Transversal CNOT gate with multicycle error correction

Younghun Kim[1,2,*] Martin Sevior[1,†] and Muhammad Usman[1,2,‡]

[1] *School of Physics, The University of Melbourne, Parkville, 3010 Victoria, Australia*

[2] *Data61, CSIRO, Clayton, 3168 Victoria, Australia*

(Received 29 June 2024; revised 18 December 2024; accepted 4 February 2025; published 28 February 2025)

A scalable and programmable quantum computer holds the potential to solve computationally intensive tasks that classical computers cannot accomplish within a reasonable time frame, achieving quantum advantage. However, the vulnerability of the current generation of quantum processors to errors poses a significant challenge towards executing complex and deep quantum circuits required for practical problems. Quantum error-correction codes such as stabilizer codes offer a promising path forward for fault-tolerant quantum computing; however, their realization on quantum hardware is an on-going area of research. In particular, fault-tolerant quantum processing must employ logical gates on logical qubits with error suppression with realistically large size codes. This work has implemented a transversal controlled-NOT (CNOT) gate between two logical qubits constructed using the repetition code with flag qubits, and demonstrated error suppression with increasing code size under multiple rounds of error detection and correction. Our results conducted on IBM quantum devices through cloud access show that despite the potential for error propagation among logical qubits during the transversal CNOT gate operation, scaling the code distance as we increase the number of physical qubits from 21 to 39 and 57 can suppress errors, which persists over ten rounds of error detection and correction. Our work establishes the feasibility of employing logical CNOT gates alongside error detection on a superconductor-based processor using current generation quantum hardware.

DOI: 10.1103/PhysRevApplied.23.024074

I. INTRODUCTION

Noisy intermediate-scale quantum (NISQ) [1] devices consist of several dozen to a few hundred noisy physical qubits accessible through various quantum processor platforms, such as superconductors [2,3], trapped ions [4,5] and neutral atoms [6]. Despite the potential of quantum processors to transcend classical computers in computational capabilities [3], solving practical problems by interpreting results from complex quantum circuits remains challenging due to errors. To obtain reliable results, it is crucial to limit errors below a target error rate [7], which can be achieved by constructing a fault-tolerant quantum computer with logical qubits and gates protected against errors by quantum error-correction codes, such as those provided by stabilizer codes [8–10]. Stabilizer codes allow the implementation of arbitrarily accurate quantum computation, provided that the error rate per physical qubit gate is kept below a certain acceptable threshold value [11–13].

Significant advances in the development of quantum devices have enabled the field of quantum error correction to shift from purely theoretical analysis to hardware benchmarking, with several recent demonstrations reporting the principles and realization of stabilizer codes employing various structures, such as the repetition code [14–17], surface code [18–22], and color code [23,24]. Experimental results have shown that, by increasing the number of physical qubits, a logical qubit can be constructed on quantum devices offering tolerance against errors with feedback based on detected errors over several rounds of the error-detection circuit, also known as a syndrome extraction circuit [15,16,18]. However, for practical quantum computing, it is essential to demonstrate the implementation and scaling of logical gates on logical qubits [25]—a feat that has not been fully achieved yet.

Amongst multiqubit logical operations, the implementation of a two-qubit logical controlled-NOT (CNOT) gate on quantum hardware will be an essential step towards achieving universal quantum gate set [26–30] and fault-tolerant quantum computing. One approach to realizing the logical CNOT gate is by performing a transversal CNOT gate [31–34]. Table I summarizes previous studies investigating the performance of multilogical qubit gates including the transversal CNOT gate on real quantum devices [35–40],

*Contact author: younghunk@student.unimelb.edu.au

†Contact author: martines@unimelb.edu.au

‡Contact author: muhammad.usman@unimelb.edu.au

TABLE I. Lists of various error-correction demonstrations with logical gates. Quantum error-correction codes are denoted as $[[n, k, d]]$, where n, k, and d represent the number of data qubits, the number of logical qubits, and the distance. Codes that can protect against either X or Z errors are denoted as $[n, k, d]$. The table indicates the number of syndrome extraction rounds, the hardware type, and error suppression achievement by increasing the size of the code block. Additionally, the table specifies the type of logical transversal gates utilized. The number of syndrome rounds refers to the number of implemented rounds of a subroutine quantum circuit that measures all eigenvalues of defined stabilizers across all logical qubits.

Reference	Year	Structure	Logical gates	Error suppression via code scaling	Syndrome rounds	Hardware modality
[35]	2022	[[7, 1, 3]]	CNOT	No	Zero	Trapped ion
[36]	2022	[[5, 1, 3]], [[7, 1, 3]]	CNOT	No	Single	Trapped ion (Quantinuum)
[37]	2023	[[8, 3, 2]]	CZ CCZ	No	Zero	Trapped ion (IonQ) Superconductors (IBM)
[38]	2023	[[9, 1, 3]]–[[49, 1, 7]], [[8, 3, 2]]	CNOT, CZ, CCZ	Yes	Single	Neutral atom (QuEra)
[39]	2024	[[7, 1, 3]], [[12, 2, 4]]	CNOT	Yes	Single	Trapped ion (Quantinuum)
[40]	2024	[[4, 1, 2]]–[[16, 1, 4]]	CNOT	No	Six[a]	Superconductors (IBM)
This work	2024	[3, 1, 3]–[7, 1, 7]	CNOT	Yes	Ten	Superconductors (IBM)

[a]We count the syndrome extraction rounds including the circuit used in initializing logical qubits.

noting that all of these studies have been reported within the last couple of years, indicating that the research on multiqubit logical gate implementation is still emerging. These studies have implemented multilogical qubit entangled states, like the logical Bell state, with a primary focus on evaluating their performance in terms of state preparation. Furthermore, previous experimental literature has explored quantum teleportation at the logical qubit level using transversal CNOT gates [41], or estimating the average fidelity per logical CNOT gate [42]. However, none of these papers have demonstrated error suppression of the logical CNOT gate with multiple cycles of syndrome extraction circuits through scaling up the code block size. Therefore, investigating the performance of a logical CNOT gate and demonstrating the possibility of error suppression while tracking errors on real quantum computers remains an open question in the NISQ era.

In this work, we design a *quantum memory experiment* to assess the effectiveness of a transversal logical CNOT gate, including multiple rounds of a syndrome extraction circuit. To evaluate the performance of a transversal CNOT gate, we prepare two logical qubits using the repetition code with flag qubits [16]. The quantum memory experiment consists of multiple rounds of syndrome extraction circuits, with the transversal CNOT gate placed between these rounds. We utilize auxiliary qubits to implement the transversal CNOT gate and map the structure onto a heavy-hexagon configuration. Our demonstrations are performed on the `ibm_sherbrooke` and `ibm_torino` quantum devices [43], which support structures ranging from 21 to 57 physical qubits, as illustrated in Fig. 1. We collect sampled data from the devices by executing the designed quantum circuit and then use a decoder to correct the final states of the two logical qubits. Logical error rates are computed by the expectation values of the logical Pauli gates known as the infidelity. Our results demonstrate that although error propagation among logical qubits can occur with the transversal CNOT gate, increasing the number of physical qubits can suppress errors by bit- or phase-flip error detection and correction on one of the IBM quantum processors. The error suppression remains consistent across ten cycles of the syndrome extraction circuit.

The paper begins by explaining the repetition code, which utilizes flag qubits, along with its syndrome extraction circuit. Next, we illustrate the transversal CNOT gate employing auxiliary qubits between two logical qubits, detailing its structure and discussing the quantum memory experiment for sampling data on the devices. Subsequently, we present a syndrome graph for error correction in the sampled data. Additionally, we visualize errors using a correlation matrix based on the data. As a result, we demonstrate the logical error probability of the transversal CNOT gate as a function of the number of physical qubits and the rounds of the syndrome extraction circuit. Finally, we summarize our findings and conclude.

II. METHOD

A. Repetition with flag qubits

Physical qubits used in stabilizer codes can be categorized as data and syndrome qubits. While data qubits are used to encode quantum information, syndrome qubits detect errors. Stabilizer operators can be defined based

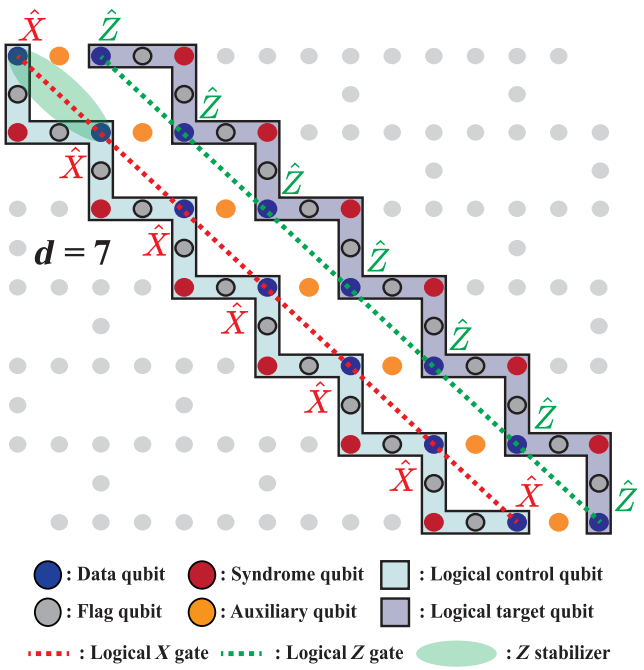

FIG. 1. Mapping the structures for $d = 7$ onto IBM's hardware. The transversal gate is designed to implement the logical CNOT gate on two logical qubits: the logical control (sky blue area) and the target qubit (purple area). Each logical qubit has data, flag, and syndrome qubits corresponding to the blue, gray, and red dots, respectively. The structure can be mapped onto IBM's quantum hardware (127 physical qubits) for different distances $d \in \{3, 5, 7\}$, corresponding to structures with 21, 39, and 57 physical qubits including auxiliary qubits for implementing the transversal CNOT gate.

on the Pauli operators of data qubits, and a syndrome extraction circuit is built using these stabilizers [44–47]. Additionally, flag qubits can be employed in the syndrome extraction circuit [48,49]. An additional flag qubit can be introduced to detect multiqubit errors in the data resulting from a single fault during the execution of the syndrome extraction circuit [23,35,36,41], or to address hardware limitations in connectivity [50–54], such as IBM's heavy-hexagon quantum hardware structure [2].

We construct a logical qubit using the repetition code with flag qubits. A single logical qubit consists of data, syndrome, and flag qubits, depicted in Fig. 1. All dots represent physical qubits; these include blue, red, and gray dots corresponding to data, syndrome, and flag qubits. The data and syndrome qubits are alternately placed but physically segregated, with a flag qubit filling the gap between them. The data qubits are encoded in the quantum state of a single logical qubit. Their state carries the observable information of interest and is subject to change through a logical gate. On the other hand, flag and syndrome qubits detect errors that may affect the state of data qubits. When the number of data qubits is d, known as a distance for the repetition code, the structure requires $d - 1$ syndrome and $2(d - 1)$ flag qubits.

Protecting a logical qubit against errors requires a stabilizer group constructed by a series of the Pauli operators of data qubits. The stabilizer group constrains the state of the data qubits to a logical state ($|\psi\rangle_L$) and enables the detection of errors that could lead the state to deviate from the logical state. The logical Pauli operators ($\hat{X}_L, \hat{Z}_L$) commute with all stabilizers. Hence, the logical Pauli operators change the state of the logical qubit without causing an error. A simple way to define them is by applying Z (X) operators on all data qubits, thereby implementing the logical Z (X) gate.

In Fig. 1, the stabilizers can be defined by the Z Pauli operators of a pair of neighbor data qubits as $\hat{Z}_i\hat{Z}_{i+1}$, where i denotes a data qubit number. When a logical qubit contains three data qubits, the logical qubit can be protected from a bit-flip error and defined in the logical states ($|0\rangle_L = |000\rangle, |1\rangle_L = |111\rangle$). The logical Pauli operators are chosen as $\hat{X}_L = \hat{X}_1\hat{X}_2\hat{X}_3$ and $\hat{Z}_L = \hat{Z}_1\hat{Z}_2\hat{Z}_3$. In our work, we use X stabilizers to prepare the logical state in the X basis, which can detect and correct phase-flip errors that cause changes in its logical state. More details of the X stabilizers can be found in Appendix C 1. It is worth noting that the repetition code can correct only one type of error by the chosen stabilizer group, for example, correcting phase-flip errors with X stabilizers or correcting bit-flip errors with Z stabilizers.

A quantum circuit subroutine known as a syndrome extraction circuit detects errors in a logical qubit. The syndrome extraction circuit is designed and performed based on the stabilizers. The syndrome extraction circuit using Z stabilizers is constructed as in Fig. 2(b). First, syndrome qubits are prepared as $|+\rangle$ through the application of Hadamard gates, and flag qubits are initialized as $|0\rangle$ ($T = 1, 2$). Secondly, each physical qubit is involved in a two-qubit gate configuration with its neighbor qubit, and the gates are applied according to a particular sequence in parallel. Two-qubit gates are only implemented between a single physical qubit and its nearest adjacent qubit at a hardware level ($T = 3$–6). Finally, we measure every flag and syndrome qubit in the Z basis ($T = 7, 8$). The syndrome extraction circuit can be repeated multiple times, and the outcomes are mapped to a binary vector (a syndrome) that can be used as input to a decoder algorithm.

In the absence of errors, the measurement outcomes from flag and syndrome qubits after syndrome extraction circuits are all "0" bits, consistently yielding the same values across multiple rounds. In this case, the syndrome is a zero vector. However, each component operation, including measurements, the initialization to the $|0\rangle$ state, as well as the one- (H) and two- [CNOT and controlled-Z (CZ)] qubit gates, are susceptible to errors during execution on quantum devices. Depending on the errors that occur as a quantum device executes syndrome extraction circuits,

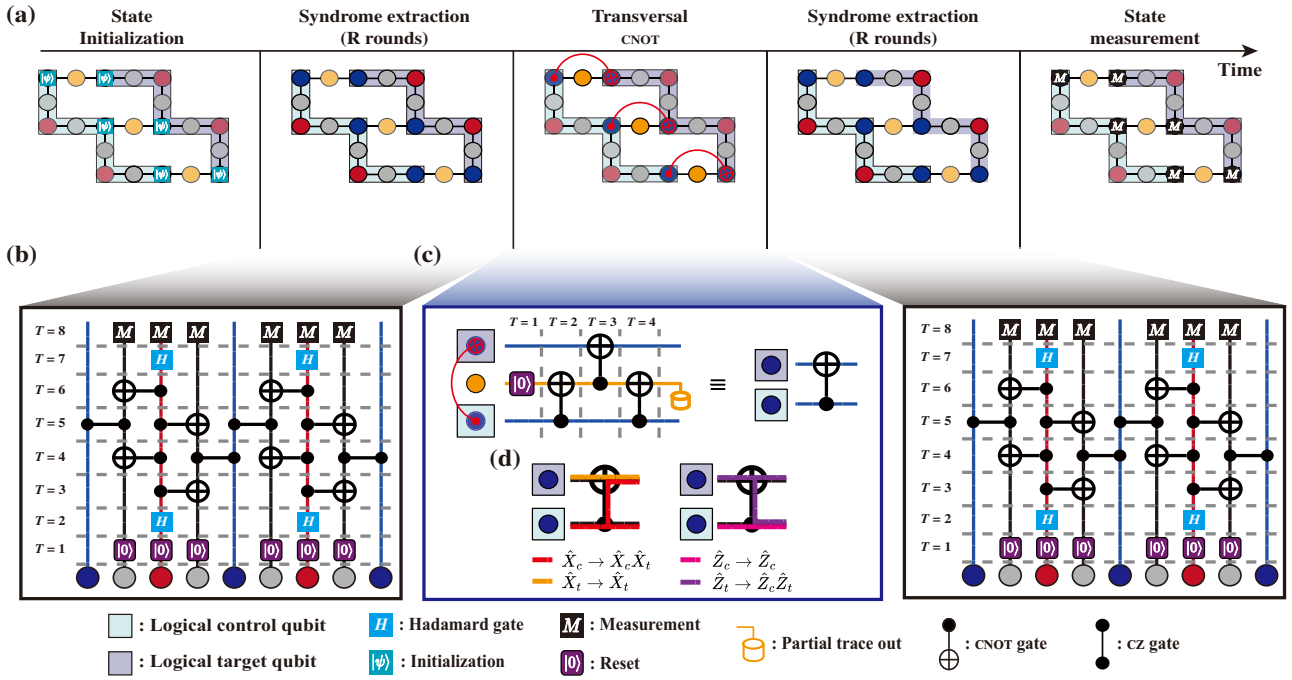

FIG. 2. Quantum memory experiment with the transversal CNOT gate. (a) A quantum circuit used for sampling data in the case of a 21-physical-qubit structure is depicted. The quantum circuit has five steps. First, two logical qubits are encoded via data qubit initialization. (b) Secondly, the syndrome extraction circuit is implemented **R** times for each logical qubit with eight steps. The subroutine circuit creates interaction between data qubits and flag and syndrome qubits to detect errors from their measurement outcomes. Each subroutine circuit measures Z stabilizers and that of measuring X stabilizers can be implemented by changing from CZ to CNOT gates, detailed in Appendix C 2. (c) Next, the transversal CNOT gate is executed with three physical CNOT gate layers employing auxiliary qubits. The CNOT gates are applied only between data qubits from two logical qubits and the auxiliary qubits between them. (d) While physical qubits undergo a physical CNOT gate, an X or Z Pauli gate can propagate from one physical qubit to another. Subsequently, another **R** rounds of the syndrome extraction circuit are considered. Finally, every data qubit is measured in the Z basis.

they flip the measured outcomes of either flag or syndrome qubits, changing them from $|0\rangle$ to $|1\rangle$ or vice versa. If measured outcomes of qubits are flipped compared to the previous round and have odd parity, it is considered a detection event resulting in a "1" bit in the syndrome [14]. Depending on the errors that occur during the execution of quantum gates, the syndrome maps onto a corresponding binary vector.

B. Quantum memory experiment

The transversal CNOT gate is applied to two logical qubits: a logical control and a target qubit. The two logical qubits are prepared with the repetition codes employing flag qubits. In Fig. 2(a), the sky blue and purple boxes represent the logical control and target qubits. Each logical qubit has three data qubits, and the same number of auxiliary qubits (orange circles) is shared.

The quantum memory experiment with the transversal CNOT gate proceeds as follows. First, we encode the two logical qubits by initializing data qubits in either the Z or X basis depending on the given logical states. We prepare both logical qubits to be initialized in the same basis. Second, we implement **R** rounds of the syndrome extraction circuit corresponding to the basis of the logical qubit, following the quantum gate sequences shown in Fig. 2(b) for each logical qubit, as discussed in Sec. II A. Next, we use auxiliary qubits and apply the transversal CNOT gate. Following the execution of the transversal CNOT gate, we perform the syndrome extraction circuit for **R** rounds. In total, the syndrome extraction circuit is executed **2R** times. Finally, we measure all data qubits in the Z basis. The structure for the transversal CNOT gate can be mapped onto the heavy hexagon structure and implemented on an IBM device. When each logical qubit consists of d data qubits, the structure requires $3d + 6(d-1)$ physical qubits, including data, syndrome, flag, and auxiliary qubits.

The method for implementing the transversal CNOT gate is to execute a physical CNOT gate between locally corresponding data qubits. The physical CNOT gate is executed with the data qubit from the logical control qubit serving as the control qubit and the data qubit from the other logical qubit serving as the target qubit. The transversal CNOT gate implements the physical CNOT gate within each pair of grouped data qubits directly. Figure 2(a) shows

two data qubits, the closest from each logical qubit, and the execution of a series of physical CNOT gates for the transversal CNOT gate in parallel. However, owing to the heavy hexagon structure of IBM hardware, the data qubits from each logical qubit are not the nearest neighbors but have a physical qubit between them. To circumvent this issue, a physical qubit is used as an auxiliary qubit to enable paired data qubits to interact for the transversal CNOT gate.

To implement a transversal CNOT gate, three physical CNOT gates are applied. Figure 2(c) shows the sequence of quantum gates for implementing the transversal CNOT gate, specifying the order of physical CNOT gates. First, we initialize an auxiliary qubit as $|0\rangle$ ($T = 1$). Next, we apply the physical CNOT gate by choosing a data qubit of the logical control qubit as a control qubit and the auxiliary qubit as a target qubit ($T = 2$). Another physical CNOT gate is applied between the auxiliary qubit and a data qubit of the logical target qubit ($T = 3$). The data qubit of the logical control qubit and the auxiliary qubit are selected for the control and target qubits, and we execute the last physical CNOT gate ($T = 4$). The auxiliary qubit is partially traced out after the physical CNOT gates have been implemented. In the demonstrations, we leave the auxiliary qubits idling without applying any further quantum gates or measurements after implementing the CNOT gates for the transversal CNOT gate. The state of the two data qubits after the series of quantum gates is the same as that resulting from the direct execution of a physical CNOT gate between them ideally [40,54]. Implementing the three layers of physical CNOT gates employed for the transversal CNOT gate on the data qubits from two logical qubits, the gates within each layer are applied in parallel.

When a physical CNOT gate is applied, a physical Pauli operator can be transferred, as shown in Fig. 2(d). For an X gate on the control qubit before the physical CNOT gate, the resulting circuit is equivalent to X gates applied to both the control and target qubits after the CNOT gate (red line). Conversely, if the gate occurs on the target qubit, it remains isolated and does not affect the other qubit (orange line). In the case of a Z gate on the control qubit, there is no gate propagation (pink line), whereas the gate on the target qubit can lead to two Z gates applied to both qubits after the CNOT gate (purple line).

The transversal CNOT gate is designed with a series of physical CNOT gates and enables the propagation of X or Z gates between each paired data qubit. Hence, the logical Pauli operators ($\hat{X}_L, \hat{Z}_L$) formed using the Pauli operators of data qubits transversally follow the same gate propagation rules as their physical-level counterparts during the execution of the transversal CNOT operation at the logical level. However, a bit- or phase-flip error on a data qubit before the transversal CNOT gate can introduce an additional error on the interacted data qubit. In other words, the transversal CNOT gate also allows for the propagation of errors among the logical qubits.

Correlated errors can arise and lead to ambiguity in calculating a correction operator by a decoder. This occurs not only because of errors within quantum hardware as it executes a quantum circuit [22,55,56], but also due to error propagation between two logical qubits through the transversal CNOT gate [30,38,40,57]. Throughout this paper, we use the term "gate-flow errors" to refer to errors caused by the transversal CNOT gate, distinguishing them from correlated errors inherent to the quantum hardware.

C. Syndrome graph

As we collect the data by running a quantum memory experiment on a quantum processor, we get the corresponding measured outcome of physical qubits based on errors that occurred as it is executed. The outcome can be mapped on the binary vector, a syndrome, considering space and time. The syndrome is calculated by checking the parity using the outcomes of each case sample as input. A detection event, resulting in a "1" syndrome bit, occurs when the measurement outcomes of a syndrome extraction circuit involving flag and syndrome qubits differ from those of the previous extraction circuit and have odd parity [16].

An error that produces two detection events can be categorized into four types: space, time, space-time, and gate-flow errors. Any arbitrary error can be decomposed into a combination of these error categories [58]. Figure 3(a) shows corresponding bit-flip errors in a syndrome extraction circuit using Z stabilizers, where blue, red, green, and orange boxes correspond to space, time, space-time, and gate-flow errors, respectively. The hexagon boxes with letters correspond to α and β for time errors, γ for a space error, δ for a space-time error, and ϵ for a gate-flow error in Figs. 3(a) and 3(b). A time error occurs when errors impact a syndrome or flag qubit, while other error types affecting data qubits differ based on when and where the error occurs within a subroutine of a quantum circuit.

Figure 3(b) illustrates how the measurement outcomes of flag and syndrome qubits change due to these categorized errors. The figure provides an example of outcomes from multiple rounds of syndrome extraction circuits, organized according to their execution times, with three data qubits for each logical qubit. A space error affects neighboring syndrome qubits horizontally, while a time error results from an error during the preparation of syndrome or flag qubits, causing outcome changes vertically. There is another type of error, the space-time error. The edge corresponds to an error that generates a pair of outcome changes diagonally during two-qubit gates in a syndrome extraction circuit on a data qubit. The case when an error spreads from one to the other logical qubit produces

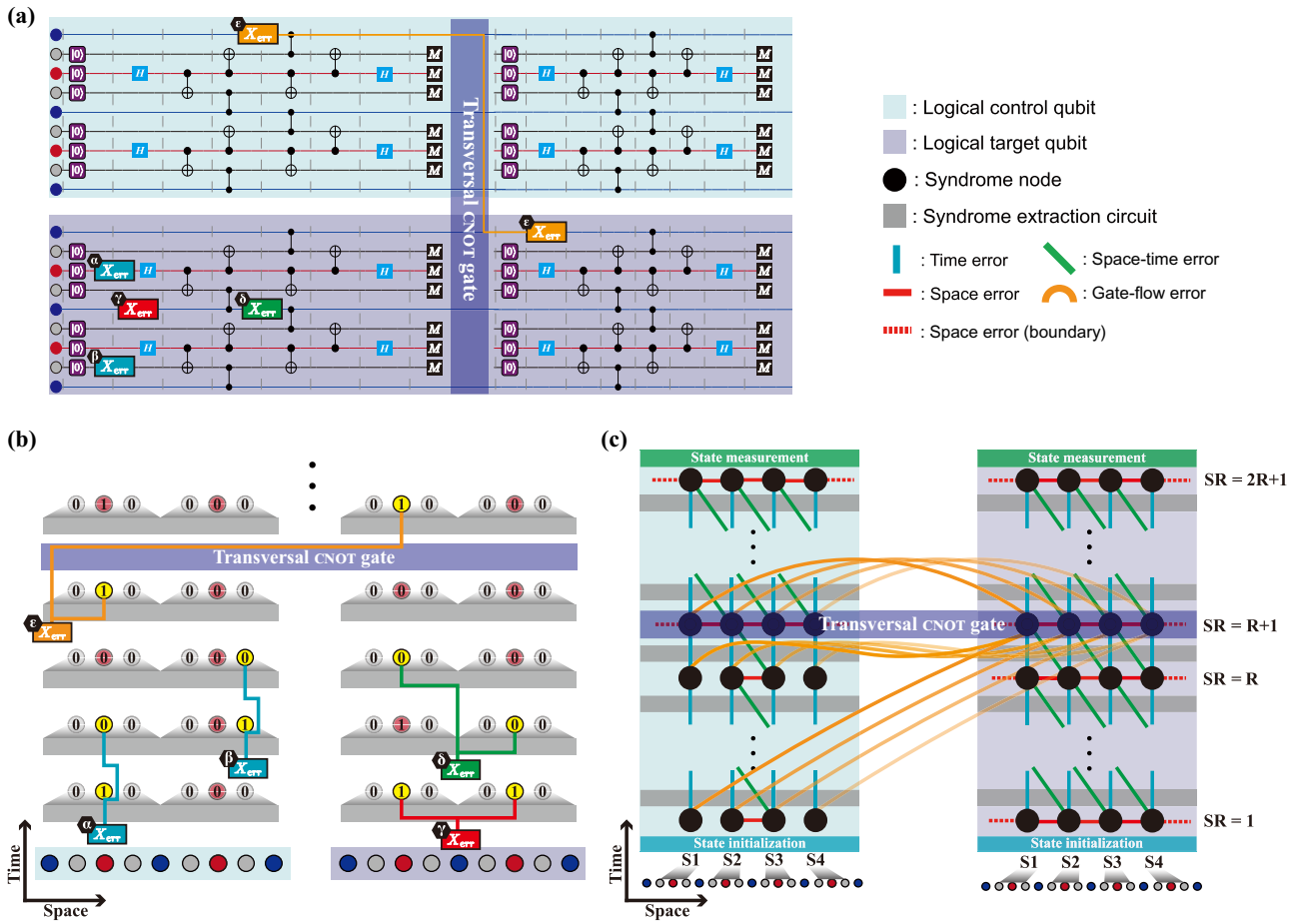

FIG. 3. Syndrome graph for two logical qubits considering the transversal CNOT gate. (a) Errors can be categorized into four types: space (red box), time (blue box), space-time (green box), and gate-flow errors (orange box). The letters within the hexagon boxes are assigned to each of these errors: α and β for time errors, γ for a space error, δ for a space-time error, and ϵ for a gate-flow error. These errors are depicted within a syndrome extraction circuit for a distance of 3. (b) An example showcases the outcomes of syndrome extraction circuits using 21 physical qubits, where an error can influence the outcomes of syndrome or flag qubits, affecting neighboring qubits horizontally, vertically, or diagonally. Additionally, a gate-flow error affects two syndrome qubits on both logical qubits. (c) The syndrome graph for the $d = 5$ structure with **2R** rounds of syndrome extraction using Z stabilizers is depicted. The horizontal and vertical axes represent the space of the syndrome nodes (**S1**, **S2**, **S3**, and **S4**) and syndrome rounds (**SR**). The logical CNOT gate is executed at $\mathbf{SR} = \mathbf{R} + \mathbf{1}$, resulting in a total of $\mathbf{2R} + \mathbf{1}$ syndrome rounds. Each logical qubit produces four syndrome nodes per syndrome round. Edges represent the correlations due to the four categorized errors (space, time, space-time, and gate-flow errors). A space error on the boundary of each logical qubit connects with the boundary except on the logical control qubit before the transversal CNOT gate has been executed.

a gate-flow error. A gate-flow error produces outcome changes on syndrome qubits from both logical qubits. Two affected syndrome qubits due to a gate-flow error locally correspond to each logical qubit.

We utilize a syndrome graph to decode errors and evaluate the performance of the transversal CNOT gate across multiple rounds of the syndrome extraction circuit. A syndrome graph is a two-dimensional graph where each node and edge represents a syndrome bit or error, with the weight of each edge corresponding to the error probability [59–61,65]. The syndrome graph is constructed by considering the four categorized errors, enabling us to decode errors and calculate a correction operator.

Figure 3(c) depicts the syndrome graph for a case where the syndrome extraction circuits are applied **2R** times, featuring a structure comprising five data qubits for the Z basis of each logical qubit. The graph consists of two regions: a logical control and a target qubit. The horizontal and vertical axes represent space and time, respectively. During each syndrome round, every logical qubit generates four syndrome bits: **S1**, **S2**, **S3**, and **S4**. Every syndrome bit is obtained by checking the parity with the measured outcomes of physical qubits while respecting their position and time within the code. Parameter **SR** implies a syndrome round in the graph obtained by comparing two consecutive results of syndrome extraction circuits. The

syndrome extraction circuits used in the quantum memory experiment are depicted as gray boxes, showing their relative implementation timing concerning syndrome rounds. The transversal CNOT gate is implemented at the time $\mathbf{SR} = \mathbf{R} + \mathbf{1}$. Executing the syndrome extraction circuits for $\mathbf{R}$ rounds both before and after the transversal CNOT gate results in a total of $\mathbf{2R} + \mathbf{1}$ syndrome rounds. These include the first and last temporal boundaries, where we consider the state of initial and final measured data qubits. An edge represents an error that produces two detection events on the connected syndrome nodes by the edge. Each edge corresponds to one of the categorized errors in the syndrome graph. The edges in the logical control qubit, before the transversal CNOT gate is applied, as shown in Fig. 3(c), differ from those in the logical target qubit. Specifically, the space and space-time edges connecting to the boundary syndrome nodes of the logical control qubit are omitted. This is because the way we handle errors in logical control and target qubits leads to different error combinations in the syndrome graph. For example, if an error in the logical control qubit propagates to the logical target qubit, it can produce up to four detection events. The interpretation of these errors depends on how the four detection events are decomposed into two categories: (1) gate-flow errors and (2) local errors (space and space-time errors). The edges connecting to the boundary syndrome nodes of the logical control qubit in the syndrome graph are primarily decomposed using gate-flow errors rather than local errors, which lead to the omission of some possible combinations in the syndrome graph. The weight of each edge in a syndrome graph corresponds to its error probability, calculated based on the error rates of individual single- and two-qubit gates according to the hardware specifications. We consider a circuit-level noise model as detailed in Appendix D 1. These calculations are performed using the Stim software package [64].

It is important to note that error propagation occurs only when an error happens within a time no later than the execution of the transversal CNOT gate. Therefore, error propagation occurs only when an error happens within a time no later than time $\mathbf{SR} = \mathbf{R} + \mathbf{1}$ corresponding to the execution of the transversal CNOT gate. As a result, all gate-flow errors connect syndrome nodes before $\mathbf{SR} = \mathbf{R} + \mathbf{1}$ on the logical control qubit and the syndrome nodes on the logical target qubit at $\mathbf{SR} = \mathbf{R} + \mathbf{1}$. A space error on boundary data qubits of a logical qubit without error propagation produces only one detection event, necessitating a connection to the boundary [14–16].

Logical qubits preserve data qubit information by detecting errors through the measured outcomes of syndrome and flag qubits from the syndrome extraction circuit and decoding them for error correction. We utilized the minimum weight perfect matching (MWPM) algorithm to rectify errors based on the outcomes obtained from the syndrome extraction circuits and the corresponding syndrome graph [14,62,63]. Given a sample obtained by conducting a quantum memory experiment on a device, we recovered the measured final data qubits using a correction operator that most likely cancels errors using the Pymatching algorithm [65,66].

III. RESULTS

The state of two logical qubits can be denoted as $|\psi_C\psi_T\rangle_L = |\psi_C\rangle_L \otimes |\psi_T\rangle_L$, where $|\psi_C\rangle_L$ and $|\psi_T\rangle_L$ represent the states of logical control and target qubits. When the logical qubits are in the Z basis, the state can be one of four computational basis states $\{|00\rangle_L, |01\rangle_L, |10\rangle_L, |11\rangle_L\}$. Since a logical X operator can propagate from the logical control qubit to the logical target qubit through the transversal CNOT gate, these logical states should be transformed into $\{|00\rangle_L, |01\rangle_L, |11\rangle_L, |10\rangle_L\}$ after passing the quantum memory experiment. In the case of the logical qubits in the X basis, a logical Z operator can be propagated reversely. The logical states $\{|{+}{+}\rangle_L, |{+}{-}\rangle_L, |{-}{+}\rangle_L, |{-}{-}\rangle_L\}$ ideally change to $\{|{+}{+}\rangle_L, |{-}{-}\rangle_L, |{-}{+}\rangle_L, |{+}{-}\rangle_L\}$ following the logical CNOT gate. The demonstrations were performed with a total of eight possible initialized two-logical-qubit states in Z (X) basis for both logical qubits: four states in the Z basis ($\{|00\rangle_L, |01\rangle_L, |11\rangle_L, |10\rangle_L\}$), and four states in the X basis ($\{|{+}{+}\rangle_L, |{+}{-}\rangle_L, |{-}{+}\rangle_L, |{-}{-}\rangle_L\}$). We collect 10^5 samples by conducting quantum memory experiments on a quantum device for each round and initial state with the *qiskit* PYTHON package [67]. A single round of the syndrome extraction circuit took approximately 1.9 and 4.1 μs on `ibm_torino` and `ibm_sherbrooke`, respectively.

Based on detection events, error patterns that include gate-flow errors can be seen through a correlation matrix. The correlation matrix can be plotted with the probability of errors as a function of the labels of syndrome nodes, i and j, corresponding to an edge connecting the two syndrome nodes in the syndrome graph. Each error probability (p_{ij}) can be calculated as follows from the sampled data [15]:

$$p_{ij} = \frac{1}{2}\left(1 - \sqrt{1 - \frac{4(\langle x_i x_j\rangle - \langle x_i\rangle\langle x_j\rangle)}{(1 - 2\langle x_i\rangle - 2\langle x_j\rangle + 4\langle x_i\rangle\langle x_j\rangle)}}\right). \tag{1}$$

While $\langle x_i\rangle$ and $\langle x_j\rangle$ are the probabilities of the detection event at ith and jth syndrome nodes, $\langle x_i x_j\rangle$ is the probability of two detection events at the same time. The correlation matrix is symmetric since $p_{ij} = p_{ji}$. We set negative values to zero to enhance visibility and clarity in the visualization. The detection event probabilities, from the demonstrations on `ibm_torino` and `ibm_sherbrooke`, for each syndrome node respecting both time and space, are provided in Appendix E 3.

Figure 4 shows the correlation matrices for the case of the structure comprising 39 physical qubits

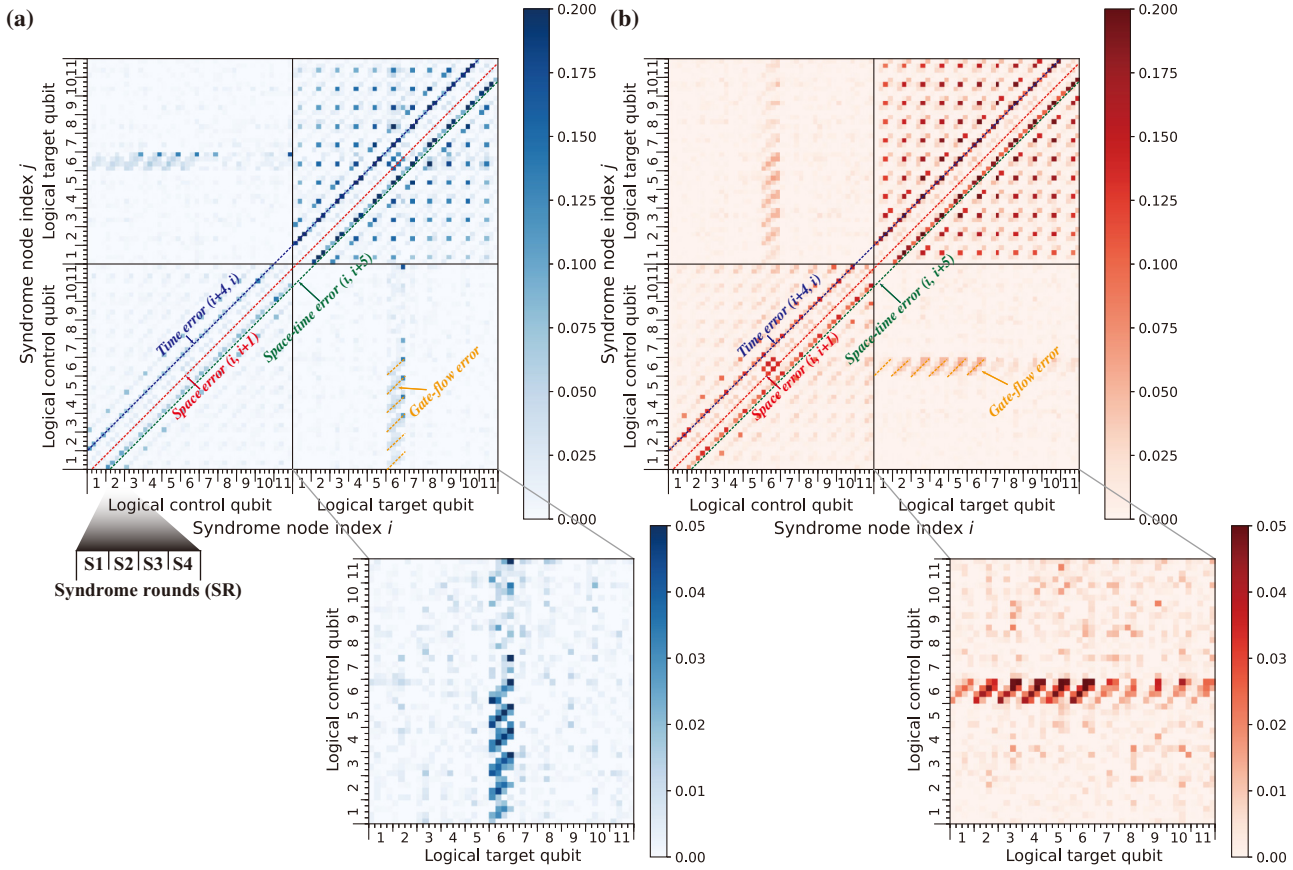

FIG. 4. Correlation matrices for the syndrome nodes (`ibm_sherbrooke`). The two logical states, the logical control and target qubits, are initialized in the $Z(X)$ basis, and each one has five data qubits. The data are obtained by sampling the quantum memory experiment 10^5 times for both (a) $|00\rangle_L$ and (b) $|{+}{+}\rangle_L$, with $\mathbf{2R = 10}$ rounds of the syndrome extraction circuit. The axes represent the 11 blocks corresponding to syndrome rounds for a logical qubit, with each block displaying four syndrome nodes (**S1**, **S2**, **S3**, and **S4**) arranged according to their locations within the code. A pixel represents the error probability that generates a "1" syndrome bit on the corresponding ith and j th syndrome nodes. Parameter **SR** corresponds to the syndrome round, and the transversal CNOT gate has been executed at $\mathbf{SR = 6}$. The probabilities of space, time, space-time, and gate-flow errors are visually represented by red, blue, green, and orange dotted lines, respectively. The color bar is truncated to 0.2 and 0.05 to enhance the visibility. Figure 9 in Appendix E 2 shows the same graph without truncation.

and $\mathbf{2R = 10}$ rounds of the syndrome extraction circuit both before and after the transversal CNOT gate. We show the correlation matrices obtained with data from `ibm_sherbrooke`, while the equivalent figures using data from `ibm_torino` employing 57 physical qubits are available in Appendix E 2. The figure shows the correlation matrices for both the X and Z bases with labeled syndrome nodes. We assign new labels to the number of syndrome nodes, using $i//d$ as the time index and $i\%d$ as the space index for each logical qubit, where i is the node number. Each correlation matrix can be divided into four regions depending on where i and j are: a pixel (i, j) that can belong to (1) the logical control qubit, (2) the logical target qubit, or (3),(4) the correlation region where i and j are in different logical qubits. There are 11 syndrome rounds, with each round producing four syndrome bits for every logical qubit. The transversal CNOT gate is performed at $\mathbf{SR = 6}$ following the five rounds of the syndrome extraction circuit.

There are spatial or temporal errors within each logical qubit region in the correlation matrix in Fig. 4 for the logical qubits in the (a) Z or (b) X basis. While measurement errors (time error) correspond to the pixels at $(i + 4, i)$, data qubit errors (space error) are represented by the pixels at $(i, i + 1)$. Space-time errors can be seen in the pixels at $(i, i + 5)$. These errors are shown with blue, red, and green diagonal dotted lines, respectively. While the probabilities of space-time errors are relatively lower than other types of errors, time errors show the high error probabilities among the categorized errors for all cases. This matches the hardware specifications, where the dominant errors are readout errors.

The pixels representing space errors at $\mathbf{SR = 6}$ are notably highlighted with high error probabilities after

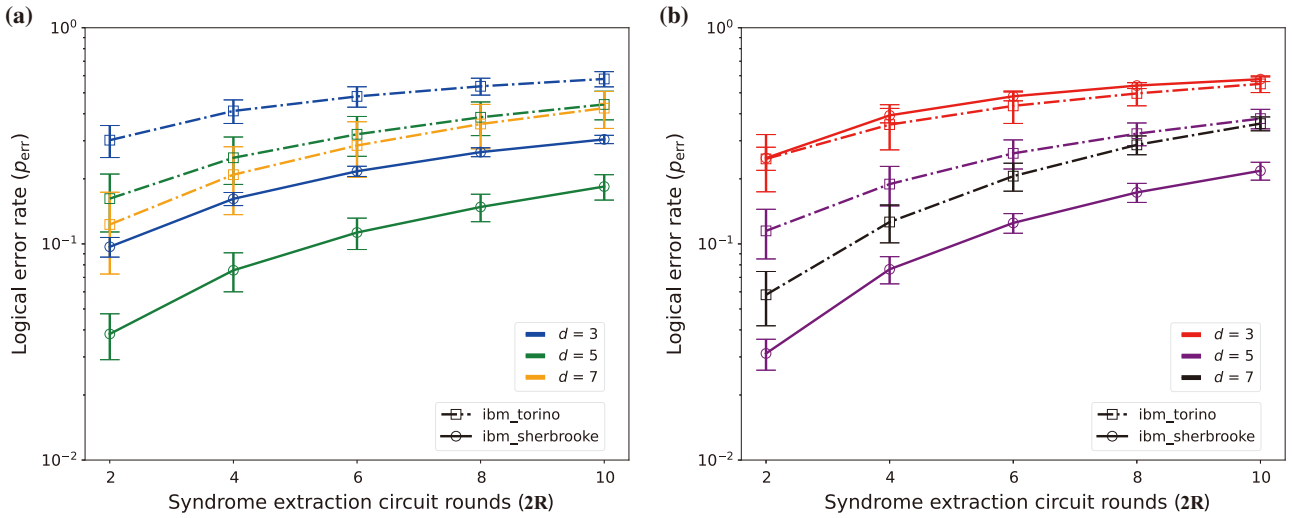

FIG. 5. Logical error rate suppression with code distance. The code considers (a) X errors using Z stabilizers or (b) Z errors using X stabilizers. The blue (red), green (purple), and orange (black) lines correspond to structure sizes of 21, 39, and 57 physical qubits correcting X (Z) errors, respectively. While solid lines denote the result from `ibm_sherbrooke`, dash-dot lines are those from `ibm_torino`. The average logical error rates for four basis states are depicted with their standard deviations as a function of the number of syndrome extraction circuit rounds (**2R**), devices, and structure size ($d \in \{3, 5, 7\}$). Each data point is obtained from 10^5 data samples collected from the corresponding devices. As the number of physical qubits increases, the average logical error rates decrease for both X and Z error cases.

executing the transversal CNOT gate. Based on the type of errors that are detected by logical qubits, the highly visible values are located at either the logical control or target qubit. For example, the error probabilities of data qubits in the logical target qubit at $\mathbf{SR} = \mathbf{6}$ are relatively high, in Fig. 4(a), where the logical qubits consider bit-flip errors. Conversely, regarding the Z errors in Fig. 4(b), spatial errors in the logical control qubit at $\mathbf{SR} = \mathbf{6}$ occur more frequently than at other times and spaces. From the perspective of the affected logical qubit, it is because additional errors are introduced from not only the other logical qubit through the transversal CNOT gate but also auxiliary qubits used to perform the transversal CNOT gate.

A gate-flow error can be seen in the area within a correlation between the logical control and target qubits. The gate-flow errors are visible in the correlation matrix in distinct ways, depending on the type of error and whether they are bit- or phase-flip errors. Gate-flow errors indeed exhibit conspicuous error probabilities, particularly at specific syndrome rounds. Notably, in Fig. 4, at $\mathbf{SR} = \mathbf{6}$, we observe this phenomenon prominently in the (a) logical target or (b) control qubit. Diagonal lines represent the dominance of gate-flow errors. These diagonal pixels represent the two syndrome nodes located at the same positions in each logical qubit, exhibiting strong correlation mainly due to error propagation while respecting spatial constraints. Additionally, notable correlations exist with locally neighboring syndrome nodes. This phenomenon suggests that the execution of the transversal CNOT gate can result in gate-flow errors, causing temporal correlations between syndrome nodes. Bluvstein *et al.* [38] demonstrated strong correlations with the stabilizers of two logical qubit blocks due to error propagation. However, the key difference lies in the number of syndrome extraction circuits. We not only show strong correlations due to gate-flow errors, but also highlight the existence of long temporal correlations between syndrome nodes by executing multiple rounds of a syndrome extraction circuit before the logical CNOT gate.

To evaluate the performance of the transversal CNOT gate, we compute the logical error rate (p_{err}), referred to as infidelity. This is determined by the expectation values of the logical Pauli operators for two logical qubits that yield a $+1$ eigenvalue for the ideally error-corrected final logical qubit states, following the methodology described in Ref. [35]. Appendix E 4 provides more details on calculating logical error rates. We consider the **2R** rounds of the syndrome extraction circuits (**R** rounds before and **R** rounds after the logical CNOT gate) and correct errors using the decoder. We obtain logical error rates for structures with distances of 3 and 5, requiring 21 and 39 physical qubits on `ibm_sherbrooke`, respectively. Despite the device's structure allowing for a mapping distance of 7, the presence of broken edges between physical qubits in a specific region, which means that two physical qubit gates have an error rate of 1, makes it challenging to execute quantum circuits, as shown in Appendix F, Fig. 24. Additionally, we evaluate the performance of

the transversal CNOT gate on ibm_torino for structures comprising 3, 5, and 7 data qubits per logical qubit. Demonstrations are performed for each round, considering initial states for both the Z and X bases. Logical error rates are evaluated for each device based on collected samples.

Figure 5 shows the logical error rates averaged across all four states in $Z\,(X)$ basis, along with their standard deviations. More details about the logical error rate results for each logical qubit state are provided in Appendix E 4 along with their simulated results using Stim. The logical error rates of the transversal CNOT gate are plotted as a function of the structure size and the number of syndrome extraction circuits for each device. The results demonstrate a reduction in the average logical error probabilities as the code size increases when the transversal CNOT gate is executed with syndrome extraction circuits. This trend remains consistent across all rounds of syndrome extraction circuits. Notably, among the selected physical qubits for demonstrations, the average error rates of two-qubit gates on ibm_torino are higher than those on ibm_sherbrooke, whereas the average readout error rates on ibm_sherbrooke are higher than on ibm_torino. More details can be found in Appendix E 1. Furthermore, according to the correlation matrices in Appendix E 2, ibm_torino has experienced more correlated errors, which inherently occur as the real device executes a quantum circuit, such as leakage errors [56]. Given that readout and correlated errors are the predominant source of errors, these device properties may contribute to higher logical error rates on ibm_torino compared to ibm_sherbrooke.

The gap in logical error rates between different structure sizes diminishes as the number of rounds of a syndrome extraction circuit increases. Notably, this phenomenon is particularly pronounced in the results from ibm_torino. This trend could be attributed to the increased rounds of syndrome extraction per logical CNOT gate, which may introduce greater ambiguity in the correction operator calculated by the decoder owing to gate-flow errors [57].

IV. CONCLUSIONS AND DISCUSSIONS

The scalability of fault-tolerant quantum computers constructed with stabilizer codes relies on the potential of error suppression achieved by increasing the size of their code block. This property must be demonstrated in both logical qubits and gates during error correction. However, demonstrating error suppression of a logical CNOT gate with multiple rounds of error correction remains an open question in the NISQ era.

In this study, we evaluated the performance of the logical CNOT gate with multicycle syndrome extraction circuits both before and after the gate ($\{\mathbf{R} = \mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4}, \mathbf{5}\}$), for a total $\mathbf{2R}$ rounds, and compared their effectiveness across different sizes ($d \in \{3, 5, 7\}$). We prepared logical qubits using the repetition code with flag qubits and implemented the logical CNOT gate using physical CNOT gates transversely, leveraging auxiliary qubits. We performed demonstrations on the IBM quantum machines ibm_sherbrooke and ibm_torino. The results show that despite the emergence of gate-flow errors caused by the transversal CNOT gate, which creates long temporal correlations, increasing the number of physical qubits on IBM quantum processors suppresses errors. Furthermore, error suppression persists during multiple rounds of the syndrome extraction circuit. Additionally, we also performed demonstrations on ibm_brisbane, another IBM hardware platform. We observed lower logical error rates as we increased the code size from three data qubits to seven per logical qubit by correcting bit-flip errors. However, in contrast, when correcting phase-flip errors, the logical error rates with seven data qubits were higher than those with three and five data qubits. Further details are given in Appendix E 4.

It is important to note that the repetition code is unable to rectify both types of errors (bit and phase flips) simultaneously. Furthermore, additional flag qubits in the repetition code may lower the effectiveness of the distance that is closely related to the number of correctable errors [16]. However, this study primarily focused on investigating the transversal CNOT gate with varying structure sizes, detecting errors multiple times, and demonstrating error suppression through increased code size. Therefore, the study showed that the transversal CNOT gate can be executed while correcting one type of error on real quantum hardware.

This work highlights an additional source of error (gate-flow errors) arising not only from the noisy quantum gates themselves, but also from the method of creating entanglement between logically encoded qubits by implementing a transversal CNOT gate. This implies that a stabilizer code capable of correcting both bit- and phase-flip errors will experience bidirectional error propagation, which could adversely affect its performance: bit-flip errors will propagate from the control logical qubit to the target logical qubit, while phase-flip errors will propagate in the reverse direction. While the transversal CNOT gate has been designed to cause fewer gate-flow errors by only interacting locally between corresponding data qubits of two logical qubits, these errors can still negatively impact logical qubits during execution on quantum hardware. This issue has also been discussed in Refs. [38,40,57].

Cain *et al.* [57] addressed the challenge of reducing detrimental effects by employing a decoder that utilizes a hypergraph for decoding complex errors resulting in more than two detection events. Furthermore, they investigated how the performance of a series of logical gates changes with varying numbers of syndrome extraction circuits per

logical gate. A promising avenue for future work would involve comparing the performance of the logical CNOT gate on real quantum devices using a correlated decoder versus the MWPM decoder, examining how this performance varies with the number of error detection cycles to find the optimized number of cycles. It would be enlightening to see whether a correlated decoder can also handle more complex gate-flow errors, which move through logical gates over time. Another line of investigation would be to train and apply machine learning–based decoders that have recently shown promising performance [68–70].

Finally, an open question for future research involves observing gate-flow errors occurring simultaneously in both directions between logical qubits. Notably, a logical Bell state has been successfully constructed using quantum hardware across different platforms [38–40]. Therefore, a promising approach is to construct a logical Bell state and calculate detection event probabilities for each syndrome bit across multiple cycles of syndrome rounds. It would be intriguing to observe how the detection event probabilities of X and Z stabilizers change before and after the execution of the transversal CNOT gate and whether the existence of temporal correlations occurred by gate-flow errors.

ACKNOWLEDGMENTS

Y.K. acknowledges support from the CSIRO Research Training Program Scholarship and a University of Melbourne Research Training Scholarship. The University of Melbourne supported the research through the establishment of the IBM Quantum Network Hub.

The authors declare no competing financial or nonfinancial interests.

DATA AVAILABILITY

The data that support the findings of this study and the source code used to generate figures in this work can be provided upon reasonable request to the corresponding author.

APPENDIX A: QUANTUM HARDWARE AND THE CLASSICAL SIMULATOR

The demonstrations were performed using three different hardware platforms, including processors covered in the main text. These platforms are `ibm_sherbrooke`, `ibm_brisbane`, and `ibm_torino`. In these hardware setups, quantum gates are executed using a basis gate set that ensures universality. IBM's hardware uses the following basis gates to execute arbitrary quantum circuits: X, $\sqrt{X}$, R_Z (single-qubit gates), and a two-qubit gate such as the CZ or CNOT gate. Additionally, some platforms use specialized gates like the ECR pulse (echoed cross-resonance gate) to achieve gate universality and create

TABLE II. List the hardware specifications used in the demonstrations. The values are obtained from the calibration table, where 21 physical qubits are used with $\mathbf{2R = 10}$ rounds of syndrome extraction circuits in the Z basis. The table displays the type of two-qubit gates used in each hardware, including ECR pulse, CZ, or CX gates. It also depicts the time length required for the execution of a single-qubit (SQ) gate, two-qubit (TQ) gate, and measurement. Furthermore, it shows the median error rate for each gate. The demonstrations on each device were performed in the duration listed in the table.

		Gate time (ns)			Median error rate (median)			Demonstration time
Hardware	TQ gate type	SQ gate	TQ gate	Measurement	SQ gate	TQ gate	Measurement	(YYYY-MM-DD, 00:00:00 + 00:00)
`ibm_sherbrooke`	ECR	56.88	533.33	1244.44	2.27×10^{-4}	7.72×10^{-3}	1.10×10^{-2}	2024-03-15 06:20:36–11:30:28
`ibm_brisbane`	ECR	60	660	4000	2.29×10^{-4}	8.02×10^{-3}	1.33×10^{-2}	2024-04-23 01:48:20–03:51:17
`ibm_torino`	CZ	32	84	1560	2.78×10^{-4}	4.78×10^{-3}	1.93×10^{-2}	2024-04-25 03:38:43–06:23:11

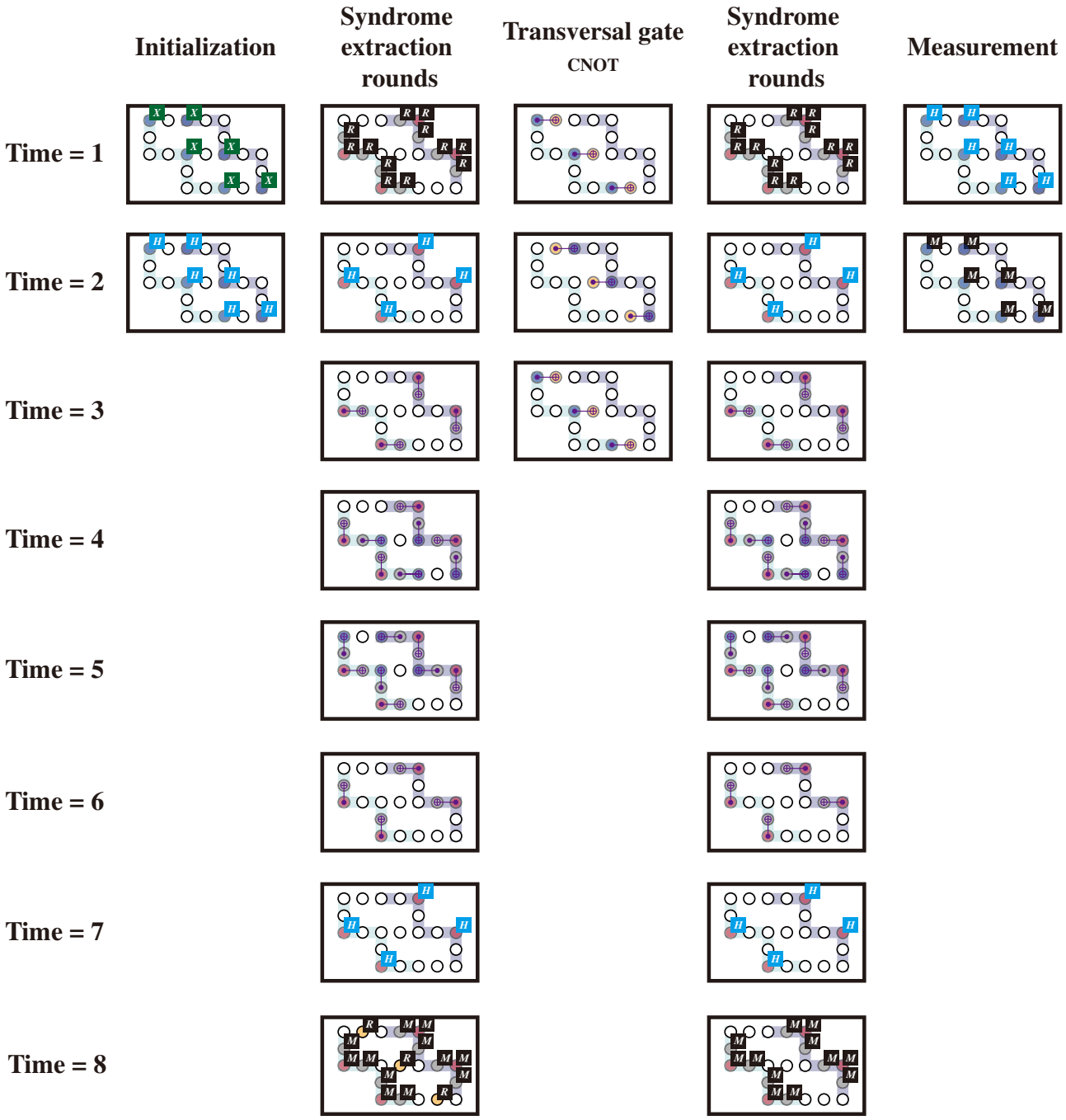

FIG. 6. Quantum memory experiment for sampling $|--\rangle_L$ with 21 physical qubits. Each node represents the physical qubits with blue, red, orange, and gray nodes corresponding to data, syndrome, auxiliary, and flag qubits. The quantum circuit consists of five steps: initialization, two syndrome extraction rounds, transversal CNOT gate, and measurement. The horizontal line represents each step, and the vertical line corresponds to the sequence of gates and time. When physical qubits are inactive and undergoing free evolution, the color of the node is filled white.

entanglement between pairs of physical qubits [71]. For example, `ibm_sherbrooke` and `ibm_brisbane` use the ECR pulse as their basis gate, enabling the construction of physical CNOT and CZ gates using single-qubit and ECR gates. Conversely, `ibm_torino` uses CZ gates to achieve quantum gate universality.

The hardware specifications, including the relaxation time (T1) and coherence time (T2) of each physical qubit, are listed in a calibration table. The calibration table data for an IBM quantum processor can be accessed via the cloud. This table also provides error rates for both single- and two-qubit gates, as well as for qubit measurements.

All physical qubit gates, except for the virtual Z gate (R_Z), require a specific execution time, which is detailed in the calibration table. Periodically, this calibration table is updated with new values. Each dataset in this work is analyzed using the corresponding device's calibration table to assess logical error rates, taking into consideration the completion time of each demonstration.

Table II presents the implementation duration for demonstrations on each hardware platform, along with gate times and median error rates for physical gates. The calibration table is derived from demonstrations utilizing 21 physical qubits with $\mathbf{2R = 10}$ rounds of a syndrome extraction circuit. Among real quantum hardware platforms, `ibm_torino` stood out with the shortest gate time and the lowest median error rate for its two-qubit gates; however, it had the highest median readout errors. In contrast, `ibm_sherbrooke` demonstrated the shortest time required and the lowest median readout errors for qubit measurements, while `ibm_brisbane` required the longest time for qubit measurements.

APPENDIX B: QUANTUM CIRCUIT OPTIMIZATION

We utilized the PYTHON library *qiskit* and transpiled circuits to execute the quantum circuits on the device. To optimize the quantum circuits and obtain results, at the hardware level, we employed the following methods.

(1) *Dynamic decoupling method (XXXX).* The method involves allowing free evolution of data qubits while syndrome and flag qubits are being measured. However, this can lead to idling errors for the data qubits. To mitigate this, we apply the dynamic decoupling method to all data qubits, minimizing their idling error [72].

(2) *Conditional reset gate.* After each round of the syndrome extraction circuit, flag and syndrome qubits are initialized to $|0\rangle$. To efficiently achieve this, we use conditional reset gates that incorporate conditional bit-flip operations based on the measured outcomes immediately after qubit measurements, reducing the resetting time [51].

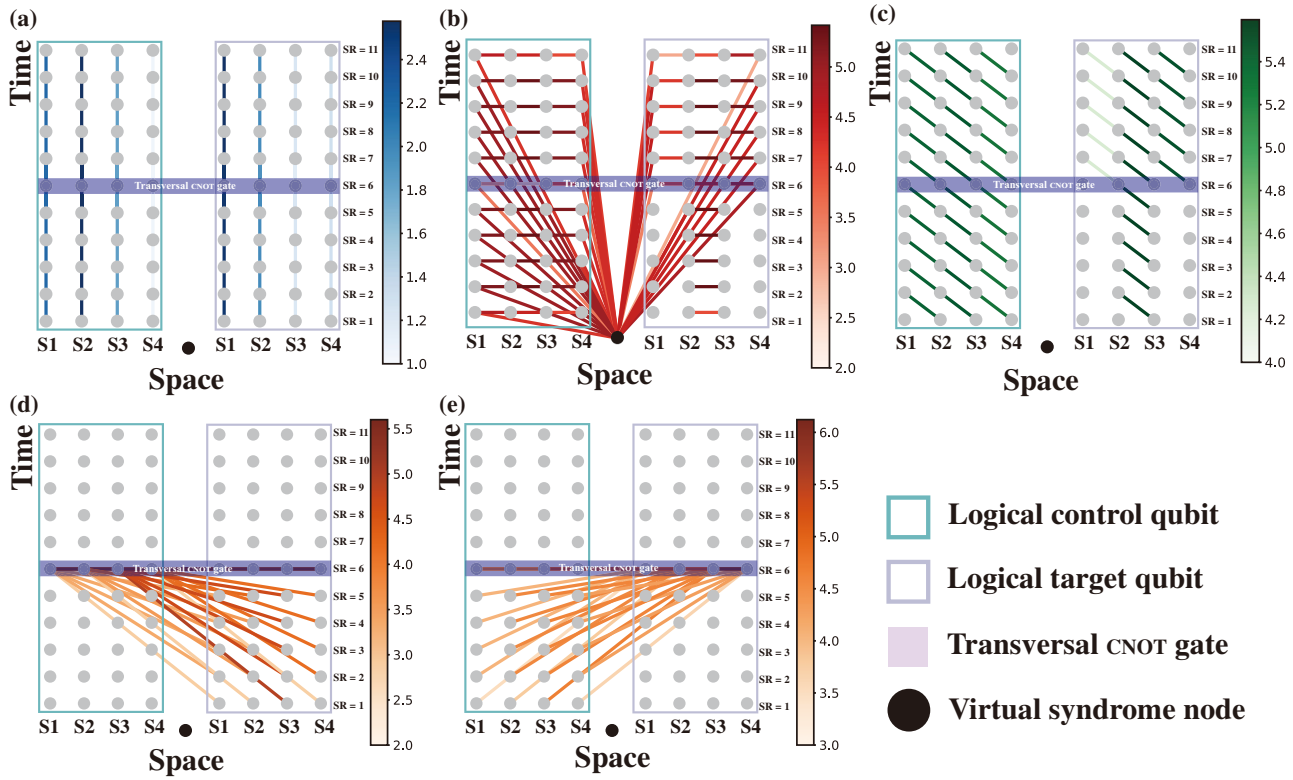

FIG. 7. Syndrome graphs with the types of edges for the structure with 39 physical qubits. The logical control and target qubits are enclosed within the blue and purple open boxes. The axes represent the space of the syndrome bit and the time of the syndrome round (**SR**). The figure is the case where the syndrome extraction is implemented five times, producing 11 syndrome rounds. The transversal CNOT gate is implemented at $\mathbf{SR = 6}$ and depicted as the filled purple box. Every node corresponds to the syndrome bit arranged by their location and time. Edges represent (a) time, (b) space, (c) space-time, and (d),(e) gate-flow errors and they have their weights corresponding to their error probabilities. The black dot represents the virtual syndrome node, used to display the S errors on the boundary of the logical qubit. The decoder considers gate-flow errors connecting one syndrome node, located on the logical (d) target or (e) control qubit before the transversal CNOT operation, to a syndrome positioned on the other logical qubit at $\mathbf{SR = 6}$. These scenarios correspond to cases where the logical qubits are initialized in Z (X) basis states.

TABLE III. Calibration table values for 21 physical qubits used for demonstrations (`ibm_sherbrooke`). The values obtained from the calibration table correspond to $|00\rangle_L$ with ten rounds. The table lists the properties of physical qubits for the $d = 3$ structure selected within the hardware. These include T1 and T2 times, as well as error rates of the readout and single- (SQ) and two-qubit (TQ) gates. More than one two-qubit gate can be implemented and their error rates are listed accordingly.

Qubit	T1 (μs)	T2 (μs)	Readout error rate	SQ gate error rate	TQ gate error rate
0	242.56	282.67	0.0109	0.000 15	[0.0053, 0.0078]
1	339.82	195.64	0.0092	0.000 14	[0.0053, 0.0058]
2	61.36	74.59	0.0071	0.000 54	[0.0162, 0.0058, 0.0061]
3	349.3	274.39	0.0132	0.000 14	[0.0058, 0.0061]
4	472.45	124.04	0.0159	0.000 18	[0.0058, 0.0071, 0.0107]
5	120.51	130.06	0.0081	0.000 36	[0.0102, 0.0107]
6	313.28	207.58	0.009	0.000 14	[0.007, 0.0102, 0.0077]
7	185.15	183.28	0.0053	0.000 17	[0.006, 0.0077]
8	437.75	359.17	0.0071	0.000 19	[0.006, 0.0085, 0.0074]
9	90.06	34.51	0.01 167	0.000 35	[0.0078, 0.008]
10	208.48	60.85	0.0119	0.000 19	[0.0071, 0.0082]
11	268.68	125.12	0.0511	0.000 52	[0.0085, 0.0077]
12	149.35	267.64	0.028 33	0.000 21	[0.008, 0.008]
13	218.14	30.84	0.0094	0.000 43	[0.0101, 0.008]
14	419.31	254.69	0.0073	0.000 13	[0.0101, 0.0057, 0.0088]
15	381.73	174.45	0.0175	0.000 15	[0.0057, 0.0079]
16	192.67	56.09	0.0278	0.000 31	[0.0082, 0.0079, 0.008]
17	386.96	13.18	0.0568	0.010 38	[0.008, 0.0051]
18	292.84	95.83	0.0048	0.000 27	[0.0041, 0.0084, 0.0051]
19	242.84	300.14	0.0222	0.000 21	[0.0041, 0.0052]
20	178.56	95.54	0.0142	0.000 17	[0.0077, 0.0096, 0.0052]

(3) *ALAP scheduling.* We utilize the "as late as possible" (ALAP) scheduling method to optimize the timing of physical qubit gates at the hardware level, ensuring efficient execution of quantum circuits [67].

APPENDIX C: SAMPLING PROCESS

A quantum memory experiment is designed to evaluate the transversal CNOT gate, incorporating either X or Z syndrome extraction circuits based on the logical qubit basis. The data were sampled by collecting outcomes from the syndrome and flag qubits for each round of the syndrome extraction circuit, including the final states of the data qubits. We conducted sampling 10^5 times for all cases, encompassing different distances, numbers of syndrome extraction circuits, and initial states.

1. X stabilizers

When the state of the logical qubit is in the Z basis, we employ Z stabilizers, as detailed in the main text. However, we can also consider the state of the logical qubit in the X basis by utilizing the phase-flip code with flag qubits to correct Z errors. The stabilizers for the X basis are defined using the X Pauli operators of neighboring data qubits, represented as X_iX_{i+1}, where i denotes the data qubit number. For example, when a logical qubit consists of three data qubits, it can be protected from phase-flip errors and defined in the logical states ($|+\rangle_L = |+++\rangle$, $|-\rangle_L = |---\rangle$). The logical Pauli operators for the logical qubit are selected similarly to when using Z stabilizers. Single- and two-qubit gates are organized in the same sequence as for Z stabilizers, with the exception that the two-qubit gates between data qubits and their neighboring flag qubits are physical CNOT gates instead of CZ gates. Hadamard gates are utilized for initializing data qubits or measuring their final state in the Z basis.

2. Quantum memory experiment

Figure 6 depicts the quantum circuit for sampling the data with the $|--\rangle_L$ structure with one round ($\mathbf{R} = \mathbf{1}$) of the syndrome extraction circuit considering X stabilizers, before and after the logical CNOT gate. The circuit can be divided into five steps as follows.

(1) *Initialization.* The logical control qubit, which comprises the three data qubits, and the other is the logical target qubit, consisting of the rest of the data qubits. Every data qubit is initialized as $|-\rangle$ by applying X and Hadamard gates. Hence, two logical qubits are initialized as $|--\rangle_L$.

(2) *Syndrome extraction rounds.* Data qubits undergo a syndrome extraction circuit considering X stabilizers. At the last round of the syndrome extraction circuit before the transversal CNOT gate, we initialize auxiliary qubits while measuring flag and syndrome qubits.

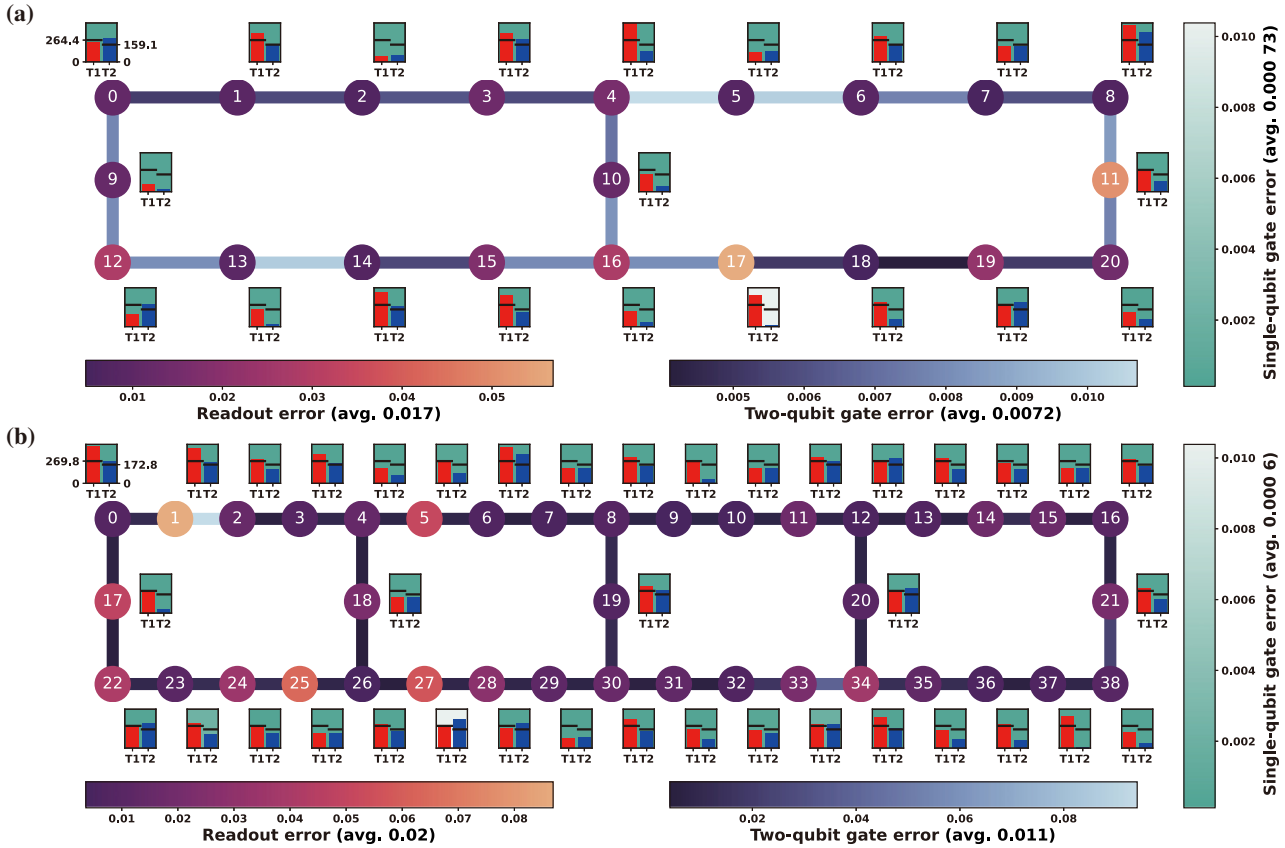

FIG. 8. Hardware specifications of physical qubits with colormaps and bar charts (`ibm_sherbrooke`). The graph with subplots shows the physical qubits' specification of the structure in (a) $d = 3$ and (b) $d = 5$, with the $|00\rangle_L$ state. Each node and edge correspond to the physical qubit and connectivity of a two-qubit gate (ECR). The error rates of the readout and two-qubit gate are displayed with colors. The subplot next to every node shows T1 and T2 times corresponding to their physical qubit. The average T1 and T2 times are plotted as the black lines on the left and right sides. Their values can be seen on the subplot of the 0th physical qubit. The background of a subplot is colored corresponding to the error rate of the single-qubit gate.

(3) *Transversal* CNOT *gate.* The transversal CNOT gate is applied to two logical qubits. Physical CNOT gates are employed between a data qubit from the logical control qubit and its closest auxiliary qubit in the structure. The physical CNOT gates are applied to grouped physical qubits to implement the transversal CNOT gate, as described in the main text.

(4) *Syndrome extraction rounds.* Another round of the syndrome extraction circuit is executed to measure syndrome and flag qubits for tracking errors in the system.

(5) *Measurement.* Finally, we measure all data qubits in the Z basis using Hadamard gates before the measurement.

3. Syndrome

Execution of noisy syndrome extraction circuits results in a binary vector determined by their outcomes. In the absence of errors, both syndrome and flag measurements are deterministic, and hence the product of measurement outcomes is also deterministic [64]. A syndrome is designed to be a zero vector when no errors are present in the system. The observation of a 1 bit in the syndrome indicates the presence of an error, and these 1 bits enable a decoder to identify the most likely correction operator based on their locations. Depending on the stabilizer group, bit- or phase-flip errors involving flag qubits result in a 1 bit in the syndrome. The repetition code with Z stabilizers can detect X errors, resulting in 1 bits in the syndrome, but it cannot detect Z errors. Conversely, the code with X stabilizers can track Z errors. We employ the same parity check process described in Ref. [16] for two logical qubits in the structure, producing the corresponding syndrome to rectify errors.

APPENDIX D: DECODING PROCESS

1. Circuit-level noise model

Each physical qubit gate is susceptible to errors. The noisy gate can be decomposed into its ideal gate and error channels. We employ a uniform depolarizing error channel simplified with Pauli operators. When the error rate p is derived from the calibration table for a basis

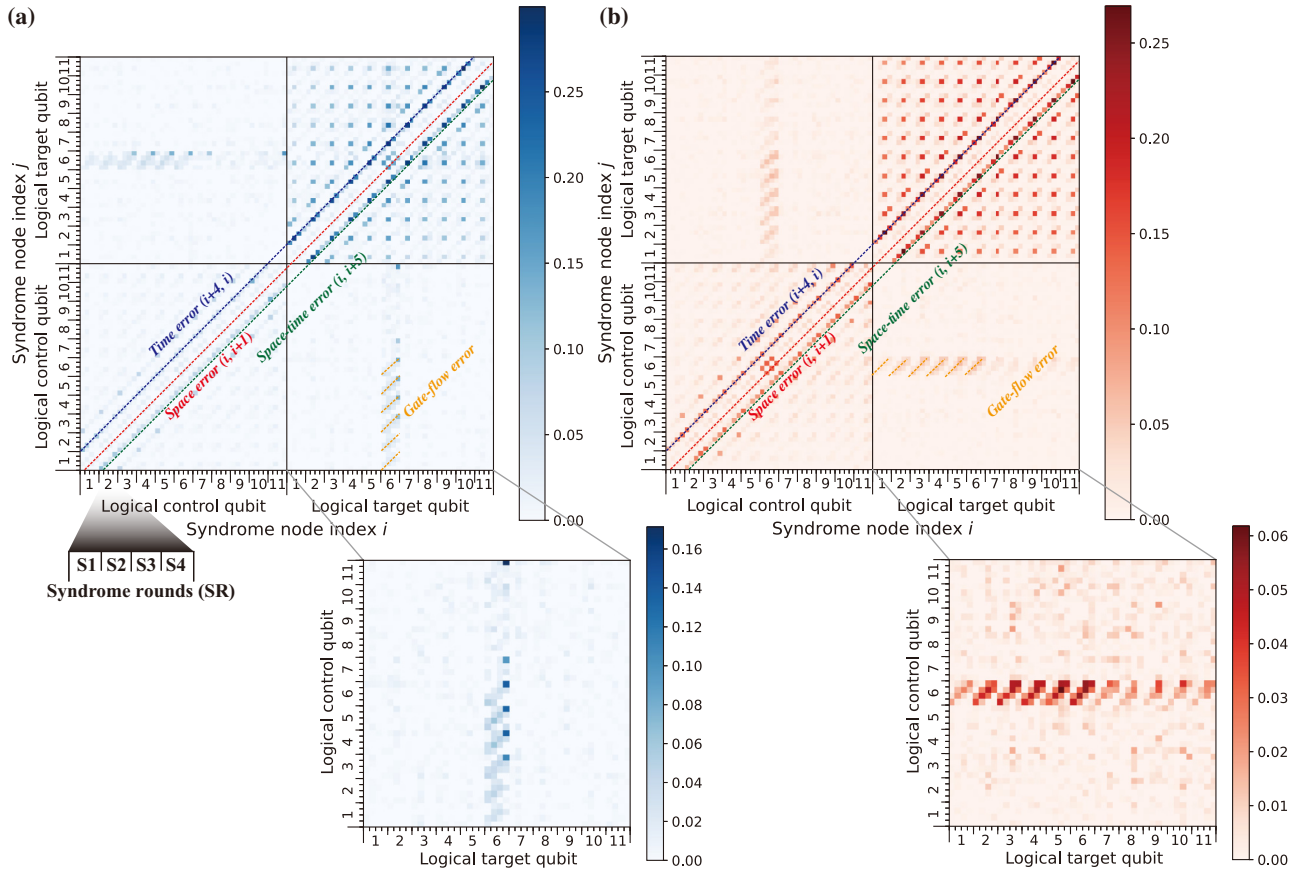

FIG. 9. Correlation matrices for the syndrome nodes (`ibm_sherbrooke`). The two logical states, the logical control and target qubits, are initialized in the $Z(X)$ basis, and each one has seven data qubits. The data are obtained by sampling the quantum memory experiment 10^5 times for both (a) $|00\rangle_L$ and (b) $|{+}{+}\rangle_L$, with $\mathbf{2R = 10}$ rounds of the syndrome extraction circuit. The axes represent the 11 blocks corresponding to syndrome rounds for a logical qubit, with each block displaying six syndrome nodes arranged according to their locations within the code. A pixel represents the error probability that generates a "1" syndrome bit on the corresponding ith and j th syndrome nodes. Parameter **SR** corresponds to the round of the detection event, and the transversal CNOT gate has been executed at $\mathbf{SR = 6}$. The probabilities of space, time, space-time, and gate-flow errors are visually represented by red, blue, green, and orange dotted lines, respectively.

gate, we implement the circuit-level noise model using the depolarizing error model as follows.

(1) The uniform Pauli error channel, characterized by p_X, p_Y, and p_Z, satisfies the condition $p_X = p_Y = p_Z = p/3$ and occurs following an ideal single-qubit gate.

(2) The uniform Pauli error channel is applied to two qubits, excluding the identity operator $(I \otimes I)$, following the ideal two-qubit (CNOT) gates. The probability of each Pauli error from $\{X, Y, Z, I\}^{\otimes 2}/\{I \otimes I\}$ is $p/15$. Additionally, two-qubit depolarization error channels are individually employed before and after the two-qubit gates, respecting their corresponding single-qubit gate error rates in the calibration table [16]. When a device uses an ECR pulse as a basic gate, we use the same error channel for a CZ gate as a physical CNOT gate. However, if the device can operate a CZ pulse at the hardware level, we only consider the two-qubit depolarization error channel.

(3) An X error channel with probability p is introduced before the measurement and after the reset gate.

(4) When a physical qubit is inactive and undergoes free evolution, it encounters the uniform Pauli error channel for a single qubit with the probability of the error being p, the idling error rate from the calibration table.

We adopt the above circuit-level noise model to calculate weights for each edge and construct a syndrome graph using the Stim code [64]; the syndrome graph is used for decoding errors using the Pymatching algorithm [65,66].

2. Syndrome graph

Figure 7 displays the syndrome graphs for the structure with five data qubits and 11 syndrome rounds, where the horizontal and vertical axes represent space and time. The graph can be divided into two regions: the logical control and target qubits. The region for the logical control qubit is

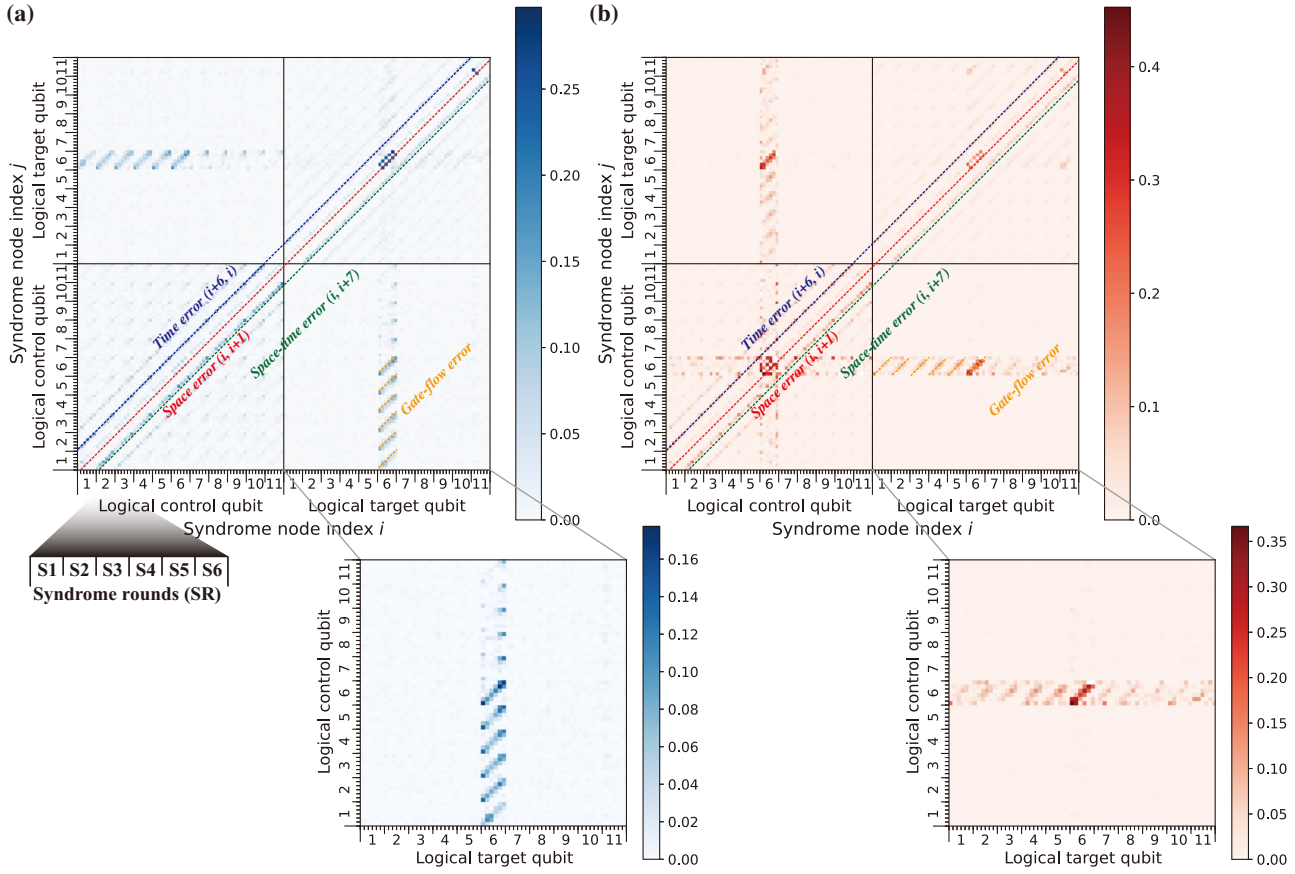

FIG. 10. Correlation matrices for the syndrome nodes (ibm_brisbane). The two logical states, the logical control and target qubits, are initialized in the $Z(X)$ basis, and each one has seven data qubits. The data are obtained by sampling the quantum memory experiment 10^5 times for both (a) $|11\rangle_L$ and (b) $|--\rangle_L$, with $\mathbf{2R = 10}$ rounds of the syndrome extraction circuit. The axes represent the 11 blocks corresponding to syndrome rounds for a logical qubit, with each block displaying six syndrome nodes arranged according to their locations within the code. A pixel represents the error probability that generates a "1" syndrome bit on the corresponding ith and j th syndrome nodes. Parameter **SR** corresponds to the round of the detection event, and the transversal CNOT gate has been executed at $\mathbf{SR = 6}$. The probabilities of space, time, space-time, and gate-flow errors are visually represented by red, blue, green, and orange dotted lines, respectively.

enclosed within the open blue box, while that of the other logical qubit is enclosed within the open purple box. Each logical qubit has four syndrome nodes (**S1**, **S2**, **S3**, and **S4**) per syndrome round (**SR**). The syndrome graphs are the same, with the only difference being the type of edge. Edges in the graph correspond to four types (space, time, space-time, and gate-flow errors) of errors that produce a pair of corresponding syndrome nodes to have detection events. We depict the categorized errors in Figs. 7(a), 7(b), 7(c), and 7(d) for the logical qubits that are initialized as $|{+}{+}\rangle_L$. Notably, the correlated edges can differ in two ways depending on what error type we consider. Hence, the gate-flow errors for the Z basis ($|00\rangle_L$) are also depicted in a different way to those of the X basis in Fig. 7(e). This is because, when the logical qubits are in the X basis, a Z error can propagate from the logical target qubit to the logical control qubit; the opposite holds for the Z basis with an X error.

A virtual syndrome node is introduced to consider an error that produces an odd number of detection events, especially in the case of a space error on the boundary data qubit of the logical control qubit in Fig. 7(b). However, when a space error occurs before the transversal CNOT gate on the boundary data qubit of the logical target qubit, the virtual node is unnecessary. This is because the error will propagate to the other logical qubit and produce another detection event, resulting in an even number of detection events being considered as a gate-flow error. The locally corresponding syndrome nodes are connected by a gate-flow error. Gate-flow errors are only connected with syndrome nodes on the logical (d) target or (e) control qubit at the time $\mathbf{SR = 6}$ when the logical gate has been executed depending on the basis of the logical qubits. Since errors propagate only when the transversal CNOT gate is implemented, no gate-flow errors are considered after the gate. This is attributed to the Pauli operators'

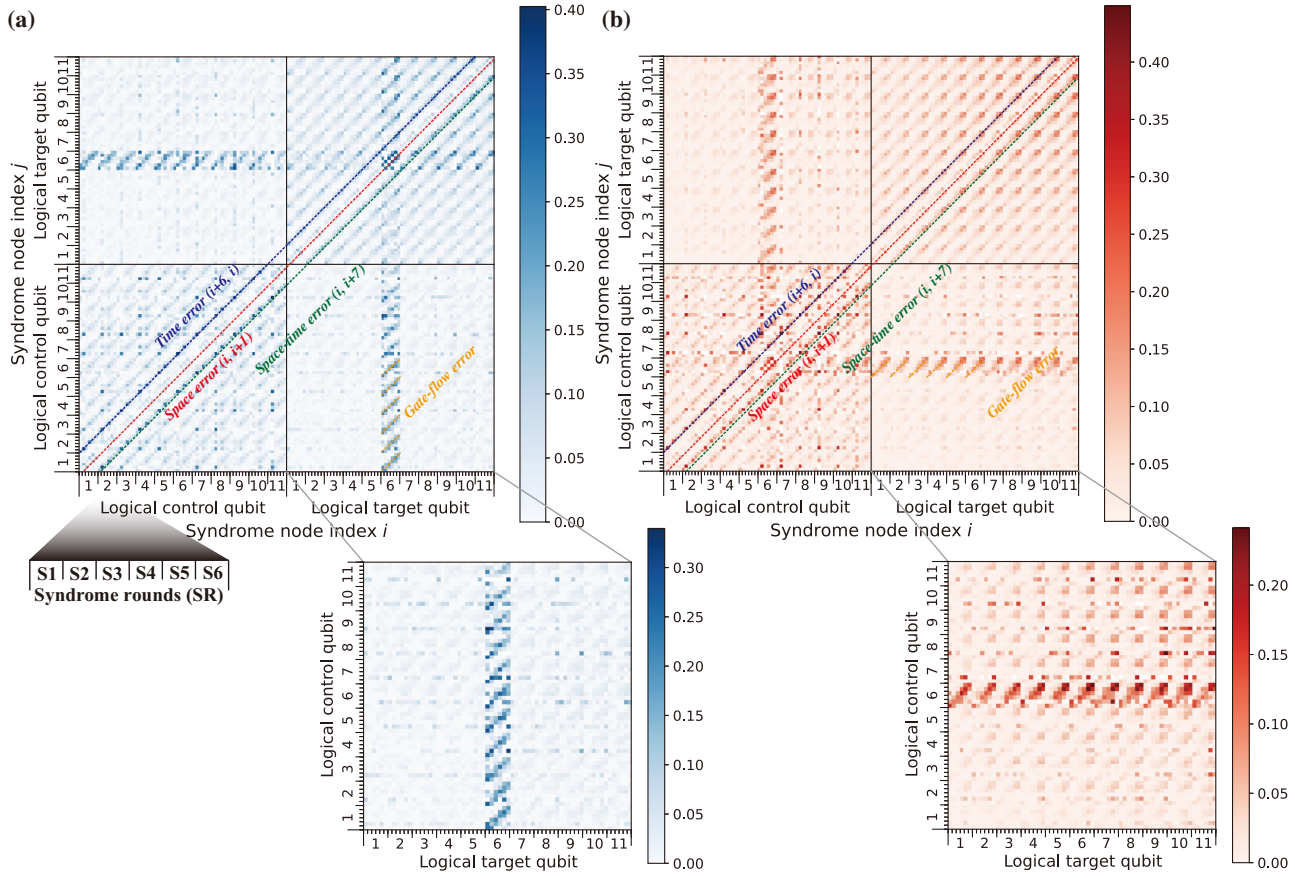

FIG. 11. Correlation matrices for the syndrome nodes (ibm_torino). The two logical states, the logical control and target qubits, are initialized in the $Z(X)$ basis, and each one has seven data qubits. The data are obtained by sampling the quantum memory experiment 10^5 times for both (a) $|11\rangle_L$ and (b) $|{-}{-}\rangle_L$, with $\mathbf{2R = 10}$ rounds of the syndrome extraction circuit. The axes represent the 11 blocks corresponding to syndrome rounds for a logical qubit, with each block displaying six syndrome nodes arranged according to their locations within the code. A pixel represents the error probability that generates a "1" syndrome bit on the corresponding *i*th and *j* th syndrome nodes. Parameter **SR** corresponds to the round of the detection event, and the transversal CNOT gate has been executed at $\mathbf{SR = 6}$. The probabilities of space, time, space-time, and gate-flow errors are visually represented by red, blue, green, and orange dotted lines, respectively.

propagation rule that has been demonstrated in the main text.

Each edge is assigned a weight based on the probability of the corresponding error, with a lower weight indicating a higher probability. When calculating a correction operator based on the syndrome, edges with lower weights are more likely to be chosen as correction operators to rectify errors. The figures are generated by considering the calibration table and using the Stim code considering the circuit-level noise. The weights of time errors have relatively lower values than those of other errors. In Fig. 7, these weights are calculated and displayed based on the calibration table of ibm_sherbrooke. Depending on what physical qubit we use, the higher the measurement error rates in the hardware specification, the lower the weight. The weights of gate-flow errors exhibit the largest gap between the highest and lowest values among the error types. The edges in the first syndrome round have relatively lower values compared to other syndrome rounds. This could be because not only the data qubit's idling error and errors during the syndrome extraction circuit, but also initialization errors on the data qubit affect it.

APPENDIX E: RESULTS

1. Selected physical qubits

Table III presents details such as the number of physical qubits, T1, T2, and error rates for gates, including single- and two-qubit gates in the case of ibm_sherbrooke. These details are relevant to our quantum memory experiment utilizing $|00\rangle_L$ with 21 physical qubits. The average values of T1 and T2 are 264.4 and 159.1 μs, respectively. The average error rate for measurement gates is 0.017, while for single-qubit gates, it is 0.00073, and for two-qubit gates, it is 0.0072. The same values for other devices are available in Appendix F.

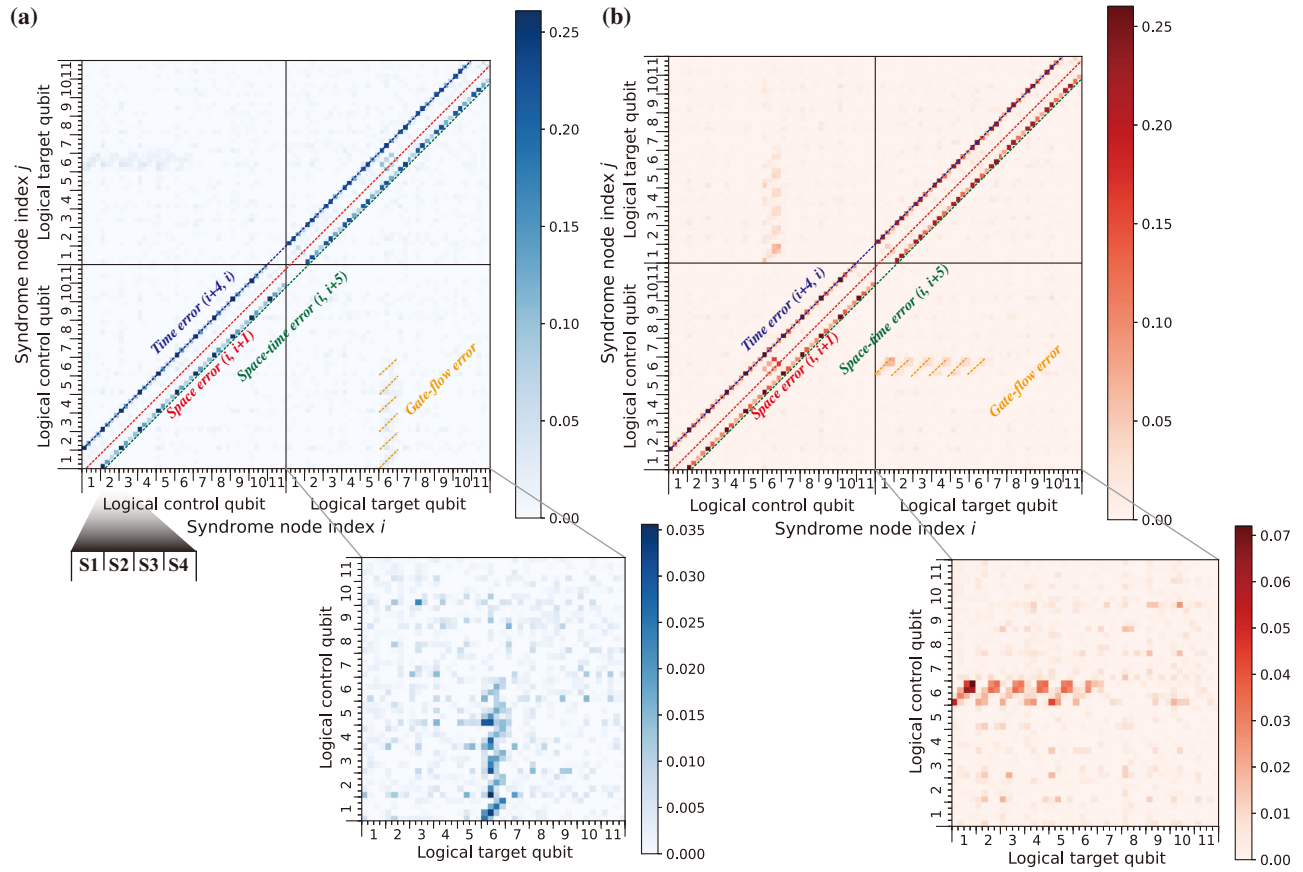

FIG. 12. Correlation matrices for the syndrome nodes obtained using Stim (`ibm_sherbrooke`). The two logical states, the logical control and target qubits, are initialized in the $Z(X)$ basis, and each one has seven data qubits. The data are obtained by sampling the quantum memory experiment 10^5 times for both (a) $|00\rangle_L$ and (b) $|{+}{+}\rangle_L$, with $\mathbf{2R = 10}$ rounds of the syndrome extraction circuit. The axes represent the 11 blocks corresponding to syndrome rounds for a logical qubit, with each block displaying four syndrome nodes arranged according to their locations within the code. A pixel represents the error probability that generates a "1" syndrome bit on the corresponding ith and jth syndrome nodes. Parameter **SR** corresponds to the round of the detection event, and the transversal CNOT gate has been executed at $\mathbf{SR = 6}$. The probabilities of space, time, space-time, and gate-flow errors are visually represented by red, blue, green, and orange dotted lines, respectively.

Figure 8 illustrates the specifications of the selected physical qubits for the $|00\rangle_L$ structure with distances 3 and 5 on `ibm_sherbrooke`. The same figures for other devices can be seen in Appendix F. Each node and edge in the figure represents a physical qubit and its connectivity with neighboring physical qubits, indicating possible direct two-qubit gate operations. The color of each node and edge corresponds to its readout error rate and two-qubit gate error rate, with brighter regions indicating higher error occurrence. Additionally, the average values of T1 and T2 among the selected physical qubits are depicted on the subplot corresponding to the 0th physical qubit. These values are represented as black lines on the left and right axes, with red and blue bars indicating T1 and T2 of the physical qubit, respectively. The background of the subplots displays the error rate of the single-qubit gate. Subplots for other physical qubits also show their T1, T2, and error rate of the single-qubit gate. It is observed that, when more physical qubits are used, error rates for readout tend to increase. Furthermore, while the average error rates of two-qubit gates on `ibm_torino` were 0.0066 and 0.0095 for distances 3 and 5, respectively, those on `ibm_sherbrooke` were higher at 0.0072 and 0.011. However, the average readout error rates on `ibm_sherbrooke` were 0.017 and 0.02, which are lower than those on `ibm_torino` at 0.02 and 0.024.

2. Correlation matrices

There are two methods to label the number of syndrome nodes: (1) labeling the nodes based on space first, and (2) labeling the nodes based on time first instead of space. We consider labeling the syndrome nodes for the logical control qubit first, followed by the logical target qubit in Fig. 4 of the main text. Figure 9 shows the same data as that without truncation. Figures 10 and 11 present correlation matrices derived from sampled data obtained from different hardware platforms, `ibm_brisbane` and

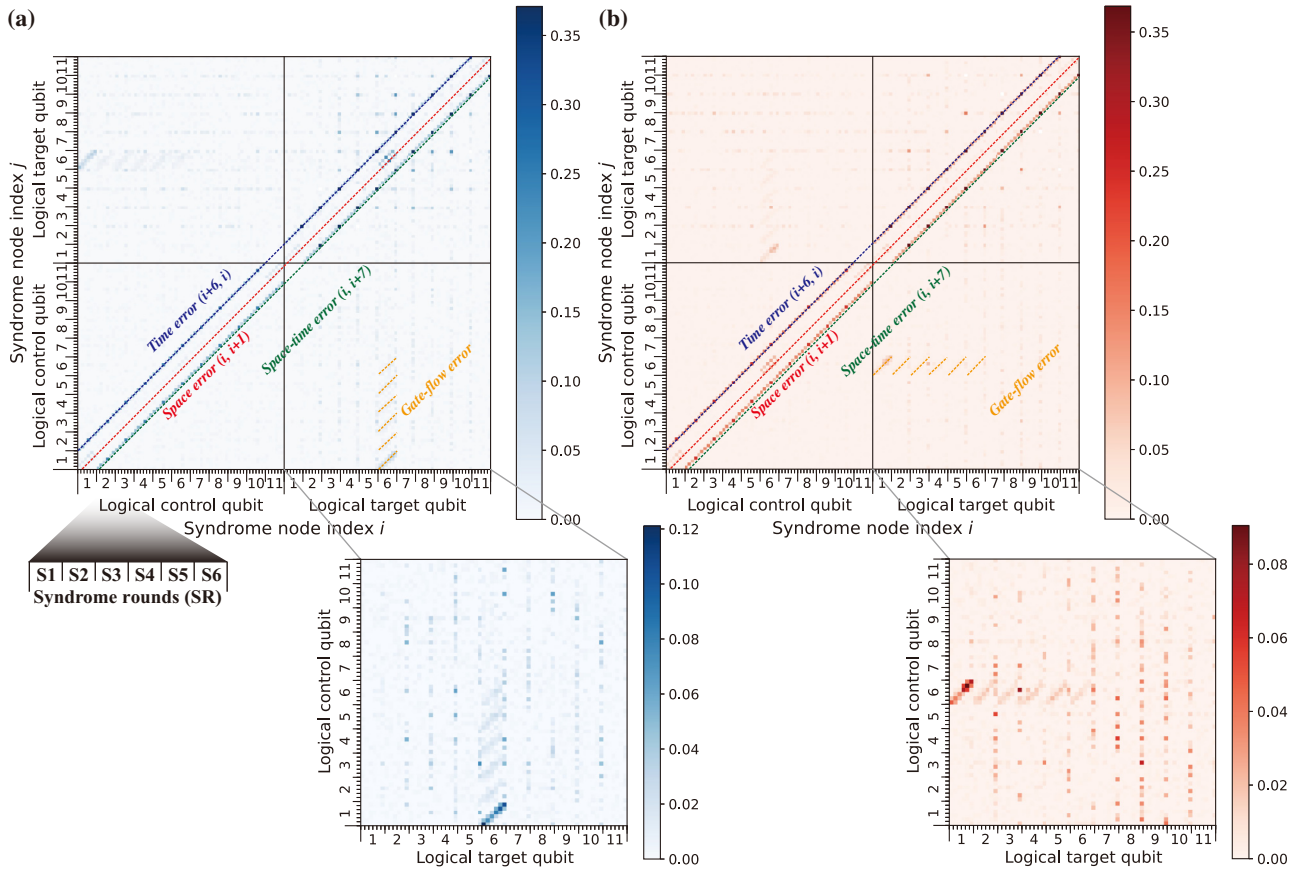

FIG. 13. Correlation matrices for the syndrome nodes obtained using Stim (`ibm_torino`). The two logical states, the logical control and target qubits, are initialized in the $Z(X)$ basis, and each one has seven data qubits. The data are obtained by sampling the quantum memory experiment 10^5 times for both (a) $|11\rangle_L$ and (b) $|{-}{-}\rangle_L$, with $\mathbf{2R} = \mathbf{10}$ rounds of the syndrome extraction circuit. The axes represent the 11 blocks corresponding to syndrome rounds for a logical qubit, with each block displaying six syndrome nodes arranged according to their locations within the code. A pixel represents the error probability that generates a "1" syndrome bit on the corresponding *i*th and *j* th syndrome nodes. Parameter **SR** corresponds to the round of the detection event, and the transversal CNOT gate has been executed at $\mathbf{SR} = \mathbf{6}$. The probabilities of space, time, space-time, and gate-flow errors are visually represented by red, blue, green, and orange dotted lines, respectively.

`ibm_torino`, respectively. Each matrix corresponds to demonstrations performed with initial states (a) $|11\rangle_L$ and (b) $|{-}{-}\rangle_L$. In this setup, a logical qubit comprises seven data qubits, necessitating a total of 57 physical qubits for the structure with ten rounds of a syndrome extraction circuit (**2R**). They show high error probabilities, particularly for time errors. Moreover, the probabilities of space errors on either the logical control or target qubits are significantly influenced by the basis of the logical qubits. The X (Z) errors on the logical target (control) qubit before the logical CNOT gate propagate to the other logical qubits, resulting in gate-flow errors, which can be seen in the figures. Strong correlations are observed at specific times, notably at $\mathbf{SR} = \mathbf{6}$, between two logical qubits from the perspective of the affected logical qubit due to gate-flow errors. Overall, these correlation matrices demonstrate similar patterns of correlations as observed in the results from both actual devices and the classical simulator. It is worth pointing out that in the results from `ibm_brisbane`, the correlation matrix with the X basis shows high error probabilities in the correlation region at $\mathbf{SR} = \mathbf{6}$ for both logical qubits. It could be related to the biased errors during the execution of the logical CNOT gate. Additionally, we obtain the correlation matrices for `ibm_sherbrooke` and `ibm_torino` using the classical simulator (Stim) with the same error rate as the hardware, as displayed in Figs. 12 and 13. We observed that the correlations between two logical qubits from the simulated results align with the results from the quantum hardware, but show relatively low error probabilities.

Figures 14 and 15 illustrate correlation matrices obtained by labeling the syndrome nodes based on time before space in the axes of the correlation matrices from `ibm_sherbrooke` and `ibm_torino`. In Fig. 14, each

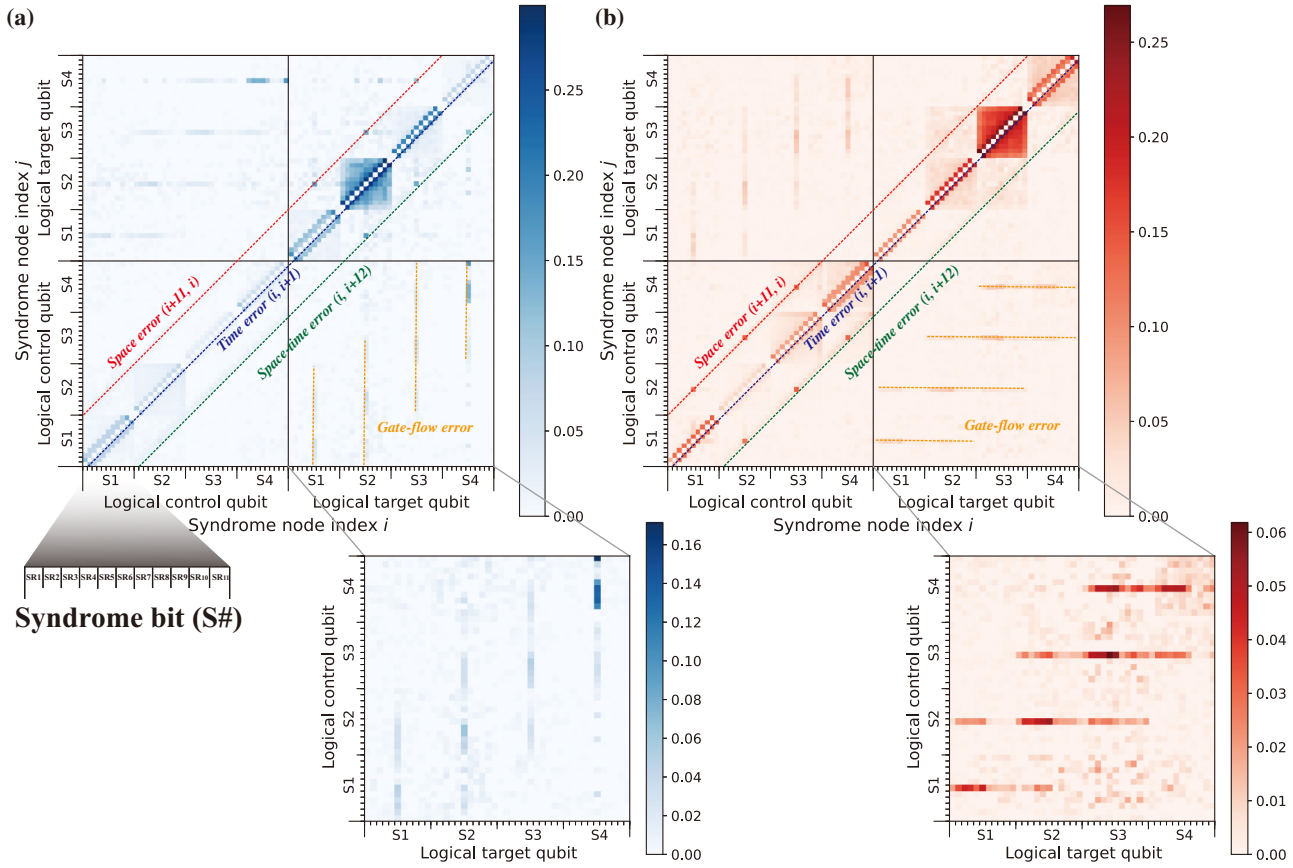

FIG. 14. Correlation matrices for the syndrome nodes (`ibm_sherbrooke`). The syndrome nodes are labeled based on time before space. The two logical states, the logical control and target qubits, are initialized in the $Z(X)$ basis, and each one has five data qubits. The data are obtained by sampling the quantum memory experiment 10^5 times for both (a) $|00\rangle_L$ and (b) $|{+}{+}\rangle_L$, with $\mathbf{2R} = \mathbf{10}$ rounds of the syndrome extraction circuit. The axes represent the four blocks corresponding to syndrome nodes for a logical qubit, with each block displaying 11 syndrome rounds arranged according to their locations within the code. A pixel represents the error probability that generates a "1" syndrome bit on the corresponding ith and j th syndrome nodes. The syndrome nodes are labeled first based on time, and then on space. Parameter **SR** corresponds to the round of the detection event, and the transversal CNOT gate has been executed at $\mathbf{SR} = \mathbf{6}$. The probabilities of space, time, space-time, and gate-flow errors are visually represented by red, blue, green, and orange dotted lines, respectively.

logical qubit consists of five data qubits, with four spatial syndrome bits per logical qubit. These syndrome bits, labeled **S1**, **S2**, **S3**, and **S4**, correspond to their positions in the code. It is important to note that each **S1** from the two logical qubits corresponds locally within the perspective of each logical qubit. The same correspondence applies to the other syndrome bits (**S2**, **S3**, **S4**). The syndrome extraction circuit is performed ten times, with the corresponding syndrome having 11 syndrome rounds (**SR**). Hence, the axes represent the four blocks corresponding to syndrome nodes for a logical qubit, with each block displaying 11 syndrome rounds arranged according to their locations within the code. The transversal CNOT gate is implemented in the middle of the syndrome extraction rounds ($\mathbf{SR} = \mathbf{6}$), allowing for the observation of the correlation between the two logical qubits, as depicted in the figures. In the case of `ibm_torino` shown in Fig. 15, the only difference is the number of data qubits per logical qubit, which is seven.

In Fig. 14, strong correlations are observed between the logical control and target qubits, particularly when the logical CNOT gate is introduced. Additionally, each syndrome bit from one logical qubit, including nearest-neighbor syndrome bits, shows correlations with the locally corresponding syndrome bit from the other logical qubit. These correlations signify error propagation affecting the logical state and are represented differently depending on the type of errors, reflecting the propagation of Pauli operators. On the other hand, a noticeable high in the error probability is observed for specific syndrome bits, such as (a) **S2** or (b) **S3**, at the logical target qubit. Although these errors can be decomposed into consecutive time errors, their actual nature may vary, including leakage or correlated errors. These differences contribute to higher overall logical error

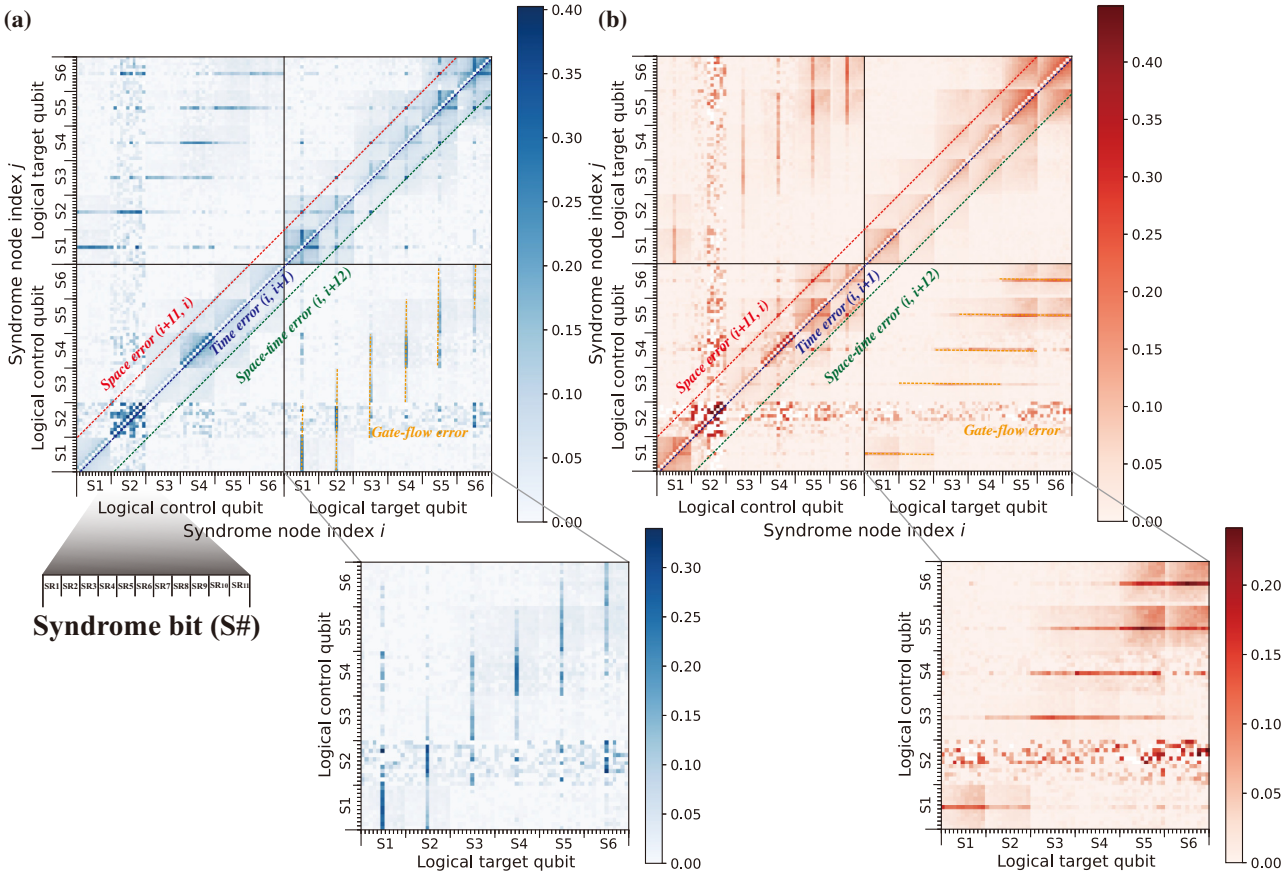

FIG. 15. Correlation matrices for the syndrome nodes (`ibm_torino`). The syndrome nodes are labeled based on time before space. The two logical states, the logical control and target qubits, are initialized in the $Z(X)$ basis, and each one has seven data qubits. The data are obtained by sampling the quantum memory experiment 10^5 times for both (a) $|11\rangle_L$ and (b) $|{-}{-}\rangle_L$, with $\mathbf{2R = 10}$ rounds of the syndrome extraction circuit. The axes represent the six blocks corresponding to syndrome nodes for a logical qubit, with each block displaying 11 syndrome rounds arranged according to their locations within the code. A pixel represents the error probability that generates a "1" syndrome bit on the corresponding ith and j th syndrome nodes. The syndrome nodes are labeled first based on time, and then on space. Parameter **SR** corresponds to the round of the detection event, and the transversal CNOT gate has been executed at $\mathbf{SR = 6}$. The probabilities of space, time, space-time, and gate-flow errors are visually represented by red, blue, green, and orange dotted lines, respectively.

rates. These properties are also observed in Fig. 15 except for the regions where syndrome bit **S2** is related.

Intriguingly, there is another type of error that causes low performance of the transversal CNOT gate. Particularly in the region related to syndrome bit **S2** within the results obtained from `ibm_torino`, most of the pixels behave differently, specifically showing 0 values. This indicates that they have negative values according to Eq. (1). These errors in syndrome bit **S2** were also observed across various distance structures and different rounds within the region where using the same physical qubits. These observations suggest that errors occurring during the execution of quantum circuits on the device primarily manifest locally. This localized error behavior, particularly around syndrome bit **S2**, highlights challenges inherent to quantum processors that warrant further investigation. We hope that our demonstration will accelerate studies addressing these challenges, including understanding and mitigating inherent errors in quantum processors, which should be a key focus of ongoing research for a fault-tolerant quantum computer.

3. Detection probability

The probability of a detection event can be calculated by the ratio of how many times a syndrome bit is observed as a detection event in each syndrome round (**SR**) and the number of trials. Figure 16 illustrates the probabilities of detection events obtained from `ibm_sherbrooke` as a function of syndrome rounds and syndrome node labels for the $d = 5$ structure on the (a) Z ($|00\rangle_L$) or (b) X ($|{+}{+}\rangle_L$) basis. The blue and red lines represent the detection event probabilities for the logical control and target qubits, respectively. The logical CNOT gate is

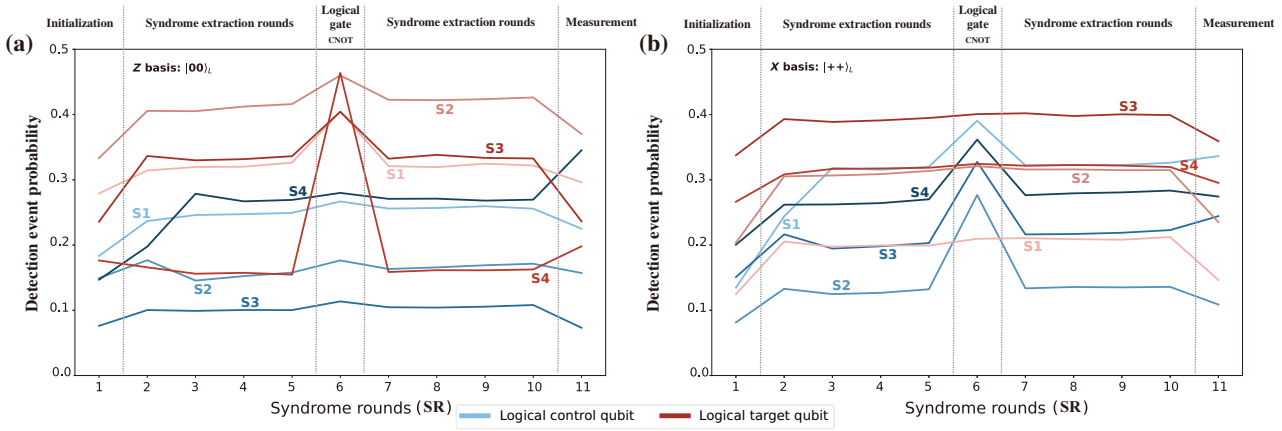

FIG. 16. Probabilities of a detection event on each syndrome node over rounds (`ibm_sherbrooke`). The data are obtained from the demonstration with initial states (a) $|00\rangle_L$ and (b) $|{+}{+}\rangle_L$ with $\mathbf{2R = 10}$ syndrome extraction rounds. The probability of the syndrome node being the detection event over time is plotted as a function of syndrome rounds. The number of syndrome rounds is 11 and steps of the quantum memory experiment can be seen based on the syndrome rounds. The transversal CNOT gate is executed at $\mathbf{SR = 6}$. The red (blue) lines correspond to the probabilities of the syndrome bits of the logical control (target) qubit. The syndrome bits are labeled **S1**, **S2**, **S3**, and **S4** according to their locations within the code. The number of samples is 10^5 to calculate the probabilities for each case.

applied after five rounds of syndrome extraction circuits, which is $\mathbf{SR = 6}$. In the case of X errors, they propagate from the logical control qubit to the logical target qubit, whereas for Z errors, the propagation occurs in the opposite direction. Notably, the detection event probabilities for all syndrome bits on the logical target (control) qubit increase significantly when the logical qubits are in the Z (X) basis at $\mathbf{SR = 6}$. Furthermore, the probabilities for syndrome bits on the other logical qubit also temporally increase when the logical CNOT gate is implemented. This increase is attributed to the use of auxiliary qubits in implementing the transversal CNOT gate, which introduces the possibility of errors associated with these auxiliary qubits during the application of the logical CNOT gate. Consequently, the detection event probabilities for syndrome nodes on both logical qubits with Z (X) stabilizers increase, reflecting the impact of gate-flow errors. It is also similarly observed from `ibm_torino` except for the **S2** syndrome bit with inherent errors, as shown in Fig. 17.

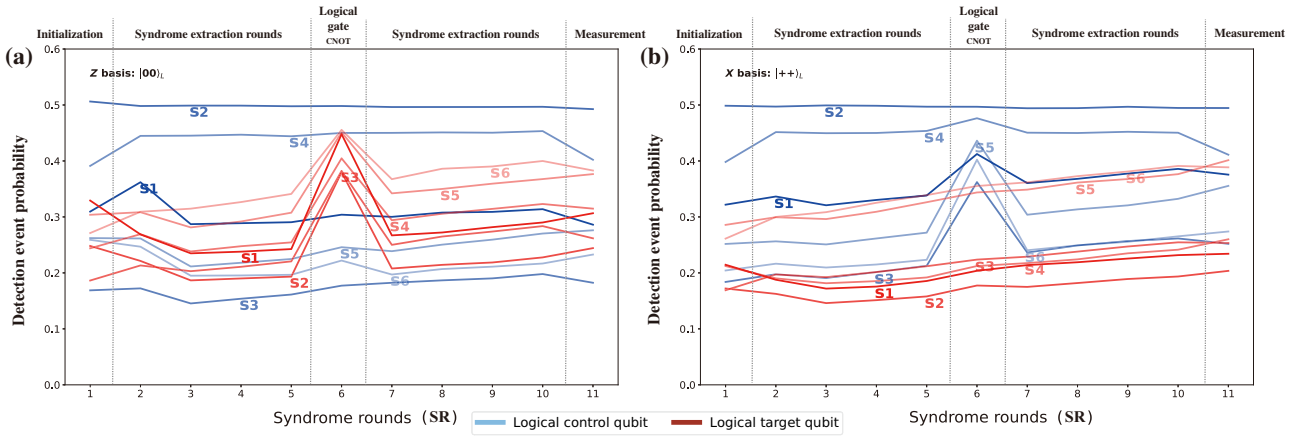

FIG. 17. Probabilities of a detection event on each syndrome node over rounds (`ibm_torino`). The data are obtained from the demonstration with initial states (a) $|00\rangle_L$ and (b) $|{+}{+}\rangle_L$ with $\mathbf{2R = 10}$ syndrome extraction rounds. The probability of the syndrome node being the detection event over time is plotted as a function of syndrome rounds. The number of syndrome rounds is 11 and steps of the quantum memory experiment can be seen based on the syndrome rounds. The transversal CNOT gate is executed at $\mathbf{SR = 6}$. The red (blue) lines correspond to the probabilities of the syndrome bits of the logical control (target) qubit. The syndrome bits are labeled **S1**, **S2**, **S3**, **S4**, **S5**, and **S6** according to their locations within the code. The number of samples is 10^5 to calculate the probabilities for each case.

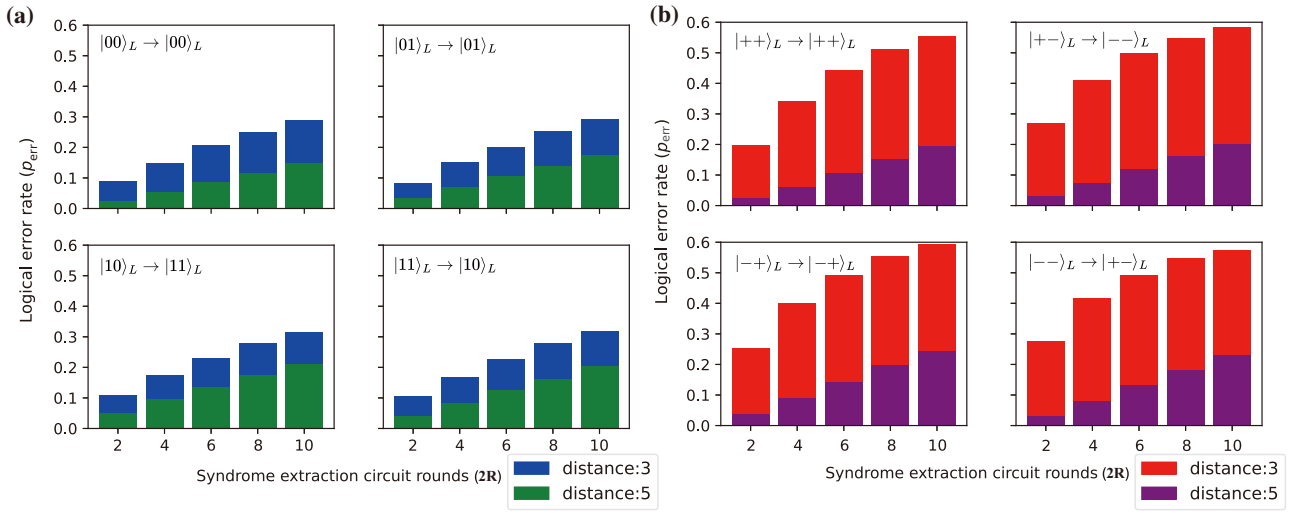

FIG. 18. Bar charts showing the logical error rates (`ibm_sherbrooke`). The logical qubits initialized in the (a) Z or (b) X basis, calculated based on data sampled 10^5 times for each case using the quantum memory experiment. The logical error rates are displayed as a function of distances and the number of rounds of the syndrome extraction circuit. The logical errors for distances 3 and 5 are plotted noncumulatively across the syndrome rounds. The state of the logical qubits can be prepared as one of four states for each basis. Subfigures illustrate the case for a certain initial state.

4. Logical error rates

Demonstrations are performed with rounds of the corresponding syndrome extraction circuit, initializing the logical qubits in four different states for each basis. Here, $|\psi_C\psi_T\rangle_L = |\psi_C\rangle_L \otimes |\psi_T\rangle_L$ denotes the product state of logical control ($|\psi_C\rangle_L$) and target ($|\psi_T\rangle_L$) qubits. For example, when using Z stabilizers, the computational states of two logical qubits can be prepared as $\{|00\rangle_L, |01\rangle_L, |10\rangle_L, |11\rangle_L\}$. Applying the logical CNOT gate to these states transforms them into $\{|00\rangle_L, |01\rangle_L, |11\rangle_L, |10\rangle_L\}$, respectively. Similarly, for the X basis, the initial states are prepared as one of $\{|++\rangle_L, |+-\rangle_L, |-+\rangle_L, |--\rangle_L\}$, which change to $\{|++\rangle_L, |--\rangle_L, |-+\rangle_L, |+-\rangle_L\}$, respectively.

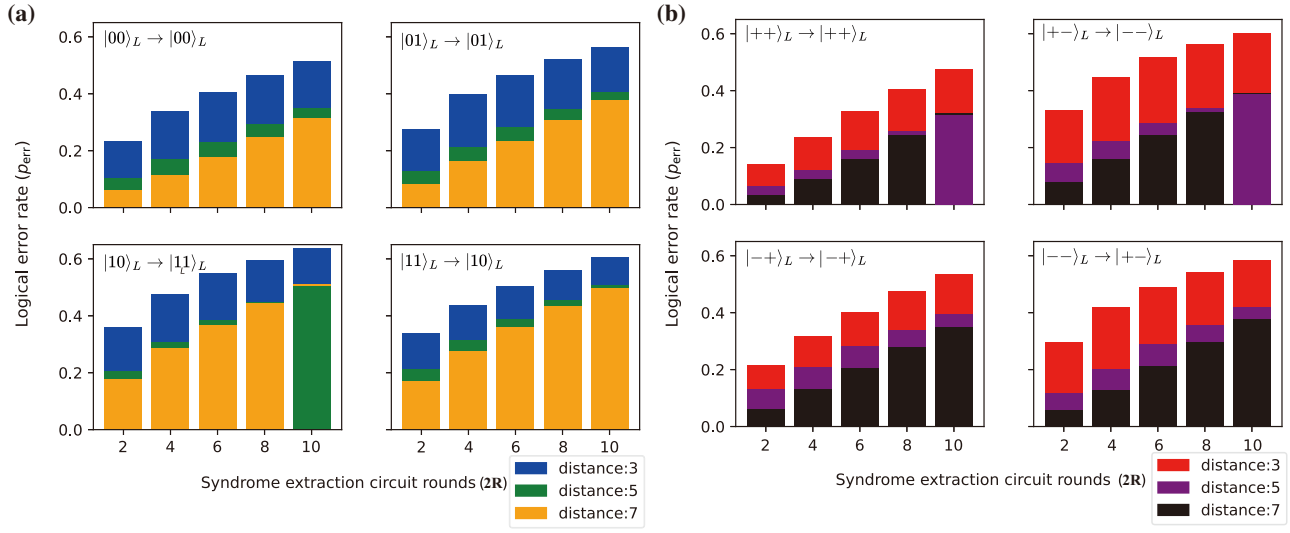

FIG. 19. Bar charts showing the logical error rates (`ibm_torino`). The logical qubits initialized in the (a) Z or (b) X basis, calculated based on data sampled 10^5 times for each case using the quantum memory experiment. The logical error rates are displayed as a function of distances and the number of rounds of the syndrome extraction circuit. The logical errors for distances 3, 5, and 7 are plotted noncumulatively across the syndrome rounds. The state of the logical qubits can be prepared as one of four states for each basis. Subfigures illustrate the case for a certain initial state.

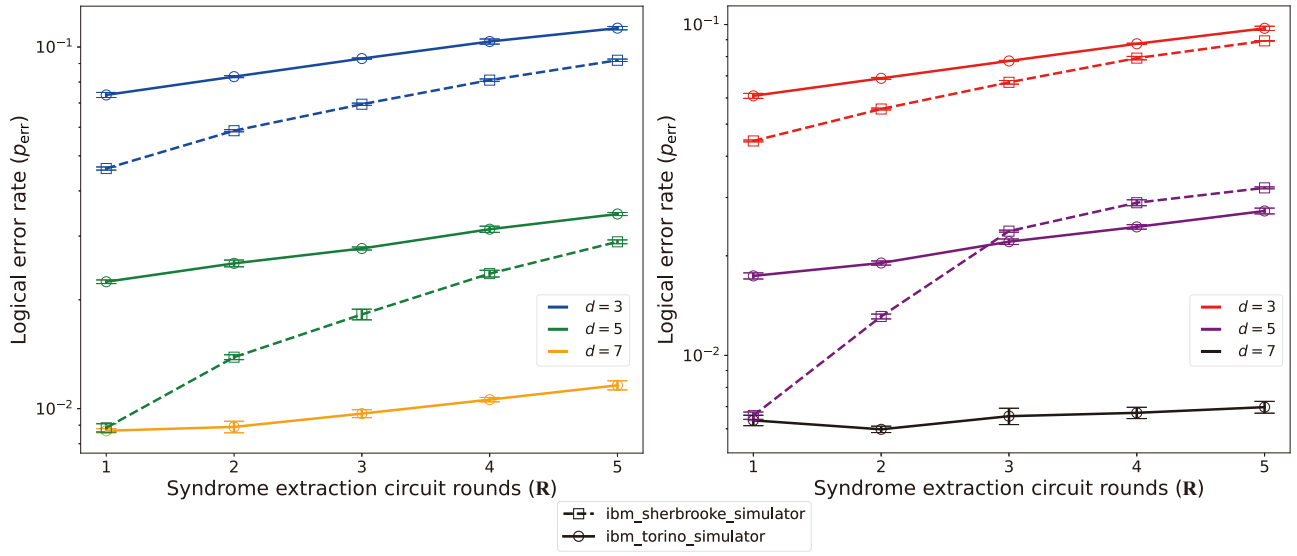

FIG. 20. Logical error rates using Stim (`ibm_torino` and `ibm_sherbrooke`). The code considers (a) X errors with Z stabilizers or (b) Z errors with X stabilizers. To assess the logical error rates, the MWPM decoder is used. The blue (red), green (purple), and orange (black) lines correspond to structure sizes of 21, 39, and 57 physical qubits correcting X (Z) errors, respectively. The lines are obtained by sampling the quantum memory experiment 10^5 times using Stim for two devices, `ibm_torino` and `ibm_sherbrooke`. The average logical error rates for four basis states are depicted with their standard deviations as a function of the number of syndrome extraction circuit rounds, and structure size. As the number of physical qubits increases, the logical error rates decrease for both X and Z error cases.

Logical error rates (p_{err}) are calculated as infidelity ($1 - F_{\psi_C,\psi_T}$), where F_{ψ_C,ψ_T} is the fidelity when the output state is $|\psi_C\psi_T\rangle_L$ after the logical CNOT gate has been executed. After decoding errors and correcting the two logical qubit states, the fidelity of the corrected logical qubits is computed using the logical qubits' Pauli operators [35]. For the Z basis, the fidelity calculations are

$$F_{00} = \tfrac{1}{4}(1 + \langle Z_L^C\rangle + \langle Z_L^T\rangle + \langle Z_L^C Z_L^T\rangle), \tag{E1}$$

$$F_{01} = \tfrac{1}{4}(1 + \langle Z_L^C\rangle - \langle Z_L^T\rangle - \langle Z_L^C Z_L^T\rangle), \tag{E2}$$

$$F_{10} = \tfrac{1}{4}(1 - \langle Z_L^C\rangle + \langle Z_L^T\rangle - \langle Z_L^C Z_L^T\rangle), \tag{E3}$$

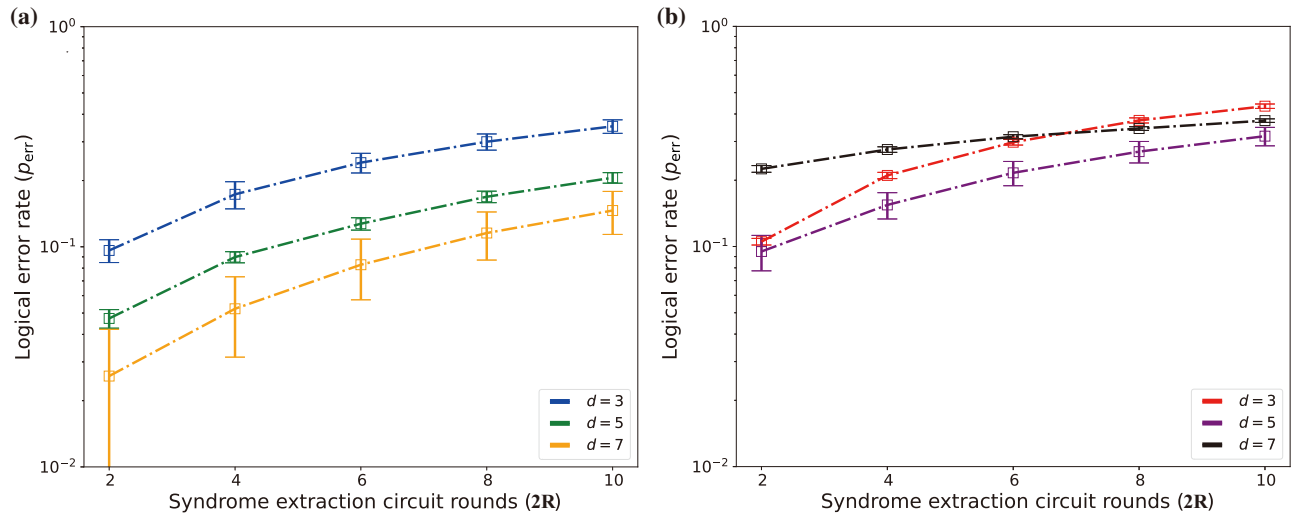

FIG. 21. Logical error rates (`ibm_brisbane`). The code considers (a) X errors with Z stabilizers or (b) Z errors with X stabilizers. To assess the logical error rates, the MWPM decoder is used. The blue (red), green (purple), and orange (black) lines correspond to structure sizes of 21, 39, and 57 physical qubits correcting X (Z) errors, respectively. The lines are obtained by sampling the quantum memory experiment 10^5 times on `brisbane`. The average logical error rates for four basis states are depicted with their standard deviations as a function of the number of syndrome extraction circuit rounds, and structure size. As the number of physical qubits increases, the logical error rates decrease for both X and Z error cases.

TABLE IV. Calibration table values for 21 physical qubits used for demonstrations (`ibm_brisbane`). The values obtained from the calibration table correspond to $|00\rangle_L$ with ten rounds. The table lists the properties of physical qubits for the $d = 3$ structure selected within the hardware. These include T1 and T2 times, as well as error rates of the readout and single- (SQ) and two-qubit (TQ) gates. More than one two-qubit gate can be implemented and their error rates are listed accordingly.

Qubit	T1 (μs)	T2 (μs)	Readout error rate	SQ gate error rate	TQ gate error rate
0	248.48	368.92	0.0133	0.000 17	[0.0137, 0.0048]
1	264.36	317.35	0.0132	0.000 14	[0.0048, 0.0041]
2	284.58	306.75	0.008	0.000 38	[0.0087, 0.0041]
3	216.28	67.9	0.0065	0.000 19	[0.0051, 0.0087]
4	430.15	78.68	0.008	0.000 15	[0.0051, 0.0051, 0.0053]
5	330.82	290.9	0.0137	0.000 28	[0.0067, 0.0053]
6	318.33	183.28	0.0136	0.000 18	[0.0086, 0.0067, 0.0051]
7	271.58	287.72	0.12	0.000 18	[0.005, 0.0051]
8	308.95	262.67	0.0134	0.000 11	[0.005, 0.0032]
9	226.93	41.98	0.0148	0.000 49	[0.0137, 0.0136]
10	493.14	381.62	0.0202	0.000 13	[0.0051, 0.0085]
11	373.5	177.92	0.0068	0.000 19	[0.0074, 0.0032]
12	227.92	60.3	0.0083	0.000 38	[0.0079, 0.0136, 0.0082]
13	363.33	81.61	0.0131	0.000 18	[0.0051, 0.0082]
14	254.07	186.99	0.0081	0.000 31	[0.0095, 0.0051, 0.0062]
15	238.31	156.11	0.0166	0.000 34	[0.0095, 0.0076]
16	139.01	91.39	0.0278	0.000 28	[0.0076, 0.0155, 0.0085]
17	307.58	261.39	0.0529	0.000 55	[0.0155, 0.0055]
18	249.95	245.07	0.0066	0.000 15	[0.0045, 0.0074, 0.0055]
19	274.84	16.84	0.0172	0.000 27	[0.0045, 0.0056]
20	174.69	202.76	0.0082	0.000 19	[0.0056, 0.0074, 0.011]

TABLE V. Calibration table values for 21 physical qubits used for demonstrations (`ibm_torino`). The values obtained from the calibration table correspond to $|00\rangle_L$ with ten rounds. The table lists the properties of physical qubits for the $d = 3$ structure selected within the hardware. These include T1 and T2 times, as well as error rates of the readout and single- (SQ) and two-qubit (TQ) gates. More than one two-qubit gate can be implemented and their error rates are listed accordingly.

Qubit	T1 (μs)	T2 (μs)	Readout error rate	SQ gate error rate	TQ gate error rate
0	142.71	116.51	0.085	0.000 53	[0.0044, 0.0083]
1	132.53	35.6	0.0143	0.000 65	[0.0082, 0.0044]
2	127.68	203.37	0.0382	0.000 37	[0.0082, 0.0164]
3	204.95	136.37	0.0077	0.000 21	[0.0164, 0.0044]
4	135.37	25.33	0.0393	0.000 67	[0.0063, 0.0079, 0.0044]
5	242.65	76.18	0.015	0.000 21	[0.0063, 0.0044]
6	124.4	187.37	0.021 67	0.000 17	[0.0062, 0.0029, 0.0044]
7	38.45	45.24	0.0321	0.002 13	[0.0062, 0.0071]
8	205.84	157.08	0.0315	0.000 24	[0.0049, 0.0071, 0.0023]
9	232.22	128.96	0.0126	0.000 22	[0.0034, 0.0083]
10	32.42	45.8	0.0131	0.000 87	[0.0085, 0.0079]
11	248.0	160.5	0.0117	0.0002	[0.0046, 0.0023]
12	215.6	122.98	0.0282	0.000 33	[0.0053, 0.0034]
13	263.96	194.63	0.0172	0.000 38	[0.0053, 0.0176]
14	160.04	214.99	0.0129	0.000 16	[0.0024, 0.0176, 0.0043]
15	174.22	200.89	0.0102	0.000 18	[0.0042, 0.0024]
16	164.92	150.59	0.017	0.000 32	[0.0042, 0.0085, 0.0085]
17	23.84	22.13	0.0243	0.001 17	[0.0085, 0.0068]
18	219.47	64.31	0.0163	0.000 27	[0.0033, 0.0029, 0.0068]
19	265.79	111.69	0.0119	0.000 32	[0.0045, 0.0033]
20	158.13	187.47	0.0143	0.000 19	[0.0045, 0.0086, 0.0046]

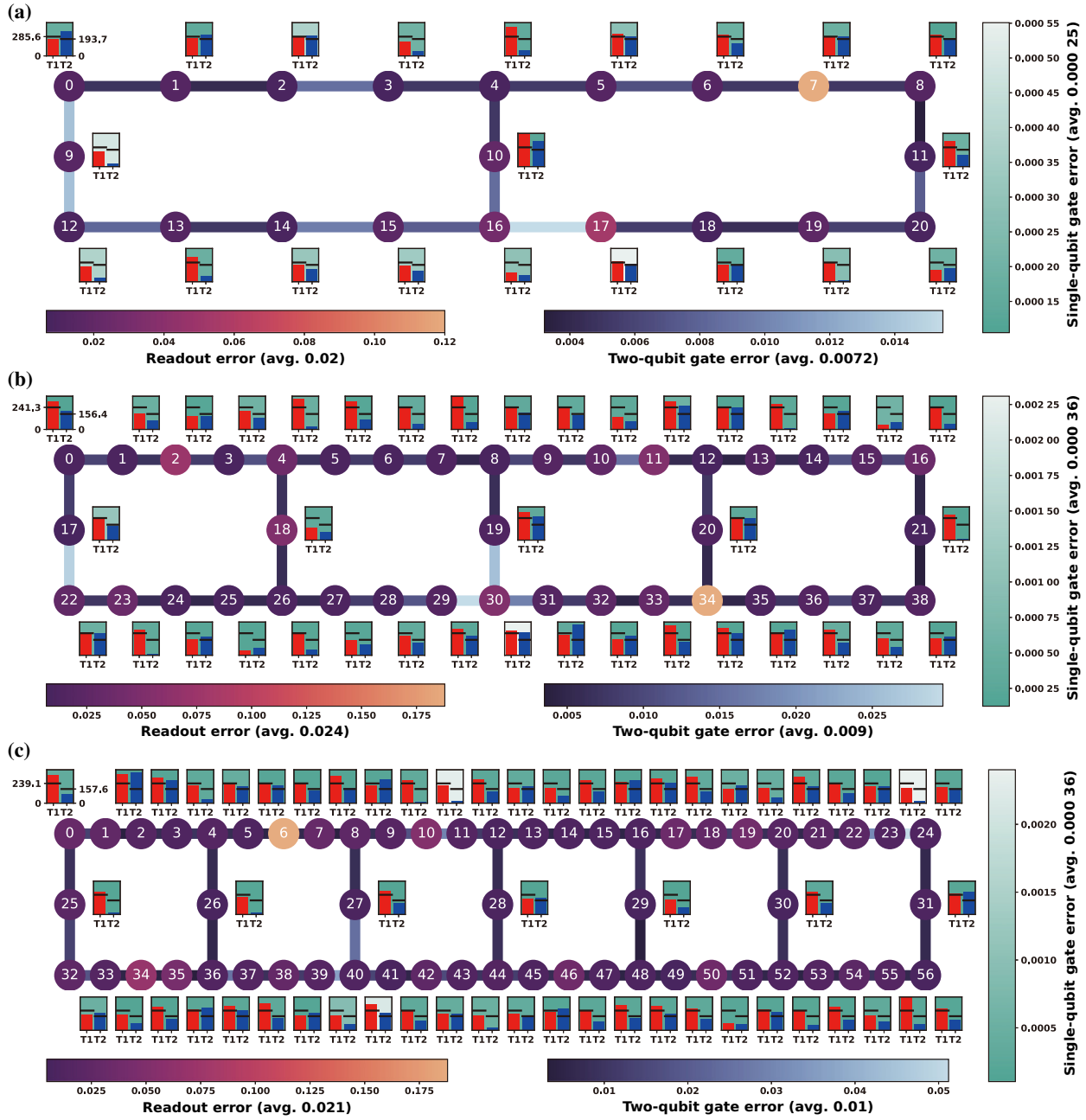

FIG. 22. Hardware specifications of physical qubits with colormaps and bar charts (`ibm_brisbane`). The graph with subplots shows the physical qubits' specification of the structure in (a) $d = 3$, (b) $d = 5$, and (c) $d = 7$ with the $|00\rangle_L$ state. Each node and edge correspond to the physical qubit and connectivity of a two-qubit gate. The error rates of the readout and two-qubit gate (ECR) are displayed with colors. The subplot next to every node shows T1 and T2 times corresponding to their physical qubit. The average T1 and T2 times are plotted as the black lines on the left and right sides. Their values can be seen on the subplot of the 0th physical qubit. The background of a subplot is colored corresponding to the error rate of the single-qubit gate.

$$F_{11} = \tfrac{1}{4}(1 - \langle Z_L^C\rangle - \langle Z_L^T\rangle + \langle Z_L^C Z_L^T\rangle). \tag{E4}$$

The fidelity calculations involve measuring the expectation values of logical Pauli operators. Specifically, $\langle Z_L^C\rangle$, $\langle Z_L^T\rangle$, and $\langle Z_L^C Z_L^T\rangle$ represent the expectation values of the logical Z operator on the logical control qubit, the logical target qubit, and both logical qubits, respectively. The same process can be applied to the logical qubits in the X basis when dealing with the phase-flip code with flag qubits:

$$F_{++} = \tfrac{1}{4}(1 + \langle X_L^C\rangle + \langle X_L^T\rangle + \langle X_L^C X_L^T\rangle), \tag{E5}$$

$$F_{+-} = \tfrac{1}{4}(1 + \langle X_L^C\rangle - \langle X_L^T\rangle - \langle X_L^C X_L^T\rangle), \tag{E6}$$

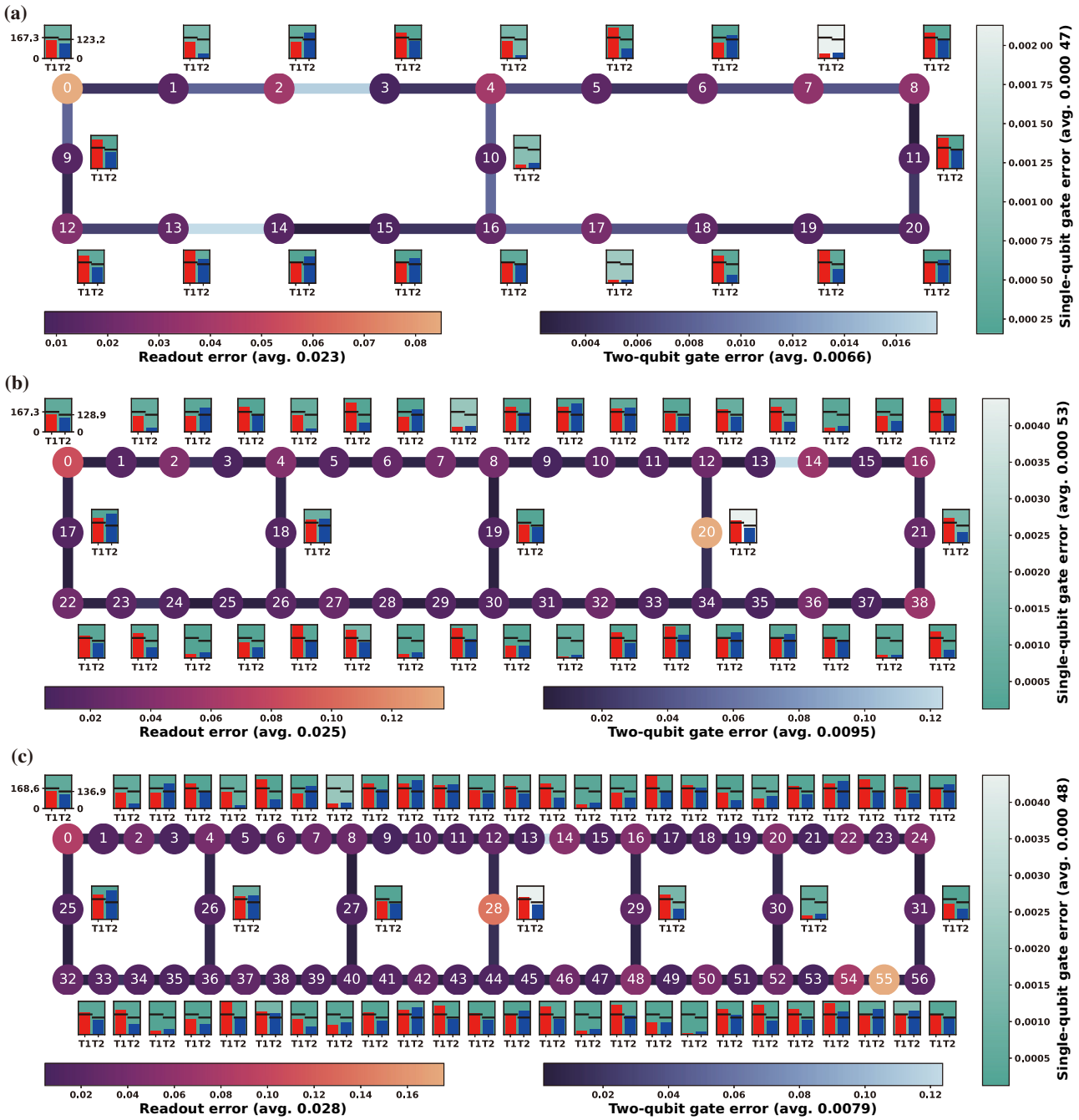

FIG. 23. Hardware specifications of physical qubits with colormaps and bar charts (`ibm_torino`). The graph with subplots shows the physical qubits' specification of the structure in (a) $d = 3$, (b) $d = 5$, and (c) $d = 7$ with the $|00\rangle_L$ state. Each node and edge correspond to the physical qubit and connectivity of a two-qubit (CZ) gate. The error rates of the readout and two-qubit gate are displayed with colors. The subplot next to every node shows T1 and T2 times corresponding to their physical qubit. The average T1 and T2 times are plotted as the black lines on the left and right sides. Their values can be seen on the subplot of the 0th physical qubit. The background of a subplot is colored corresponding to the error rate of the single-qubit gate.

$$F_{-+} = \tfrac{1}{4}(1 - \langle X_L^C\rangle + \langle X_L^T\rangle - \langle X_L^C X_L^T\rangle), \quad \text{(E7)}$$

$$F_{--} = \tfrac{1}{4}(1 - \langle X_L^C\rangle - \langle X_L^T\rangle + \langle X_L^C X_L^T\rangle). \quad \text{(E8)}$$

Figure 18 displays the logical error rates for four states with noncumulative bar charts across the syndrome rounds, obtained from `ibm_sherbrooke`. The reduction in logical errors regardless of the initial state and basis is observed when more physical qubits are utilized. However, we observe that logical errors for the X basis are higher than those for the Z basis across all syndrome extraction rounds. However, the $d = 5$ structure exhibits a lower

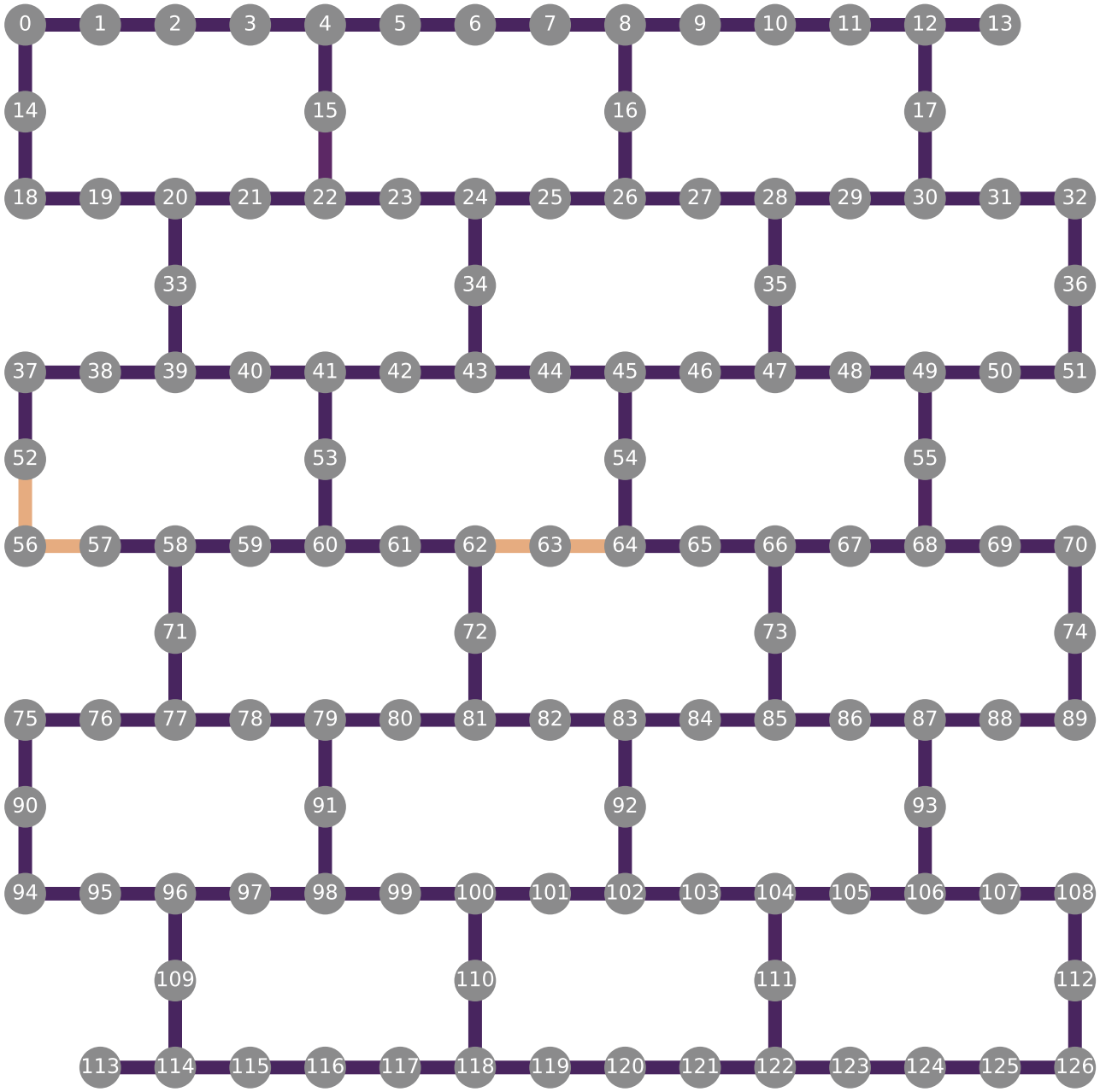

FIG. 24. Connectivity of `ibm_sherbrooke`. Each node and edge correspond to a qubit and its connectivity with other nearest-neighbor qubits, respectively. The edges highlighted in orange represent the connections where the two-qubit (ECR) gate has an error rate of 1. In contrast, other edges, indicated in purple, have error rates lower than 1.

logical error rate than the $d = 3$ structure for all cases. The same figure obtained from `ibm_torino` is shown in Fig. 19. Even though, in some cases, the larger structure does not have lower logical errors than the smaller one the trends of logical error rates across all syndrome extraction circuit rounds indicate the error suppression overall. This could be attributed to the inherent errors in the **S2** syndrome bit region that has been covered in Appendix E 3.

Figure 20 shows the logical error rates as a function of the number of syndrome extraction circuits and the code distance, obtained from the classical simulator (Stim) for `ibm_torino` and `ibm_sherbrooke`. The structure with more physical qubits has a lower logical error rate for all syndrome circuit rounds. As we increase the number of syndrome extraction circuit rounds, the logical error rates increase for both bases regardless of devices and distance. This is attributed to the increase in the possibility

of gate-flow errors, which may lead to ambiguity for a decoder when calculating a correction operator due to correlated errors.

Figure 21 presents results obtained from `ibm_brisbane`. When the logical qubits are prepared in the Z basis, the logical error rates notably decrease as we increase the number of physical qubits. However, while there is a decrease in logical error rates between $d = 3$ and $d = 5$, the $d = 7$ structure exhibits higher logical error rates across the syndrome extraction rounds than that of the $d = 5$ structure. It could be related to the biased errors during the execution of the logical CNOT gate, as discussed in Appendix E 2.

APPENDIX F: FURTHER FIGURES

We performed the demonstrations on April 23, 2024, using the `ibm_brisbane` device and on April 23 using the `ibm_torino` device. In this section, we present the calibration tables for these demonstrations for the $d = 3$ structure in Tables IV and V, as well as the hardware specifications of physical qubits across different code distances ($d = \{3, 5, 7\}$) in Figs. 22 and 23, respectively. Furthermore, Fig. 24 illustrates the infeasibility of certain two-qubit gates on `ibm_sherbrooke`, which prevented the embedding of the proposed quantum memory experiment with the $d = 7$ structure.

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

[2] Y. Kim, A. Eddins, S. Anand, K. X. Wei, E. van den Berg, S. Rosenblatt, H. Nayfeh, Y. Wu, M. Zaletel, K. Temme, A. Kandala, *et al.*, Evidence for the utility of quantum computing before fault tolerance, Nature **618**, 500 (2023).

[3] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell, *et al.*, Quantum supremacy using a programmable superconducting processor, Nature **574**, 505 (2019).

[4] S. Moses, C. Baldwin, M. Allman, R. Ancona, L. Ascarrunz, C. Barnes, J. Bartolotta, B. Bjork, P. Blanchard, M. Bohn, *et al.*, A race-track trapped-ion quantum processor, Phys. Rev. X **13**, 041052 (2023).

[5] M. DeCross, R. Haghshenas, M. Liu, E. Rinaldi, J. Gray, Y. Alexeev, C. H. Baldwin, J. P. Bartolotta, M. Bohn, E. Chertkov, *et al.*, The computational power of random quantum circuits in arbitrary geometries, ArXiv:2406.02501.

[6] J. Wurtz, A. Bylinskii, B. Braverman, J. Amato-Grill, S. H. Cantu, F. Huber, A. Lukin, F. Liu, P. Weinberg, J. Long, *et al.*, Aquila: QuEra's 256-qubit neutral-atom quantum computer, ArXiv:2306.11727.

[7] C. Gidney and M. Ekerå, How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits, Quantum **5**, 433 (2021).

[8] D. Gottesman, Stabilizer codes and quantum error correction, ArXiv:quant-ph/9705052.

[9] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, Surface codes: Towards practical large-scale quantum computation, Phys. Rev. A **86**, 032324 (2012).

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

[11] E. Knill, R. Laflamme, and W. H. Zurek, Resilient quantum computation: Error models and thresholds, Proc. R. Soc. Lond., Ser. A **454**, 365 (1998).

[12] J. Preskill, Reliable quantum computers, Proc. R. Soc. A **454**, 385 (1998).

[13] D. S. Wang, A. G. Fowler, A. M. Stephens, and L. C. L. Hollenberg, Threshold error rates for the toric and planar codes, Quantum Inf. Comput. **10**, 456 (2010).

[14] J. Kelly, R. Barends, A. G. Fowler, A. Megrant, E. Jeffrey, T. C. White, D. Sank, J. Y. Mutus, B. Campbell, Y. Chen, *et al.*, State preservation by repetitive error detection in a superconducting quantum circuit, Nature **519**, 66 (2015).

[15] Google Quantum AI, Exponential suppression of bit or phase errors with cyclic error correction, Nature **595**, 383 (2021).

[16] Y. Kim, H. Kim, J. Kang, W. Choi, and Y. Kwon, Effectiveness of the syndrome extraction circuit with flag qubits on IBM quantum hardware, ArXiv:2403.10217.

[17] J. R. Wootton, Benchmarking near-term devices with quantum error correction, Quantum Sci. Technol. **5**, 044004 (2020).

[18] Google Quantum AI, Suppressing quantum errors by scaling a surface code logical qubit, Nature **614**, 676 (2023).

[19] C. K. Andersen, A. Remm, S. Lazar, S. Krinner, N. Lacroix, G. J. Norris, M. Gabureac, C. Eichler, and A. Wallraff, Repeated quantum error detection in a surface code, Nat. Phys. **16**, 875 (2020).

[20] S. Krinner, N. Lacroix, A. Remm, A. D. Paolo, E. Genois, C. Leroux, C. Hellings, S. Lazar, F. Swiadek, J. Herrmann, *et al.*, Realizing repeated quantum error correction in a distance-three surface code, Nature **605**, 669 (2022).

[21] Y. Zhao, Y. Ye, H.-L. Huang, Y. Zhang, D. Wu, H. Guan, Q. Zhu, Z. Wei, T. He, S. Cao, *et al.*, Realization of an error-correcting surface code with superconducting qubits, Phys. Rev. Lett. **129**, 030501 (2022).

[22] R. Harper and S. T. Flammia, Learning correlated noise in a 39-qubit quantum processor, PRX Quantum **4**, 040311 (2023).

[23] C. Ryan-Anderson, J. G. Bohnet, K. Lee, D. Gresh, A. Hankin, J. P. Gaebler, D. Francois, A. Chernoguzov, D. Lucchetti, N. C. Brown, *et al.*, Realization of real-time fault-tolerant quantum error correction, Phys. Rev. X **11**, 041058 (2021).

[24] D. Nigg, M. Mueller, E. A. Martinez, P. Schindler, M. Hennrich, T. Monz, M. A. Martin-Delgado, and R. Blatt, Quantum computations on a topologically encoded qubit, Science **345**, 302 (2014).

[25] M. H. Devoret and R. J. Schoelkopf, Superconducting circuits for quantum information: An outlook, Science **339**, 1169 (2013).

[26] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter, Elementary gates for quantum computation, Phys. Rev. A **52**, 3457 (1995).

[27] P. O. Boykin, T. Mor, M. Pulver, V. Roychowdhury, and F. Vatan, A new universal and fault-tolerant quantum basis, Inf. Process. Lett. **75**, 101 (2000).

[28] E. T. Campbell, B. M. Terhal, and C. Vuillot, Roads towards fault-tolerant universal quantum computation, Nature **549**, 172 (2017).

[29] B. Eastin and E. Knill, Restrictions on transversal encoded quantum gate sets, Phys. Rev. Lett. **102**, 110502 (2009).

[30] G. P. Gehér, C. McLauchlan, E. T. Campbell, A. E. Moylett, and O. Crawford, Error-corrected Hadamard gate simulated at the circuit level, Quantum **8**, 1394 (2024).

[31] B. Zeng, A. Cross, and I. L. Chuang, Transversality versus universality for additive quantum codes, IEEE Trans. Inf. Theory **57**, 6272 (2011).

[32] A. Kubica and M. E. Beverland, Universal transversal gates with color codes: A simplified approach, Phys. Rev. A **91**, 032330 (2015).

[33] T. Jochym-O'Connor and R. Laflamme, Using concatenated quantum codes for universal fault-tolerant quantum gates, Phys. Rev. Lett. **112**, 010505 (2014).

[34] J. T. Anderson, G. Duclos-Cianci, and D. Poulin, Fault-tolerant conversion between the Steane and Reed-Muller quantum codes, Phys. Rev. Lett. **113**, 080501 (2014).

[35] L. Postler, S. Heußen, I. Pogorelov, M. Rispler, T. Feldker, M. Meth, C. D. Marciniak, R. Stricker, M. Ringbauer, R. Blatt, P. Schindler, M. Müller, and T. Monz, Demonstration of fault-tolerant universal quantum gate operations, Nature **605**, 675 (2022).

[36] C. Ryan-Anderson, N. C. Brown, M. S. Allman, B. Arkin, G. Asa-Attuah, C. Baldwin, J. Berg, J. G. Bohnet, S. Braxton, N. Burdick, J. P. Campora, A. Chernoguzov, J. Esposito, B. Evans, D. Francois, J. P. Gaebler, T. M. Gatterman, J. Gerber, K. Gilmore, D. Gresh, *et al.*, Implementing fault-tolerant entangling gates on the five-qubit code and the color code, ArXiv:2208.01863.

[37] D. Honciuc Menendez, A. Ray, and M. Vasmer, Implementing fault-tolerant non-Clifford gates using the [[8,3,2]] color code, Phys. Rev. A **109**, 062438 (2024).

[38] D. Bluvstein, S. J. Evered, A. A. Geim, S. H. Li, H. Zhou, T. Manovitz, S. Ebadi, M. Cain, M. Kalinowski, D. Hangleiter, *et al.*, Logical quantum processor based on reconfigurable atom arrays, Nature **626**, 58 (2024).

[39] M. P. da Silva, C. Ryan-Anderson, J. M. Bello-Rivas, A. Chernoguzov, J. M. Dreiling, C. Foltz, F. Frachon, J. P. Gaebler, T. M. Gatterman, L. Grans-Samuelsson, D. Hayes, N. Hewitt, J. Johansen, D. Lucchetti, M. Mills, S. A. Moses, B. Neyenhuis, A. Paz, J. Pino, P. Siegfried, *et al.*, Demonstration of logical qubits and repeated error correction with better-than-physical error rates, ArXiv:2404.02280.

[40] B. Hetenyi and J. R. Wootton, Creating entangled logical qubits in the heavy-hex lattice with topological codes, Quantum **5**, 040334 (2024).

[41] C. Ryan-Anderson, N. C. Brown, C. H. Baldwin, J. M. Dreiling, C. Foltz, J. P. Gaebler, T. M. Gatterman, N. Hewitt, C. Holliman, C. V. Horst, J. Johansen, D. Lucchetti, T. Mengle, M. Matheny, Y. Matsuoka, K. Mayer, M. Mills, S. A. Moses, B. Neyenhuis, J. Pino, *et al.*, High-fidelity teleportation of a logical qubit using transversal gates and lattice surgery, Science **385**, 1327 (2024).

[42] K. Mayer, C. Ryan-Anderson, N. Brown, E. Durso-Sabina, C. H. Baldwin, D. Hayes, J. M. Dreiling, C. Foltz, J. P. Gaebler, T. M. Gatterman, *et al.*, Benchmarking logical three-qubit quantum Fourier transform encoded in the Steane code on a trapped-ion quantum computer, ArXiv:2404.08616.

[43] D. C. McKay, I. Hincks, E. J. Pritchett, M. Carroll, L. C. Govia, and S. T. Merkel, Benchmarking quantum processor performance at scale, ArXiv:2311.05933.

[44] D. P. DiVincenzo and P. W. Shor, Fault-tolerant error correction with efficient quantum codes, Phys. Rev. Lett. **77**, 3260 (1996).

[45] A. M. Steane, Active stabilization, quantum computation, and quantum state synthesis, Phys. Rev. Lett. **78**, 2252 (1997).

[46] E. Knill, Quantum computing with realistically noisy devices, Nature **434**, 39 (2005).

[47] J. Roffe, Quantum error correction: An introductory guide, Contemp. Phys. **60**, 226 (2019).

[48] R. Chao and B. W. Reichardt, Quantum error correction with only two extra qubits, Phys. Rev. Lett. **121**, 050502 (2018).

[49] R. Chao and B. W. Reichardt, Flag fault-tolerant error correction for any stabilizer code, PRX Quantum **1**, 010302 (2020).

[50] E. H. Chen, T. J. Yoder, Y. Kim, N. Sundaresan, S. Srinivasan, M. Li, A. D. Corcoles, A. W. Cross, and M. Takita, Calibrated decoders for experimental quantum error correction, Phys. Rev. Lett. **128**, 110504 (2022).

[51] N. Sundaresan, T. J. Yoder, Y. Kim, M. Li, E. H. Chen, G. Harper, T. Thorbeck, A. W. Cross, A. D. Córcoles, and M. Takita, Demonstrating multi-round subsystem quantum error correction using matching and maximum likelihood decoders, Nat. Commun. **14**, 2852 (2023).

[52] Y. Kim, J. Kang, and Y. Kwon, Design of quantum error correcting code for biased error on heavy-hexagon structure, Quantum Inf. Process **22**, 230 (2023).

[53] C. Chamberland, G. Zhu, T. J. Yoder, J. B. Hertzberg, and A. W. Cross, Topological and subsystem codes on low-degree graphs with flag qubits, Phys. Rev. X **10**, 011022 (2020).

[54] C. Benito, E. Löpez, B. Peropadre, and A. Bermudez, Comparative study of quantum error correction strategies for the heavy-hexagonal lattice, Quantum **9**, 1623 (2025).

[55] N. H. Nickerson and B. J. Brown, Analysing correlated noise on the surface code using adaptive decoding algorithms, Quantum **3**, 131 (2019).

[56] M. McEwen, D. Kafri, Z. Chen, J. Atalaya, K. J. Satzinger, C. Quintana, P. V. Klimov, D. Sank, C. Gidney, A. G. Fowler, *et al.*, Removing leakage-induced correlated errors in superconducting quantum error correction, Nat. Commun. **12**, 1761 (2021).

[57] M. Cain, C. Zhao, H. Zhou, N. Meister, J. P. Ataides, A. Jaffe, D. Bluvstein, and M. D. Lukin, Correlated decoding of logical algorithms with transversal gates, Phys. Rev. Lett. **133**, 240602 (2024).

[58] A. Paler and A. G. Fowler, Pipelined correlated minimum weight perfect matching of the surface code, Quantum **7**, 1205 (2023).

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

[60] S. Bravyi, G. Duclos-Cianci, D. Poulin, and M. Suchara, Subsystem surface codes with three-qubit check operators, ArXiv:1207.1443.

[61] J. P. Bonilla Ataides, D. K. Tuckett, S. D. Bartlett, S. T. Flammia, and B. Brown, The XZZX surface code, Nat. Commun. **12**, 2172 (2021).

[62] A. G. Fowler, A. C. Whiteside, and L. C. L. Hollenberg, Towards practical classical processing for the surface code, Phys. Rev. Lett. **108**, 180501 (2012).

[63] A. G. Fowler, Minimum weight perfect matching of fault-tolerant topological quantum error correction in average $O(1)$ parallel time, Quantum Inf. Comput. **15**, 145 (2015).

[64] C. Gidney, Stim: A fast stabilizer circuit simulator, Quantum **5**, 497 (2021).

[65] O. Higgott, PyMatching: A Python package for decoding quantum codes with minimum-weight perfect matching, ACM Trans. Quantum Comput. **3**, 1 (2022).

[66] O. Higgott and C. Gidney, Sparse blossom: Correcting a million errors per core second with minimum-weight matching, Quantum **9**, 1600 (2025).

[67] IBM Quantum and Community. Qiskit: An open-source framework for quantum computing, Zenodo.

[68] S. Gicev, L. C. L. Hollenberg, and M. Usman, A fast and scalable artificial neural network syndrome decoder for surface codes, Quantum **7**, 1058 (2023).

[69] B. Hall, S. Gicev, and M. Usman, Artificial neural network syndrome decoding on IBM quantum processors, Phys. Rev. Res. **6**, L032004 (2024).

[70] J. Bausch, A. W. Senior, F. J. H. Heras, T. Edlich, A. Davies, M. Newman, C. Jones, K. Satzinger, M. Y. Niu, S. Blackwell, G. Holland, D. Kafri, J. Atalaya, C. Gidney, D. Hassabis, S. Boixo, H. Neven, and P. Kohli, Learning to decode the surface code with a recurrent, transformer-based neural network, ArXiv:2310.05900.

[71] N. Sundaresan, I. Lauer, E. Pritchett, E. Magesan, P. Jurcevic, and J. M. Gambetta, Reducing unitary and spectator errors in cross resonance with optimized rotary echoes, PRX Quantum **1**, 020318 (2020).

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