO) under grant W911NF-23-1-0051. K.C. would like to thank Yue Wu, Aleksander Kubica, and Bailey Gu for useful discussions. K.C. and S.S. also thank Akshay Kootandavida for discussions on experimental implementations of the dual rail.

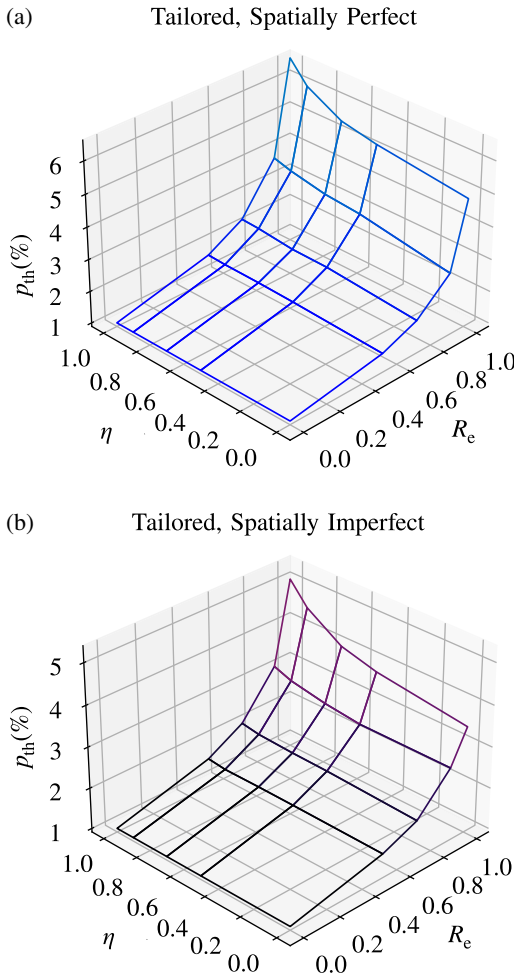


FIG. 8. (a) For spatially perfect checks and the tailored Pauli model, we plot the threshold (Z axis) parameterized over η and R_e . (b) Similarly, we repeat this for spatially imperfect checks with the same leakage-induced Pauli model. The threshold values in these plots are listed in table form in Appendix H.

DATA AVAILABILITY

The data is not publically available. The data is available from the authors upon reasonable request.

APPENDIX A: GATE WITH BUILT-IN ERASURE CHECK

In this gate, which is described in more detail in Appendices B and C, the gate and erasure check are implemented simultaneously by coupling the dual rail with a transmon. At the end of this gate, the transmon is measured. Depending on the measured state, one can deduce if there was an error during the gate, flagging the gate [20,28]. The transmon raises two kinds of flags: one due to a relaxation event (T_1 event) in the transmon itself and the other due to a dephasing event in the transmon or due to a relaxation or photon loss event in the dual rail. There are a few important points to note. Firstly, the transmon flag cannot differentiate between a dephasing event and an erasure in the dual-rails. Secondly, it cannot resolve between erasures in the two dual-rail qubits participating in the gate. Thirdly, once a transmon T_1 error is flagged, both the dual-rails must be reset as they may have incurred an unknown leakage error. Thus, the above observations imply that the gate has imperfect spatial resolution because, as soon as any flag is seen, there is an uncertainty in the type of error and both dual-rails must be reset. Finally, the transmon flags erasure or photon loss in the dual-rail only with 50% probability. Thus, it also has an imperfect temporal resolution

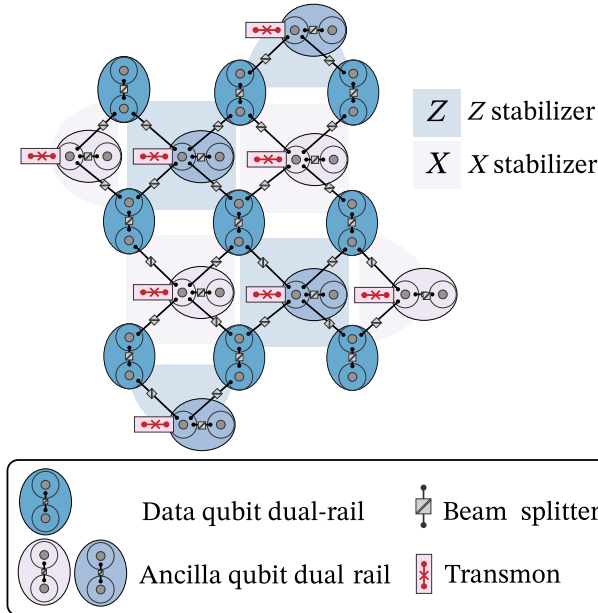


FIG. 9. Potential layout of transmons and cavities in a surface code implemented with cavity-based dual-rail qubits with built-in erasure checks. Instead of explicit erasure checks on each qubit, which would require an extra transmon and readout line on every physical qubit, the built-in nature of these checks halves the number of these ancillary transmons.

which depends on the rate of single-photon loss in the dual-rails (κ), and the rates of dephasing $1/T_\phi$ and relaxation in the transmon $1/T_1$ (note that here T_ϕ and T_1 are that of the f -level of the transmon), $\eta = (\kappa + 1/T_1 + 1/T_\phi)/(2\kappa + 1/T_1 + 1/T_\phi)$. For conservative experimental parameters of $T_1 = 30 \mu\text{s}$, $T_\phi = 25 \mu\text{s}$, and $\kappa = 1 \text{ ms}^{-1}$, we have $\eta \sim 0.986$.

Due to the built-in nature of this erasure check, in principle, one only needs one ancilla transmon for every ancilla dual-rail qubit on the surface code. We illustrate a potential layout of hardware elements in Fig. 9 of a 3×3 rotated code. This would only require eight readout lines (one per ancilla transmon), in comparison to 17 readout lines for explicit erasure checks on every physical qubit. More details on hardware costs of different architectures are given in Appendix D.

APPENDIX B: FIRST-ORDER ERROR DETECTION IN GATES WITH AND WITHOUT BUILT-IN CHECKS

Now, we describe in more detail two scenarios of erasure checks presented in Fig. 10. We first consider two-qubit gates as proposed in Refs. [20,28] followed by explicit erasure checks for cavity leakage. The scheme for the two-qubit gate leverages the dispersive interaction between a transmon and one of the cavities to perform a $ZZ(\pi/2)$ maximally entangling gate. Transmon errors, if undetected, propagate onto the cavities as either dephasing or leakage errors. However, detecting a transmon error via

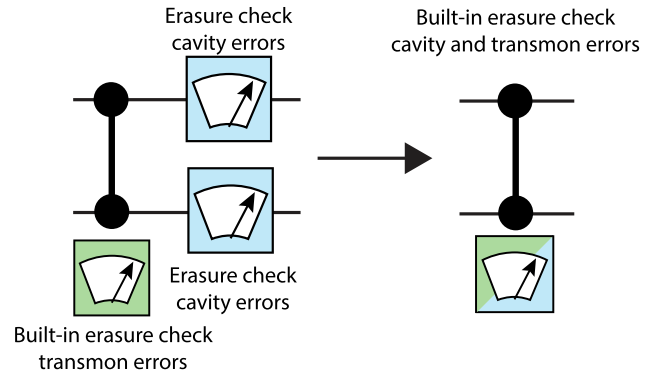


FIG. 10. (Left) Typical use case for an entangling gate between two erasure qubits. Erasure checks are performed after the gate on both qubits. In the gate proposal in Ref. [20] for dual-rail cavity qubits, the gate comes with “built-in” error detection for transmon ancilla errors which may occur during the gate. However, separate erasure checks are needed after the gate to detect cavity errors. (Right) Our proposed gate modification combines the error detection of transmon and cavity errors into a single measurement (of the transmon). As a consequence of this simplification, we will be unable to detect cavity photon loss errors if they occur midway through the gate with 100% probability, thus motivating our need to study delayed erasure detection.

transmon readout after the gate allows us to convert these errors to erasures. One prominent feature of the gate is the ability to read out the state of the transmon at the end of the gate, which allows us to detect whether transmon errors have occurred during the gate. In particular, transmon decay errors are mapped to the $|e\rangle$ state while transmon dephasing errors are mapped to the $|f\rangle$ state. Measuring the $|g\rangle$ state signals that no transmon error has occurred during the gate. If we measure the transmon to be in $|e\rangle$ or $|f\rangle$ after the gate, we must reset both dual-rail cavity qubits.

Detecting transmon errors allows us to use the dispersive interaction between the transmon and the cavity to perform a dual-rail entangling gate reasonably quickly (of the order of $\sim 1\text{-}\mu\text{s}$ gate time) while preventing our Pauli error rate from being limited by transmon decoherence.

However, it is not sufficient to detect transmon ancilla errors in a dual-rail cavity qubit alone. We must also detect photon loss errors and convert these to erasures (even though they are expected to be 10 times rarer). Otherwise, photon loss will lead to a disastrous accumulation of leakage errors.

The ZZ gate proposed in Ref. [20] actually already has the ability to detect some photon loss errors if they occur during the gate. In particular, a photon loss error dephases the transmon ancilla (by inducing a Z rotation by an angle which depends on the exact time of the photon loss), and so it will be detected in the $|f\rangle$ state with 50% probability at the end of the gate. In order to catch any remaining leakage or leakage, erasure checks must be performed by measuring the joint-parity operator of each dual-rail qubit after the gate [20].

These erasure checks take time and, in turn, expose us to more cavity and transmon errors. Instead, we propose modifying the ZZ gate protocol to have an in-built check for photon loss errors in the cavities, thereby eliminating the need for costly dedicated erasure check operations. This new paradigm is illustrated in Fig. 10 and also allows us to place transmon ancillae more sparsely in a dual-rail architecture, in every other physical qubit.

Our new gate will instead have the following error detection properties.

- (1) Transmon decay during the gate $\rightarrow |e\rangle$.
- (2) Transmon dephasing during the gate $\rightarrow |f\rangle$.
- (3) Leakage on a dual-rail qubit *during* the gate $\rightarrow |f\rangle$ (with $\approx 50\%$ probability).
- (4) One dual-rail qubit is leaked *before* the gate $\rightarrow |f\rangle$.
- (5) Both dual-rail qubits are leaked *before* the gate $\rightarrow |g\rangle$.

To reiterate, properties (1)–(3) and (5) are inherent to the ZZ gate proposed in Ref. [20] while (4) results from our proposed modification. Item (5) is an undesirable property still present in our modification in that we are unable to

reliably detect if both dual-rail qubits are leaked before or during the gate. However, this is a second-order error which could be detected with sparsely interspersed erasure check operations.

APPENDIX C: PROTOCOL FOR THE TWO-QUBIT GATE WITH LEAKAGE DETECTION

Next, we briefly describe the ZZ gate in Ref. [20]. First, we write the four dual-rail basis states in the Z basis as $\{|01, 01\rangle, |01, 10\rangle, |10, 01\rangle, |10, 10\rangle\}$, using the notation $|N_{a_1}, N_{b_1}, N_{a_2}, N_{b_2}\rangle$. If we examine the joint parity of the middle two cavities, a_2 and b_1 , we observe that $\{|01, 01\rangle, |10, 10\rangle\}$ have odd joint parity while $\{|01, 10\rangle, |10, 01\rangle\}$ have even joint parity. The ZZ gate puts a relative phase between these two sets of logical states and can be achieved by the circuit in Fig. 11(a).

It is apparent that if a dual-rail qubit begins in the leakage state $|0, 0\rangle$, there is no way to detect this by only measuring the middle two cavities with the transmon ancilla. [e.g., state $|00\rangle$ in dual rail (a_1, b_1) looks “identical” to state $|10\rangle$ if only cavity b_1 interacts with the transmon].

We can allow the transmon to measure the state of all four cavities during the ZZ gate by simply adding cavity-cavity SWAP gates between pairs (a_1, b_1) and (a_2, b_2) in between the control joint-parity unitaries, as shown in Fig. 12(a). After the gate, we may optionally perform another set of cavity swaps or merely track this as additional X gates on both dual-rail qubits.

For the transmon to be in $|g\rangle$ at the end of the gate, both controlled joint-parity unitaries must induce the same phase shift (0 or π) on the ancilla transmon.

Suppose that we began in state $|00, 01\rangle$ before we perform the gate. The first control joint-parity unitary will induce 0-phase shift on the transmon ancilla. However, after the swaps, the state in the cavities will be $|00, 10\rangle$ for the second control-joint-parity unitary, resulting in a π phase shift (Z operator) on the transmon. Consequently, we measure the transmon to be in $|f\rangle$ instead of $|g\rangle$, thereby detecting this photon loss error.

In other words, adding these swap operations means that we are also measuring if the joint parity of (b_1, a_2) is the same as the joint parity of (a_1, b_2) . Crucially, if we are still in the dual-rail codespace, we also do the desired ZZ entangling gate.

In practice, performing the swap on cavity pair (a_2, b_2) while the transmon is still in a superposition is thought to be impractical, because the strength of the dispersive interaction is comparable to the typical strength of the parametric beam splitter, and thus the beam-splitter interaction will be off-resonance when the transmon is in the $|f\rangle$ state. (It is possible to perform an “unconditional” SWAP gate, as shown in Ref. [28], but this is comparable or longer than a control joint-parity operation).

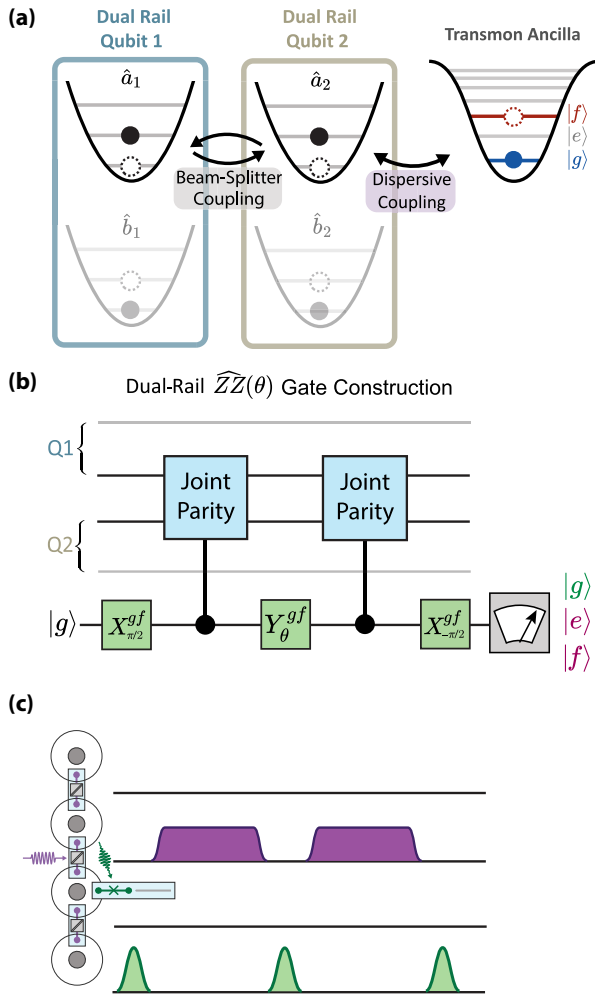


FIG. 11. (a) Dual-rail ZZ gate as proposed in Ref. [20]. With only a parametric beam splitter between the middle two cavity modes, (b_1, a_2), and a transmon dispersively coupled to only cavity a_2 , it is possible to perform a ZZ(θ) entangling gate between the two dual-rail qubits. (b) Quantum circuit construction for the dual-rail ZZ(θ) gate. This construction already has some degree of error detection built in. Measuring the transmon in $|g\rangle$ at the end of the gate signals no transmon errors have occurred during the gate. By performing transmon ancilla rotations in the g - f manifold instead of the g - e manifold, we can detect transmon decay errors to first order, which results in measuring the transmon in $|e\rangle$ at the end of the gate. Transmon dephasing errors are detected by measuring the transmon to be in $|f\rangle$. Photon loss errors in either dual-rail qubit during the gate result in a 50-50 probability of measuring $|g\rangle$ or $|f\rangle$ at the end of the gate and hence are only partially error detected. (c) Hardware layout and pulse sequence required for the dual-rail ZZ(θ) gate. The pulse sequence is a qualitative representation of the controls required and typical durations for implementing the gate.

In comparison, the (a_1, b_1) swap does not have this issue, since any transmon ancillae coupled to these cavities should remain in their ground state throughout the gate.

The final circuit and pulse sequence we suggest for experimental implementation is shown in Figs. 12(b)

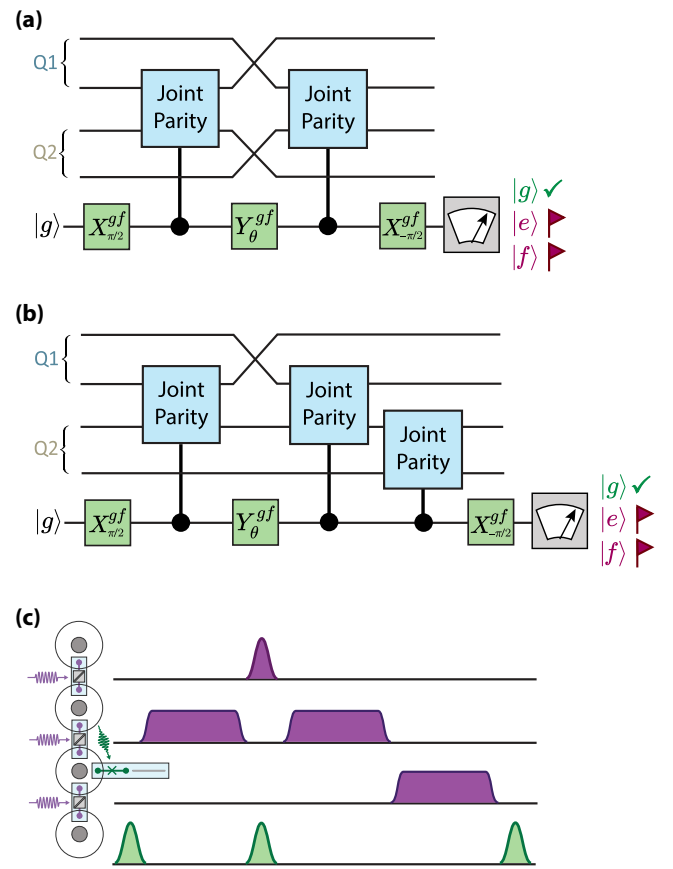


FIG. 12. (a) Modified quantum circuit for implementing a dual-rail ZZ(θ) gate with in-built erasure detection for cavity photon loss errors. Without additional SWAP operations after the gate, the circuit implements the entangling gate $X_1 X_2 ZZ(\theta)$. If one out of the two dual-rail qubits begins in a leakage state such as $|00\rangle$ or $|11\rangle$, we measure the transmon to be in $|f\rangle$ at the end of the gate. This allows us to remove the expensive dedicated erasure check operations from our implementation of the surface code. If photon loss happens midway through the gate, we only have a 50% chance of measuring the transmon in $|f\rangle$, leading to $\eta \approx 0.5$ for photon loss erasures when using this gate. (b) A more realistic proposal which does not rely on performing fast cavity swaps between (a_2, b_2) relative to the dispersive interaction strength. This sequence has the same error-detection properties as the circuit in (a) and implements the entangling gate $X_1 ZZ(\theta)$. (c) Qualitative pulse sequence and hardware required to implement the circuit in (b). Compared to the original ZZ gate scheme, our proposal does not require any additional hardware but lengthens the gate duration by around 50%.

and 12(c), in which we perform the (a_1, b_1) SWAP gate halfway through the gate, and add an extra control joint-parity unitary acting on (a_2, b_2). Overall, our modifications make the gate $\approx 50\%$ longer in duration but now effectively lump the check for transmon errors and erasure checks on the dual-rails into a single transmon measurement, leading to an overall reduction in the time needed to perform a round of stabilizer measurement and few transmon ancilla readout operations.

APPENDIX D: ACCOUNTING FOR THE NUMBER OF TRANSMONS AND CONTROL LINES IN DIFFERENT PROPOSED DUAL-RAIL ARCHITECTURES

Dual-rail qubits are inherently composite systems and hence require more hardware to realize each physical qubit. At the same time, their properties as erasure qubits promise more efficient error correction which has the potential to greatly reduce the number of total hardware elements required to reach some target logical error rate. It is important that we can quantify and estimate the hardware savings we can achieve with dual-rail qubits, which in turn requires quantifying these two affects.

We first begin by accounting for the number of hardware elements needed to realize what we term the “tile-able unit cell” for various possible superconducting qubit architectures, as shown in Fig. 13 and Table II. The tile-able unit cell includes not only the hardware needed to realize a physical qubit but also the couplings to surrounding qubits needed to realize a square lattice of physical qubits with nearest-neighbor connectivity, achieved solely by tiling the unit cell.

One way to compare the hardware resource overhead is to compare the number of transmonlike elements required to realize a surface code of a particular effective code distance. Compared to microwave resonators, transmons (which may be used as the physical qubit, tunable couplers, or beam-splitter couplers, as shown in Fig. 13) can be more resource intensive to fabricate at scale and so can serve as a point of reference for comparing different architectures.

This is not the only way to compare different architectures. For instance, we could also compare the total number of readout lines or control lines instead. Nonetheless, using the total number of transmons as our metric, we can obtain useful insights into how dual-rail erasure qubits can lead to hardware savings despite having a more complicated unit cell.

Now, let us fix the number of transmons and compare the effective distance d_{eff} when we use dual-rail transmon qubits with erasure noise [19,25] and transmons as qubits with Pauli noise. Suppose that we have N transmons, and we let n be the number of transmons per tile-able unit cell which makes a functioning physical qubit in a surface code. Physical qubits may include data and ancilla qubits, and n includes tunable couplers as described in Table II. A surface code contains of the order of d^2 physical qubits; thus, $N/n = d^2$. Furthermore, we know that, for pure erasure noise, the effective distance is $d_{\text{eff}}^{\text{eras}} = d$, while for Pauli noise, $d_{\text{eff}}^{\text{Pauli}} \approx d/2$. Thus, we quantify the gain factor in effective distance by using an erasure-based architecture with N transmons:

$$\frac{d_{\text{eff}}^{\text{eras}}}{d_{\text{eff}}^{\text{Pauli}}} = \frac{\sqrt{N/n_{\text{eras}}}}{\frac{1}{2}\sqrt{N/n_{\text{Pauli}}}} = 2\sqrt{\frac{n_{\text{Pauli}}}{n_{\text{eras}}}}. \quad (\text{D1})$$

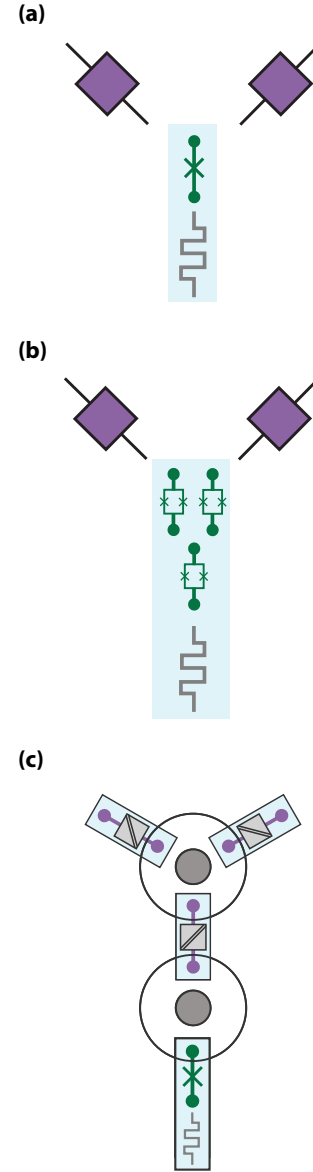


FIG. 13. Tile-able unit cells for different superconducting qubit architectures. These unit cells can be tiled in a plane to produce a square lattice of physical qubits with nearest-neighbor connectivity in order to realize a surface code. Counting the number of hardware elements per unit cell is a useful point of reference when comparing different possible qubit architectures. (a) Possible unit cell for transmon qubits (green) with tunable couplers (purple boxes). These tunable couplers are presumed to be transmonlike elements required for entangling gates. Each transmon is also presumed to have its own readout resonator (gray). (b) Possible unit cell for dual-rail transmon qubits as demonstrated in Ref. [25] (in another version, each transmon may have its own dedicated readout resonator). Entangling gates are presumed to require some kind of separate coupling element (purple). (c) A possible unit cell for dual-rail cavity qubits. Three beam-splitter couplers (purple) are required per unit cell, two of which are needed for entangling gates while the third actuates single-qubit gates. The ancilla transmon (green) and readout resonator (gray) are only needed in half the unit cells, for the ancilla dual-rail qubits in a surface code.

TABLE II. Accounting for the number of hardware elements for three different possible superconducting qubit architectures to realize a square lattice of physical qubits with nearest-neighbor entangling gate connectivity. We attempt to do our unit-cell accounting in a way which is as agnostic to the specific hardware implementation as possible. For the dual-rail cavity qubit architecture, only half the unit cells require an ancilla transmon and readout line, as reflected by values in the table.

Average number of elements per tile-able unit cell	Architecture		
	Transmons w/ tunable couplers (frequency tunable)	Dual-rail transmons w/ tunable couplers [19] (symmetric readout resonator)	Dual-rail cavities w/ beam-splitter couplers [20]
Transmonlike elements	3	5 (4)	3.5
Readout lines	1	3 (1)	0.5
Two-qubit gate control lines	2	2 (2)	2
Single-qubit control lines	1 (2)	5 (3)	1
Resonators	1	3 (1)	2.5

In a dual-rail transmon architecture, we assume that the number of transmons per physical qubit (including tunable couplers) is $n_{\text{eras}} = 5$. Using transmon qubits with tunable couplers requires $n_{\text{Pauli}} = 3$ transmons per physical qubit. Thus, our gain in effective distance by using dual-rail transmon qubits is $d_{\text{eff}}^{\text{eras}}/d_{\text{eff}}^{\text{Pauli}} = 2\sqrt{3/5} = 1.55$. Therefore, for the same number of transmons N , there is only a 55% increase in d_{eff} when using dual-rail transmon qubits. If we exclude the couplers in the transmon count then we adjust the number of transmons per unit cell to $n_{\text{Pauli}} = 1$ and $n_{\text{eras}} = 3$. Then, the new d_{eff} ratio is $d_{\text{eff}}^{\text{eras}}/d_{\text{eff}}^{\text{Pauli}} = 2\sqrt{1/3} = 1.15$. Thus, there is only a 15% increase in d_{eff} when using dual-rail transmon qubits instead of transmon qubits with Pauli noise. This is in contrast to a 100% increase in d_{eff} , as naively expected by virtue of using erasure qubits instead of regular qubits with Pauli noise. This motivates the desire for hardware-efficient erasure architectures that can achieve a ratio closer to the theoretical gain in effective distance. Repeating this exercise for a dual-rail cavity architecture with built-in checks and assuming that $n_{\text{eras}} = 3.5$, we see that the gain in effective distance is 1.85, or equivalently an 85% increase in d_{eff} .

Now suppose that we want to achieve a target effective error distance d_{eff} . We compare how many fewer transmons we need to reach this target distance on a surface code with dual-rail cavity qubits than transmon qubits. We set the effective distances equal to each other,

$$d_{\text{eff}}^{\text{eras}} = d_{\text{eff}}^{\text{Pauli}}, \quad (\text{D2})$$

$$\sqrt{\frac{N_{\text{eras}}}{n_{\text{eras}}}} = \frac{1}{2} \sqrt{\frac{N_{\text{Pauli}}}{n_{\text{Pauli}}}}, \quad (\text{D3})$$

$$\frac{N_{\text{Pauli}}}{N_{\text{eras}}} = 4 \left(\frac{n_{\text{Pauli}}}{n_{\text{eras}}} \right). \quad (\text{D4})$$

Therefore, by using built-in erasure checks, we use $N_{\text{Pauli}}/N_{\text{eras}} = 4(3/3.5) = 3.4$ times fewer transmons. In other words, we need 29% of the transmons to achieve the same d_{eff} compared to using transmon qubits with Pauli

noise. If we were to use a dual-rail transmon architecture (with the physical qubit shown in Ref. [25]), we would instead need 41% of the transmons to achieve the same d_{eff} .

APPENDIX E: DECODING IMPERFECT ERASURES

In Sec. IV, we described the population of errors depending on the spatial and temporal resolution of erasure checks. We now describe in detail how these error mechanisms lead to edge weights that are defined on the decoding graph.

1. Decoding preliminaries

We decode Z_L and X_L information independently by defining two decoding graphs with nodes created from X -type and Z -type syndromes, respectively. In any decoding graph, an edge with a lower weight means that the error that creates syndromes at its ends is more probable. The standard way of weighing an edge is $\text{wt} = \log(1-p)/p$, where p is the probability of an error mechanism that creates syndromes at the ends of that edge [40]. The probability of Pauli errors due to noisy gates results in an edge weight distribution that is static in that it does not change with every memory experiment. If there are no Pauli errors, such as in the case where the only errors are erasures, then all edges have static weights $\text{wt} = \infty$. Assuming this static, Pauli decoding graph as a starting point, we will talk about how the decoder uses the location of erasure flags to dynamically modify the weights of edges.

2. Perfect erasures

In the case of perfect erasures, leakage is always accurately detected and immediately reinitialized into the computational mixed state $I/2$ following detection. This depolarizing error results in specific syndromes which appear as two nodes in the X decoding graph with probability $1/2$ and Z decoding graph with probability $1/2$. Thus, the

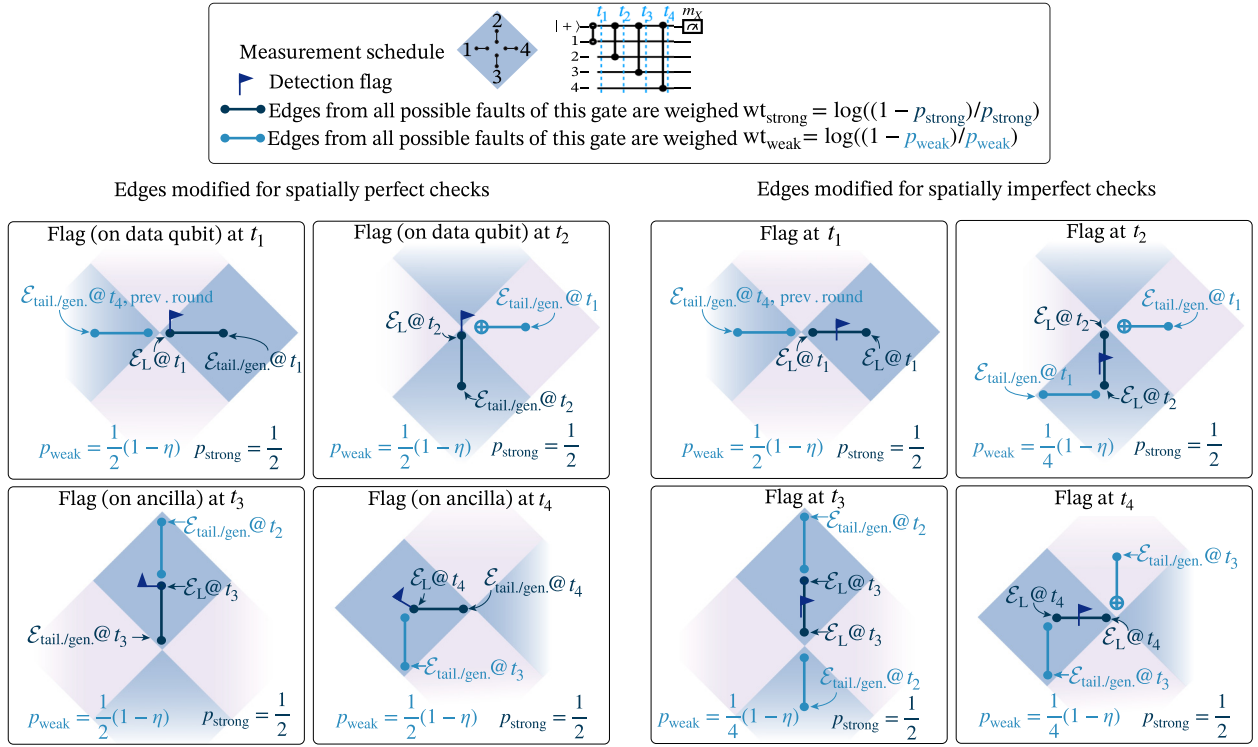


FIG. 14. Given a flag from a spatially perfect, temporally imperfect check (left four boxes) or spatially imperfect, temporally imperfect check (right four boxes), we show which gates and their errors chosen from sets labeled $\mathcal{E}_L$, $\mathcal{E}_{\text{tail}}$, or $\mathcal{E}_{\text{gen}}$, should be strongly erased (dark blue) or weakly erased (light blue). See Appendix E 3 for an in-depth discussion.

edges between these syndromes are weighed $wt_{\text{strong}} = \log((1/2)/(1/2)) = 0$ and the syndromes are automatically paired up in a 0-parity cluster. We use the subscript strong to indicate that we know with certainty its syndromes should be joined within a 0-parity cluster in the decoding graph.

3. Imperfect erasures

Now we discuss our decoder for imperfect erasure detection. This decoder takes the locations of flagged erasures and dynamically modifies edges depending on the leakage-induced Pauli model, spatial resolution, and temporal resolution η . In Fig. 14, given a spatial resolution (perfect or imperfect) and imperfect temporal resolution which is continuously varied by η , we illustrate which edges require weight modification for a flag at each step of the stabilizer measurement circuit ($t_1 - t_4$). The left four boxes indicate spatially perfect checks that flag individual qubits, while the right four boxes represent spatially imperfect checks that flag gates. For each flag in Fig. 14, we identify all gates with potential errors that occur at a time step immediately after the gate, at t_1, t_2, t_3 , or t_4 . These errors are chosen from either the sets $\mathcal{E}_L$, $\mathcal{E}_{\text{gen}}$, or $\mathcal{E}_{\text{tail}}$, in accordance to how the models are defined in Sec. IV. Note that due to separate X - and Z -syndrome decoding graphs,

when collecting edges that result from faults drawn from $\mathcal{E}_{\text{tail}}$, edges may be asymmetrically modified in one graph compared to the other.

A dark blue *strongly erased* gate indicates our certainty that an error was definitely present after this gate, chosen out of the set of possible errors (denoted by $\mathcal{E}_L$, $\mathcal{E}_{\text{tail}}$, or $\mathcal{E}_{\text{gen}}$). Conversely, a light blue *weakly erased* gate suggests that errors may have happened with a probability depending on η . Errors following a weakly erased gate are selected from either $\mathcal{E}_{\text{tail}}$ or $\mathcal{E}_{\text{gen}}$, depending on the leakage-induced Pauli model. Lower values of η correspond to poorer temporal resolution, resulting in a higher likelihood of such errors after the weakly erased gate. The probability that an error selected from these sets will lead to an edge on either the X - or Z -decoding graph is determined by the probabilities p_{strong} (for errors at strongly erased gates) and p_{weak} (for errors at weakly erased gates). Therefore, we assign weights to the edges on the X - and Z -decoding graphs that are formed by these error sets with either $wt_{\text{strong}} = \log(1 - p_{\text{strong}})/p_{\text{strong}} = 0$ or $wt_{\text{weak}} = \log(1 - p_{\text{weak}})/p_{\text{weak}} > 0$ for errors at strongly or weakly erased gates, respectively. Figure 14 only display flags at CZ gates, but the same analysis would be applicable for any flagged CX gate by flipping the type (CX or CZ) of all weakly erased gates.

APPENDIX F: ROTATED CODES

Although the numerical analysis in this work is for the unrotated code, all the thresholds are also applicable to the rotated code. In Fig. 15, we illustrate the rotated surface code for a lattice size of $d = 3$, where qubits are on the vertices, X stabilizers are represented by light faces, and Z stabilizers are represented by dark faces. We give a representative of the X_L operator acting on the top three qubits and the Z_L operator on three qubits on the left boundary. We now analytically examine the effect of imperfect checks on the effective error distance on rotated surface codes. To determine if the rotated surface code experiences error distance reduction, we examine the standard measurement schedule [37] of the rotated code, where numbers 1–4 in Fig. 15 indicate the order of the two-qubit gates. This particular scheduling is to avoid *hook errors* in a circuit-level noise model in which Pauli errors may occur in between any gate in the stabilizer measurement circuit [37]. Errors in the middle of the syndrome extraction circuit may lead to correlated errors on the data qubits. The measurement schedule shown in Fig. 15 ensures that these correlated errors do not result in a decrease in the effective distance [37]. As introduced in Sec. IV, under delayed detection, a leaked qubit A interacts with two qubits, B and C , populating them with errors according to the general or tailored Pauli model. Moreover, we saw in Sec. V A that the effective distance is reduced to the Pauli noise case when qubits B and C have errors that lead to the same logical error. We see that the measurement schedule leads

to configurations of qubits B and C along both X_L and Z_L operators. As a result, under the general Pauli model, qubits B and C are depolarized, thus reducing the effective distance to X_L and Z_L . Now let us examine what happens in the tailored Pauli model. Here, the errors on qubit B are chosen from the set $\mathcal{E}_{\text{tail.}}$, which will not have support on the same X_L or Z_L operator that qubit C lies on. Thus, even in the rotated code, we expect the tailored Pauli model with delayed erasure checks to have unchanged effective error distance to both logical errors.

APPENDIX G: CONFIRMING EFFECTIVE DISTANCE SCALING WITH DIFFERENT LATTICE SIZES

In Fig. 16, we observe that the diminished effective distance to a Z_L error approaches the same error distance for Pauli noise, $(d + 1)/2$ (solid line), for various unrotated lattice sizes $d = 3, 5$, and 7 (diamonds), as predicted in Sec. IV A. Here we have considered the general Pauli model with spatially imperfect checks [see Fig. 3(a)] at temporal resolution $\eta = 0$. As we have observed that the effective error distance is independent of spatial resolution (see Sec. V A), we do not expect these results to change for spatially perfect checks. The fact that the error distance approaches $(d + 1)/2$ is unsurprising because each leakage event leads to potentially two data qubit Z errors along Z_L . This means that, for an odd number of qubits in Z_L , exactly $(d + 1)/2$ leakage events may cause Z errors on all the qubits in support of the logical.

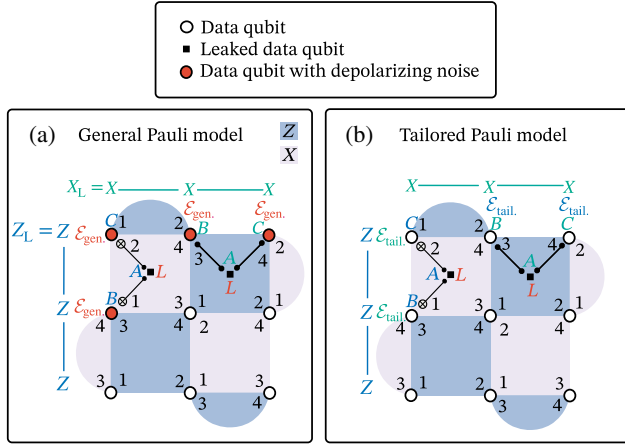


FIG. 15. (a) General Pauli model. The effective error distance to a Z_L and X_L error should be reduced. Leakage of the ancilla at an X stabilizer on the first CX gate leads to two depolarized qubits along the Z_L operator. Similarly, leakage of the ancilla at a Z stabilizer third CZ gate leads to the same scenario along the X_L operator. (b) We show that, under the tailored Pauli model, these cases do not lead to multiple errors with support along the logical operators.

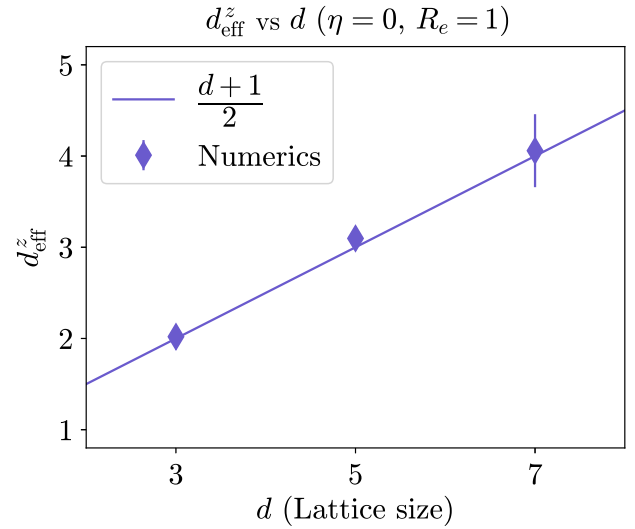


FIG. 16. Numerically obtained effective error distances to a Z_L error at $\eta = 0$ for lattice sizes $d = 3, 5, 7$. The solid line is the Pauli noise effective error distance, $(d + 1)/2$.

General Pauli model and spatially imperfect checks				
	$R_e = 0.52$	$R_e = 0.71$	$R_e = 0.91$	$R_e = 1.0$
$\eta = 0.0$	1.40 ± 0.01	1.66 ± 0.02	2.11 ± 0.02	2.59 ± 0.01
$\eta = 0.5$	1.48 ± 0.01	1.80 ± 0.02	2.35 ± 0.02	2.98 ± 0.01
$\eta = 0.7$	1.53 ± 0.01	1.88 ± 0.02	2.50 ± 0.02	3.27 ± 0.01
$\eta = 0.9$	1.57 ± 0.02	1.98 ± 0.02	2.83 ± 0.02	3.93 ± 0.01
$\eta = 1.0$	1.60 ± 0.02	2.08 ± 0.02	3.12 ± 0.04	5.09 ± 0.02

General Pauli model and spatially perfect checks				
	$R_e = 0.52$	$R_e = 0.71$	$R_e = 0.91$	$R_e = 1.0$
$\eta = 0.0$	1.32 ± 0.01	1.63 ± 0.02	2.14 ± 0.02	2.96 ± 0.02
$\eta = 0.5$	1.51 ± 0.01	1.86 ± 0.02	2.49 ± 0.02	3.33 ± 0.02
$\eta = 0.7$	1.54 ± 0.01	1.90 ± 0.02	2.66 ± 0.02	3.57 ± 0.03
$\eta = 0.9$	1.59 ± 0.02	2.03 ± 0.02	2.88 ± 0.02	4.17 ± 0.03
$\eta = 1.0$	1.60 ± 0.02	2.08 ± 0.02	3.12 ± 0.04	5.09 ± 0.02

Tailored Pauli model and spatially imperfect checks				
	$R_e = 0.52$	$R_e = 0.71$	$R_e = 0.91$	$R_e = 1.0$
$\eta = 0.0$	1.50 ± 0.01	1.85 ± 0.02	2.70 ± 0.02	3.49 ± 0.02
$\eta = 0.5$	1.54 ± 0.02	1.95 ± 0.02	2.70 ± 0.02	3.79 ± 0.02
$\eta = 0.7$	1.55 ± 0.02	1.97 ± 0.02	2.80 ± 0.02	4.02 ± 0.02
$\eta = 0.9$	1.55 ± 0.02	2.01 ± 0.02	2.96 ± 0.02	4.58 ± 0.02
$\eta = 1.0$	1.60 ± 0.02	2.08 ± 0.02	3.12 ± 0.04	5.09 ± 0.02

Tailored Pauli model and spatially perfect checks				
	$R_e = 0.52$	$R_e = 0.71$	$R_e = 0.91$	$R_e = 1.0$
$\eta = 0.0$	1.48 ± 0.01	1.92 ± 0.02	2.84 ± 0.02	4.87 ± 0.05
$\eta = 0.5$	1.62 ± 0.02	2.16 ± 0.02	3.32 ± 0.02	5.24 ± 0.03
$\eta = 0.7$	1.66 ± 0.02	2.18 ± 0.02	3.42 ± 0.03	5.47 ± 0.04
$\eta = 0.9$	1.68 ± 0.02	2.30 ± 0.02	3.64 ± 0.03	6.07 ± 0.03
$\eta = 1.0$	1.71 ± 0.02	2.33 ± 0.03	3.80 ± 0.03	6.71 ± 0.05

APPENDIX H: THREE-DIMENSIONAL THRESHOLD PLOTS

We present the thresholds in Figs. 7 and 8 in table form. For constructing these figures, we evaluate the threshold at four different values of erasure fraction $R_e = 0.5, 0.7, 0.9, 1.0$ and five different values of $\eta = 0.0, 0.5, 0.7, 0.9, 1.0$ each. Although not displayed, when there is only Pauli noise at $R_e = 0.0$, the threshold is $1.00 \pm 0.01\%$ for all four noise models.

- [1] P. Aliferis, and J. Preskill, Fault-tolerant quantum computation against biased noise, *Phys. Rev. A* **78**, 052331 (2008).
- [2] P. Aliferis, F. Brito, D. P. DiVincenzo, J. Preskill, M. Steffen, and B. M. Terhal, Fault-tolerant computing with

- biased-noise superconducting qubits: A case study, *New J. Phys.* **11**, 013061 (2009).
- [3] P. Brooks, and J. Preskill, Fault-tolerant quantum computation with asymmetric Bacon-Shor codes, *Phys. Rev. A* **87**, 032310 (2013).
- [4] A. M. Stephens, W. J. Munro, and K. Nemoto, High-threshold topological quantum error correction against biased noise, *Phys. Rev. A* **88**, 060301(R) (2013).
- [5] P. Webster, S. D. Bartlett, and D. Poulin, Reducing the overhead for quantum computation when noise is biased, *Phys. Rev. A* **92**, 062309 (2015).
- [6] D. K. Tuckett, A. S. Darmawan, C. T. Chubb, S. Bravyi, S. D. Bartlett, and S. T. Flammia, Tailoring surface codes for highly biased noise, *Phys. Rev. X* **9**, 041031 (2019).
- [7] D. K. Tuckett, S. D. Bartlett, and S. T. Flammia, Ultrahigh error threshold for surface codes with biased noise, *Phys. Rev. Lett.* **120**, 050505 (2018).
- [8] D. K. Tuckett, S. D. Bartlett, S. T. Flammia, and B. J. Brown, Fault-tolerant thresholds for the surface code in excess of 5% under biased noise, *Phys. Rev. Lett.* **124**, 130501 (2020).
- [9] J. P. Bonilla Ataides, D. K. Tuckett, S. D. Bartlett, S. T. Flammia, and B. J. Brown, The XZZX surface code, *Nat. Commun.* **12**, 2172 (2021).
- [10] A. Dua, A. Kubica, L. Jiang, S. T. Flammia, and M. J. Gullans, Clifford-deformed surface codes, *PRX Quantum* **5**, 010347 (2024).
- [11] O. Higgott, T. C. Bohdanowicz, A. Kubica, S. T. Flammia, and E. T. Campbell, Improved decoding of circuit noise and fragile boundaries of tailored surface codes, *Phys. Rev. X* **13**, 031007 (2023).
- [12] Q. Xu, N. Mannucci, A. Seif, A. Kubica, S. T. Flammia, and L. Jiang, Tailored XZZX codes for biased noise, *Phys. Rev. Res.* **5**, 013035 (2023).
- [13] S. Singh, A. S. Darmawan, B. J. Brown, and S. Puri, High-fidelity magic-state preparation with a biased-noise architecture, *Phys. Rev. A* **105**, 052410 (2022).
- [14] A. S. Darmawan, B. J. Brown, A. L. Grimsmo, D. K. Tuckett, and S. Puri, Practical quantum error correction with the XZZX code and Kerr-cat qubits, *PRX Quantum* **2**, 030345 (2021).
- [15] D. Gottesman, Stabilizer codes and quantum error correction. Caltech Ph. D. Thesis, eprint: [quant-ph/9705052](https://arxiv.org/abs/quant-ph/9705052) (1997).
- [16] M. Grassl, T. Beth, and T. Pellizzari, Codes for the quantum erasure channel, *Phys. Rev. A* **56**, 33 (1997).
- [17] 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).
- [18] 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).
- [19] A. Kubica, A. Haim, Y. Vaknin, H. Levine, F. Brandão, and A. Retzker, Erasure qubits: Overcoming the T_1 limit in superconducting circuits, *Phys. Rev. X* **13**, 041022 (2023).
- [20] J. D. Teoh, P. Winkel, H. K. Babla, B. J. Chapman, J. Claes, S. J. de Graaf, J. W. Garmon, W. D. Kalfus, Y. Lu, A. Maiti *et al.*, Dual-rail encoding with superconducting cavities, *Proc. Natl. Acad. Sci.* **120**, e2221736120 (2023).

- [21] W. C. Campbell, Certified quantum gates, *Phys. Rev. A* **102**, 022426 (2020).
- [22] 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).
- [23] 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* **622**, 279 (2023).
- [24] 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, and M. H. Devoret, Erasure detection of a dual-rail qubit encoded in a double-post superconducting cavity, *Phys. Rev. Lett.* **132**, 180601 (2024).
- [25] H. Levine *et al.*, Demonstrating a long-coherence dual-rail erasure qubit using tunable transmons, *Phys. Rev. X* **14**, 011051 (2024).
- [26] 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).
- [27] S. J. de Graaf, S. H. Xue, B. J. Chapman, J. D. Teoh, T. Tsunoda, P. Winkel, J. W. O. Garmon, K. M. Chang, L. Frunzio, S. Puri, and R. J. Schoelkopf, A mid-circuit erasure check on a dual-rail cavity qubit using the joint-photon number-splitting regime of circuit QED, [arXiv:2406.14621](https://arxiv.org/abs/2406.14621).
- [28] T. Tsunoda, J. D. Teoh, W. D. Kalfus, S. J. de Graaf, B. J. Chapman, J. C. Curtis, N. Thakur, S. M. Girvin, and R. J. Schoelkopf, Error-detectable bosonic entangling gates with a noisy ancilla, *PRX Quantum* **4**, 020354 (2023).
- [29] I. L. Chuang, and Y. Yamamoto, Simple quantum computer, *Phys. Rev. A* **52**, 3489 (1995).
- [30] E. Knill, R. Laflamme, and G. J. Milburn, A scheme for efficient quantum computation with linear optics, *Nature* **409**, 46 (2001).
- [31] E. Zakka-Bajjani, F. Nguyen, M. Lee, L. R. Vale, R. W. Simmonds, and J. Aumentado, Quantum superposition of a single microwave photon in two different ‘colour’ states, *Nat. Phys.* **7**, 599 (2011).
- [32] Y.-P. Shim, and C. Tahan, Semiconductor-inspired design principles for superconducting quantum computing, *Nat. Commun.* **7**, 11059 (2016).
- [33] D. L. Campbell, Y.-P. Shim, B. Kannan, R. Winik, D. K. Kim, A. Melville, B. M. Niedzielski, J. L. Yoder, C. Tahan, S. Gustavsson *et al.*, Universal nonadiabatic control of small-gap superconducting qubits, *Phys. Rev. X* **10**, 041051 (2020).
- [34] Y. Wu, and L. Zhong, Fusion blossom: Fast MWPM decoders for QEC, [ArXiv:2305.08307](https://arxiv.org/abs/2305.08307).
- [35] O. Higgott, and C. Gidney, Sparse blossom: Correcting a million errors per core second with minimum-weight matching, *Quantum* **9**, 1600 (2025).
- [36] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, Surface codes: Towards practical large-scale quantum computation, *Phys. Rev. A—At., Mol. Opt. Phys.* **86**, 032324 (2012).
- [37] Y. Tomita, and K. M. Svore, Low-distance surface codes under realistic quantum noise, *Phys. Rev. A* **90**, 062320 (2014).
- [38] S. Gu, A. Retzker, and A. Kubica, Fault-tolerant quantum architectures based on erasure qubits, *Phys. Rev. Res.* **7**, 013249 (2025).
- [39] S. Gu, Y. Vaknin, A. Retzker, and A. Kubica, Optimizing quantum error-correction protocols with erasure qubits, *PRX Quantum* **6**, 040354 (2025).
- [40] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, Topological quantum memory, *J. Math. Phys.* **43**, 4452 (2002).
- [41] P. Aliferis, and B. M. Terhal, Fault-tolerant quantum computation for local leakage faults, *Quantum Inf. Comput.* **7**, 139 (2007).
- [42] A. G. Fowler, Coping with qubit leakage in topological codes, *Phys. Rev. A* **88**, 042308 (2013).
- [43] J. Ghosh, and A. G. Fowler, Leakage-resilient approach to fault-tolerant quantum computing with superconducting elements, *Phys. Rev. A* **91**, 020302(R) (2015).
- [44] M. Suchara, A. W. Cross, and J. M. Gambetta, Leakage suppression in the toric code, *Quantum Inf. Comput.* **15**, 997 (2015).
- [45] N. C. Brown, and K. R. Brown, Leakage mitigation for quantum error correction using a mixed qubit scheme, *Phys. Rev. A* **100**, 032325 (2019).
- [46] F. Battistel, B. M. Varbanov, and B. M. Terhal, Hardware-efficient leakage-reduction scheme for quantum error correction with superconducting transmon qubits, *PRX Quantum* **2**, 030314 (2021).
- [47] M. McEwen, D. Kafri, Z. Chen, J. Atalaya, K. Satzinger, C. Quintana, P. V. Klimov, D. Sank, C. Gidney, A. Fowler *et al.*, Removing leakage-induced correlated errors in superconducting quantum error correction, *Nat. Commun.* **12**, 1761 (2021).
- [48] J. Marques, H. Ali, B. Varbanov, M. Finkel, H. Veen, S. van der Meer, S. Valles-Sanclemente, N. Muthusubramanian, M. Beekman, N. Haider *et al.*, All-microwave leakage reduction units for quantum error correction with superconducting transmon qubits, *Phys. Rev. Lett.* **130**, 250602 (2023).
- [49] 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).
- [50] N. C. Brown, and K. R. Brown, Comparing Zeeman qubits to hyperfine qubits in the context of the surface code: $^{174}\text{Yb}^+$ and $^{171}\text{Yb}^+$, *Phys. Rev. A* **97**, 052301 (2018).
- [51] A. Robertson, C. Granade, S. D. Bartlett, and S. T. Flammia, Tailored codes for small quantum memories, *Phys. Rev. Appl.* **8**, 064004 (2017).
- [52] L. Hägggeli, M. Heinze, and R. König, Enhanced noise resilience of the surface–Gottesman–Kitaev–Preskill code via designed bias, *Phys. Rev. A* **102**, 052408 (2020).
- [53] S. Huang, and K. R. Brown, Fault-tolerant compass codes, *Phys. Rev. A* **101**, 042312 (2020).
- [54] O. Higgott, and N. P. Breuckmann, Subsystem codes with high thresholds by gauge fixing and reduced qubit overhead, *Phys. Rev. X* **11**, 031039 (2021).
- [55] J. Guillaud, and M. Mirrahimi, Error rates and resource overheads of repetition cat qubits, *Phys. Rev. A* **103**, 042413 (2021).

- [56] C. Chamberland, K. Noh, P. Arrangoiz-Arriola, E. T. Campbell, C. T. Hann, J. Iverson, H. Putterman, T. C. Bohdanowicz, S. T. Flammia, A. Keller *et al.*, Building a fault-tolerant quantum computer using concatenated cat codes, [PRX Quantum 3, 010329 \(2022\)](#).
- [57] J. Roffe, L. Z. Cohen, A. O. Quintavalle, D. Chandra, and E. T. Campbell, Bias-tailored quantum LDPC codes, [Quantum 7, 1005 \(2023\)](#).
- [58] C. H. Bennett, D. P. DiVincenzo, and J. A. Smolin, Capacities of quantum erasure channels, [Phys. Rev. Lett. 78, 3217 \(1997\)](#).
- [59] N. Delfosse, and G. Zémor, Linear-time maximum likelihood decoding of surface codes over the quantum erasure channel, [Phys. Rev. Res. 2, 033042 \(2020\)](#).
- [60] N. Delfosse, and N. H. Nickerson, Almost-linear time decoding algorithm for topological codes, [Quantum 5, 595 \(2021\)](#).
- [61] S. Huang, M. Newman, and K. R. Brown, Fault-tolerant weighted union-find decoding on the toric code, [Phys. Rev. A 102, 012419 \(2020\)](#).
- [62] C. Wang, J. H. Harrington, and J. Preskill, Confinement-Higgs transition in a disordered gauge theory and the accuracy threshold for quantum memory, [Ann. Phys. 303, 31 \(2003\)](#).
- [63] T. M. Stace, S. D. Barrett, and A. C. Doherty, Thresholds for topological codes in the presence of loss, [Phys. Rev. Lett. 102, 200501 \(2009\)](#).
- [64] S. D. Barrett, and T. M. Stace, Fault tolerant quantum computation with very high threshold for loss errors, [Phys. Rev. Lett. 105, 200502 \(2010\)](#).
- [65] 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](#).
- [66] G. Baranes, M. Cain, J. Ataides, D. Bluvstein, J. Sinclair, V. Vuletic, H. Zhou, and M. D. Lukin, Leveraging atom loss errors in fault tolerant quantum algorithms, [ArXiv:2502.20558](#).
- [67] J. Claes, and S. Puri, Estimating the bias of CX gates via character randomized benchmarking, [PRX Quantum 4, 010307 \(2023\)](#).