
Teleportation transition of surface codes on a superconducting quantum processor

In the format provided by the
authors and unedited

CONTENTS

S1. Experimental setup and noise characterization	1
A. Effective incoherent noise strength	1
B. Qubits allocation	2
C. Readout error mitigation	3
D. Numerical simulations of parity values	3
S2. Rényi relative entropy and Ising phase transition	4
A. Analytical mapping	5
B. Classical error injection for experimental bitstrings	5
S3. Tunable teleportation phase transition	6
A. Teleportation in X basis	6
B. Teleportation in a self-dual basis	6
C. Learnability and teleportation transition	7
S4. Surface code state preparation	7
A. Circuit for preparing Alice in a logical state	7
B. Circuit for preparing both Alice and Bob in a logical state	9
C. Preparing logical states by measuring stabilizers	9
S5. Decoding methods	10
S6. Integration of quantum error correction with the logical teleportation protocols	12
References	13

S1. EXPERIMENTAL SETUP AND NOISE CHARACTERIZATION

A. Effective incoherent noise strength

We estimate the effective incoherent noise strength p for our teleportation experiment using the measured parities without applying readout error correction. Since the stabilizer expectation values depend on the stabilizer weight, we extract p via

$$1 - 2p = \left(\frac{\sum_{s \in \mathcal{S}_2} \langle A_s \rangle^2 + \sum_{s \in \mathcal{S}_4} \langle A_s \rangle}{\# \mathcal{S}_2 + \# \mathcal{S}_4} \right)^{1/4}, \quad (\text{S1})$$

where $\mathcal{S}_w$ denote the set of Z -type faces defined by $w \in \{2, 4\}$ qubits and $\# \mathcal{S}_w$ is the number of elements in $\mathcal{S}_w$. The extracted values are provided in Table S1.

TABLE S1. **Effective incoherent noise strength for each code distance.** These noise strength values are derived from measured stabilizer parities with teleportation parameters fixed at $\theta = \pi/2$ and $t = 0$. The values being a few percent are due to the joint action of the mid-circuit noises and the readout errors that lead to the infidelity of the measurement outcomes.

Code distance, d	2	3	4	5	6	7
Effective incoherent noise strength, p	0.0237(5)	0.0306(5)	0.0364(4)	0.0380(3)	0.0484(4)	0.0505(2)

The unitary circuit of preparing the surface code scales linearly with the code distance. Assuming that the accumulated noise increases with the circuit depth, the finite noise threshold poses an upper bound on the circuit depth and thus the code distance. We can approximate the effective noise strength as a linear function $p(d) = a(d + d_0)$, where the maximum code distance that remains below the threshold $p_c = 0.109$ is then given by

$$d_{\max} = \left\lfloor \frac{p_c}{a} - d_0 \right\rfloor. \quad (\text{S2})$$

Fitting the experimentally observed noise levels in our quantum processor provides $a = 0.0054$ and $d_0 = 2.5$. Consequently, the maximum distance for a surface code below the threshold is approximately $d_{\max} = 17$, which can be further increased by improving qubit coherence times and gate fidelities.

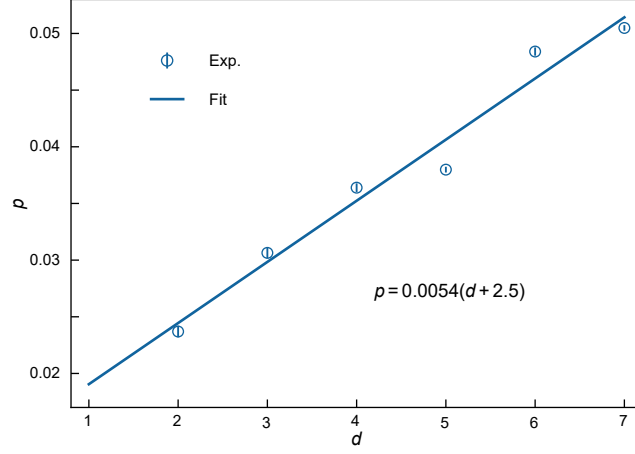


FIG. S1. **Effective bit-flip error strength accumulated during circuit implementation.** The effective error strength p is modelled as a linear function of the code distance d , yielding the relationship $p = 0.0054(d + 2.5)$. Experimental results are represented by circles with error bars, while the dashed line indicates the linear fit.

B. Qubits allocation

To implement the teleportation experiment, we interlace Alice's and Bob's qubits such that each of Bob's qubits is positioned at the immediate bottom-right lattice site relative to its corresponding qubit in Alice. However, due to the presence of non-functional qubits along the boundaries, this precise configuration is feasible only for code distances $d \leq 6$. Consequently, as illustrated in Fig. S2, we introduce a spatial rearrangement of Bob's qubits, positioning each at a nearest-neighbour lattice site (not precisely the bottom-right) relative to its corresponding qubit in Alice.

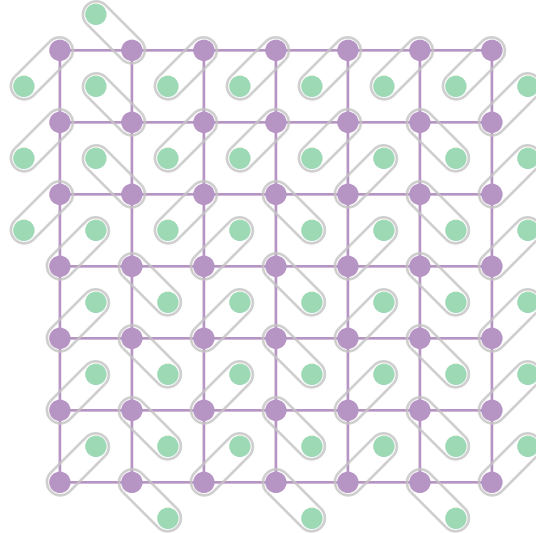


FIG. S2. **Spatial arrangement of Alice's and Bob's qubits for the distance-7 teleportation experiment.** Purple circles denote Alice's qubits, and green circles denote Bob's qubits. Grey outlines indicate the one-to-one pairing of corresponding qubits used in the teleportation experiment.

C. Readout error mitigation

Here we show details for readout error mitigation for measuring the stabilizer expectation values to validate our prepared surface codes targetting the stabilizer states, which should be distinguished from the task of decoding a logical qubit. Measurement of superconducting qubits is vulnerable to various types of errors. As a result, the measured observables can be significantly degraded with respect to their actual values. The dominant error mechanisms include: (1) separation errors arising from a finite signal-to-noise ratio in the amplifying chain, which limit the state distinguishability, (2) qubit energy relaxation occurring during the readout window, which can change the qubit state before the measurement is completed, and (3) unwanted coherent transitions between qubit states, induced, for instance, by stray couplings or measurement-induced leakage out of the computational subspace.

In quantum error correction circuits, measurement errors are typically addressed through decoding algorithms that analyse the measured syndrome patterns to identify and correct the errors. However, this approach is not directly applicable to measurement outcomes of the physical qubits themselves. For quantities such as single-qubit expectation values and multi-qubit parities, they are not topologically protected and there is no associated decoding algorithm. The measurement results must be corrected by an explicit readout-error mitigation procedure.

The effect of readout errors can be quantitatively modelled by a response matrix [1]. In this description, the noisy measurement process is represented as a classical channel that maps the actual bitstring s' to the observed bitstring s with a certain conditional probability. Denoting by $P_a(s')$ the actual probability of obtaining bitstring s' and by $P_m(s)$ the measured probability of s , the two distributions are related via

$$P_m(s) = \sum_{s'} M(s | s') P_a(s'), \quad (\text{S3})$$

where $M(s | s')$ are the elements of the response matrix $\mathbf{M}$. By construction, $M(s | s')$ specifies the probability of measuring the measurement outputs s given that the system is prepared in the state $|s'\rangle$. In matrix notation, this relationship is written as $\mathbf{P}_m = \mathbf{M} \mathbf{P}_a$.

In the experiment, the response matrix is calibrated by preparing the system in a complete set of product states $\{|s'\rangle\}$ and performing many repeated measurements for each preparation. For a given input state $|s'\rangle$, the observed empirical frequencies of the outcomes s define the column of $\mathbf{M}$ associated with s' . Assuming that the response matrix is invertible and that the calibration is sufficiently accurate, one can then estimate the actual distribution by applying the inverse matrix to the measured data $\mathbf{P}_a = \mathbf{M}^{-1} \mathbf{P}_m$. Subsequently, corrected expectation values of observables, such as multi-qubit parities, can be obtained by evaluating them with respect to the reconstructed distribution $\mathbf{P}_a$.

While conceptually straightforward, this full response matrix approach faces severe scalability limitations. The dimension of the response matrix is $2^N \times 2^N$, so both its experimental calibration and numerical inversion require resources that grow exponentially with the qubit number N . Consequently, a brute-force characterization and inversion of the full response matrix is feasible only for small systems and quickly becomes impractical as the number of qubits increases.

In superconducting qubit platforms, however, the dominant contribution to measurement errors arises from processes that affect each qubit approximately independently, such as individual readout misclassification or single-qubit relaxation during measurement. Under this assumption, the measurement noise can be accurately described by a tensor-product noise model, in which the full response matrix is constructed as

$$\mathbf{M} = \bigotimes_{n=1}^N \mathbf{M}_n, \quad (\text{S4})$$

where $\mathbf{M}_n$ is the single-qubit response matrix associated with qubit Q_n . This tensor product structure dramatically reduces the experimental and computational overhead. Each single-qubit response matrix is a 2×2 matrix that can be calibrated independently using only the two single-qubit preparations $\{|0\rangle, |1\rangle\}$, and the inversion of $\mathbf{M}$ reduces to inverting N single-qubit response matrices and taking their tensor product, $\mathbf{M}^{-1} = \bigotimes_{n=1}^N \mathbf{M}_n^{-1}$.

D. Numerical simulations of parity values

We perform numerical simulations using the *Stim* [2] package to obtain the simulated parity values before and after teleportation, modeling the experimentally introduced noise as Pauli error channels. The corresponding error rates are shown in Fig. 6 of the main text, where the median Pauli error rates of single-qubit and two-qubit gates

are approximately 0.08% and 0.44%, respectively. For qubits idling during the CZ-gate layers, we include a Pauli noise channel with a median error rate of 0.10%. All these Pauli error rates are characterized using cross-entropy benchmarking (XEB). Readout errors are not included in the simulation, because readout error mitigation is applied to the experimental data. As shown in Fig. S3, the simulated results are in close agreement with the experimental observations, also showing closely correlated non-uniform stabilizer fidelity distribution in space. We attribute the remaining discrepancy mainly to correlated noise mechanisms that cannot be accurately captured within our Pauli-error model. A correlated noise study is an open question for future work.

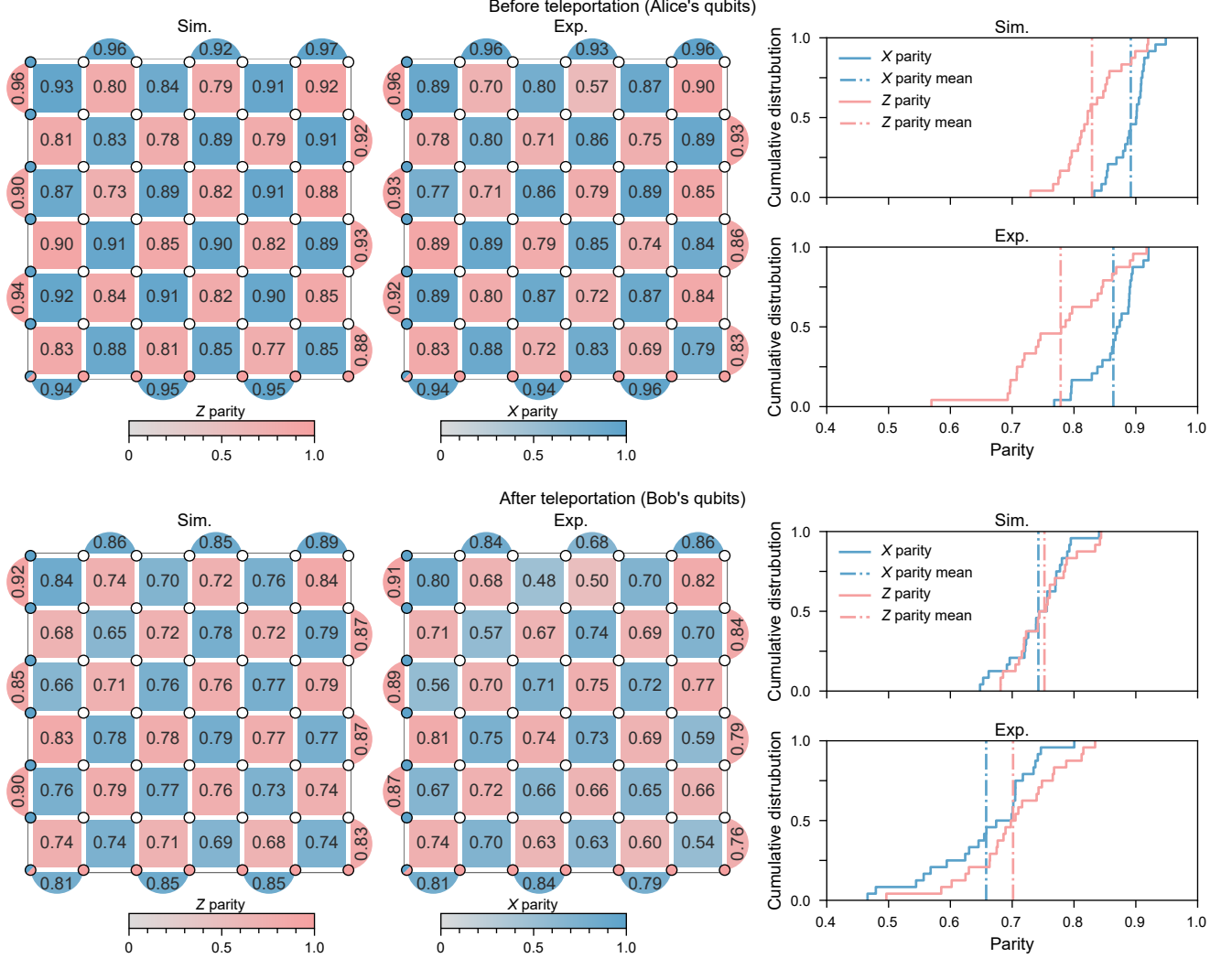


FIG. S3. Parities obtained from experiment and simulation for code distance $d = 7$. The top panel shows results before teleportation (Alice's qubits), and the bottom panel shows results after teleportation (Bob's qubits). To display the cumulative distributions, weight-2 parities are squared prior to analysis, as described in the main text. Before teleportation, the simulated mean X-type and Z-type parities are 0.89(3) and 0.83(5), respectively. The corresponding experimental mean values are 0.86(4) for X-type parities and 0.78(8) for Z-type parities. After teleportation, the simulated mean values are 0.74(5) and 0.75(5), while the experimental mean values are 0.66(8) and 0.70(8).

S2. RÉNYI RELATIVE ENTROPY AND ISING PHASE TRANSITION

The topological order entails two locally indistinguishable yet globally orthogonal logical states, even in the presence of finite noise below the threshold. This can be detected by the quantum coherent information or quantum relative entropy (Kullback-Leibler divergence) [3]. However, von Neumann entropy is not directly accessible in experiment,

and we turn to the second Rényi relative entropy accessible in experiment:

$$D_2(\rho_0||\rho_1) = -\ln \text{tr}(\rho_0\rho_1) + \ln \text{tr}(\rho_0^2) . \quad (\text{S5})$$

The purity can be measured by a swap measurement jointly performed by Alice and Bob:

$$\text{tr}(\rho_\mu\rho_\nu) = \text{tr}(\rho_\mu \otimes \rho_\nu \cdot \text{SWAP}^{\otimes N}) , \quad (\text{S6})$$

where $\text{SWAP} = (1 + X \otimes X + Y \otimes Y + Z \otimes Z)/2$ is the swap operator. The SWAP expectation value can be measured by a Bell measurement i.e., applying a controlled-NOT (CNOT) and then measure the two qubits in X and Z basis, respectively, which gives bit-string for XX , ZZ , as well as YY interaction, see Fig. 7 in the main text. In this way, we can measure the relative Rényi entropy to signal whether the experimentally prepared surface codes remain under the 2-replica threshold (the Ising threshold, slightly larger than the 1-replica Nishimori threshold) [3]. In classical numerical simulation, we consider independent bit-flip and phase-flip noises to be compared with the experimental results.

A. Analytical mapping

The independent phase-flip and bit-flip errors in Alice and Bob lead to the same consequence as independent readout errors. The readout errors can be pulled back the circuit into only phase-flip in Alice and bit-flip in Bob, see Fig. 7 in the main text. In this case, $\ln(\rho_0^2)$ becomes $\ln(\rho_{A_0} \cdot \rho_{B_0})$, where ρ_{A_0} and ρ_{B_0} represent the density matrices for Alice and Bob respectively. Since Alice and Bob suffer from different kind of noise, the contribution of overlap of ρ_{A_0} and ρ_{B_0} from those error chain E_A and E_B such that E_A is a contractible loop and E_B is a loop. In this case, we have:

$$\ln \text{tr}(\rho_{A_0} \cdot \rho_{B_0}) = P_{E_A}(P'_{E_B} + P'_{E_B Z_L})$$

where P_{E_A} and P_{E_B} is given by the partition function of the Ising model:

$$P_{E_A} = \sum_{\{s\}} \prod_{\langle ij \rangle} e^{\beta J_x s_i s_j} ,$$

$$P'_{E_B} = \sum_{\{s\}} \prod_{\langle ij \rangle} e^{\beta J_z s'_i s'_j} .$$

And as the same, for $\ln(\rho_{A_0} \cdot \rho_{B_1})$, we have:

$$\ln \text{tr}(\rho_{A_0} \cdot \rho_{B_1}) = P_{E_A}(P'_{E_B X_L} + P'_{E_B X_L Z_L}) .$$

So that finally we have the Rényi relative entropy:

$$D = -\ln \frac{\text{tr}(\rho_{A_0} \cdot \rho_{B_1})}{\text{tr}(\rho_{A_0} \cdot \rho_{B_0})} = -\ln \frac{P_{E_A}(P'_{E_B X_L} + P'_{E_B X_L Z_L})}{P_{E_A}(P'_{E_B} + P'_{E_B Z_L})} = -\ln \frac{P'_{E_B X_L} + P'_{E_B X_L Z_L}}{P'_{E_B} + P'_{E_B Z_L}} . \quad (\text{S7})$$

Note that from the equality: $P'_{E_B} P'_{E_B X_L Z_L} = P'_{E_B X_L} P'_{E_B Z_L}$ the formula S7 can be reformulated as:

$$D = -\ln \frac{P'_{E_B X_L}}{P'_{E_B}} = -\ln \frac{\tilde{Z}_{\mathcal{I}}(J_z)}{\tilde{Z}_{\mathcal{I}}(J_z)} , \quad (\text{S8})$$

where $\mathcal{Z}_{\mathcal{I}}$ denote the partition function of the Ising model. The numerical result is shown in Fig. 7 of the main text.

B. Classical error injection for experimental bitstrings

To investigate the performance of the Rényi relative entropy under different noise levels, we implement a classical post-processing method to simulate increased noise. First, we characterize the intrinsic noise of the system by determining an effective uniform bit-flip error strength p_0 . This parameter is inferred from the measured purity $\text{tr}(\rho_0^2)$, using a relationship pre-calibrated through numerical simulations. Subsequently, we increase the noise level

by injecting additional errors into the measured bitstrings. This is achieved by applying a classical bit-flip channel, where each bit is flipped with a probability p_{flip} . The combination of the intrinsic system noise and the injected classical noise results in a boosted effective bit-flip error strength p , given by:

$$p = (1 - p_0) p_{\text{flip}} + (1 - p_{\text{flip}}) p_0 \quad (\text{S9})$$

By varying the flipping probability p_{flip} in the range $[0, 1/2]$, we can scan the effective bit-flip error strength p from its baseline value of p_0 up to the theoretical maximum of $1/2$. To ensure robust statistical estimates, this error injection procedure is repeated 100 times for each measured bitstring.

S3. TUNABLE TELEPORTATION PHASE TRANSITION

In the classical simulation, the core of the task is to sample the data bitstring σ , which is given by the Born's rule:

$$P(\sigma) = \langle \langle \sigma | \mathcal{N} | 0_L \rangle \rangle, \quad (\text{S10})$$

cast in the double Hilbert space for brevity. When $\mathcal{N} = (1 - p_X - p_Y - p_Z) + p_X X^{\otimes 2} + p_Y Y^{\otimes 2} + p_Z Z^{\otimes 2}$ becomes *diagonal* Pauli noise, the sampling is straightforward and can be done by an uncorrelated sampling. When $\mathcal{N} = 1 - p + p(\frac{X+Z}{\sqrt{2}})^{\otimes 2}$ involves the off-diagonal Pauli noise (sometimes referred to as coherent noise, albeit slightly different from the unitary coherent error), the sampling is less straightforward, but we have mature experience to use a double layer tensor network calculation for the Born sampling. Then we feed the data bitstring to the tensor network decoder to obtain the result.

A. Teleportation in X basis

Here we take $\theta = \pi/2$. The post-teleportation state that Bob receives (after local correction with Alice's outcomes) effectively suffers from the following noise:

$$\mathcal{N}_Z \circ \mathcal{M}_X \circ \mathcal{N}'_X (|0_L\rangle) , \quad (\text{S11})$$

where $\mathcal{N}'_X = (1 - p) + pX^{\otimes 2}$ denotes the bit-flip channel arising from the incoherent error p in the preparation of the surface code, $\mathcal{M}_X$ expresses the measurement channel induced by the non-maximally entangled teleportation, and $\mathcal{N}_Z$ denotes the dephasing channel owing to the measurement error of the Alice's qubits. We measure the qubits in the Z basis and feed the data bitstring to the classical computer to generate the decoded logical $\langle Z_L \rangle$. The last $\mathcal{N}_Z$ is absorbed by the measurement basis, and thus the noise action reduces to

$$\mathcal{N}_X \circ \mathcal{N}'_X = (\cos^2(t) + \sin^2(t)X^{\otimes 2}) \circ ((1 - p) + pX^{\otimes 2}) = \cos^2(t_{\text{eff}}) + \sin^2(t_{\text{eff}})X^{\otimes 2}, \quad (\text{S12})$$

where $\mathcal{N}_X = \cos^2(t) + \sin^2(t)X^{\otimes 2}$ is the bit-flip channel caused by the ‘‘coherent error’’ t . When only Z measurement is performed, m syndrome is extracted from the data qubits. In this case,

$$\sin^2(t_{\text{eff}}) = (1 - \sin^2(t))p + \sin^2(t)(1 - p) = \sin^2(t) + \cos(2t)p, \quad (\text{S13})$$

is the effective ‘‘coherent error’’ rate that creates the m syndrome. Thus we can predict how the logical error p_{X_L} depends on p and t . This effective reparametrization of coherent error is used in Fig. 3 in the main text.

B. Teleportation in a self-dual basis

The self-duality can maximally enhance the threshold of teleportation. Here we take $\theta = \pi/4$, which results in the following state:

$$\mathcal{N}_{\frac{X-Z}{\sqrt{2}}} \circ \mathcal{M}_{\frac{X+Z}{\sqrt{2}}} \circ \mathcal{N}'_X (|0_L\rangle) . \quad (\text{S14})$$

Here we take the same strategy for experimental measurement and decoding, to check the dependence of the infinite threshold on the noise probability.

If the measurement and local feedback is replaced by a unitary gate, and Alice is traced out afterwards, Bob's surface code can be approximated by the following channel:

$$\begin{aligned}
\mathcal{N} &= \mathcal{N}_{\frac{X+Z}{\sqrt{2}}} \circ \mathcal{N}'_X \\
&= \left(\cos^2(t) + \sin^2(t) \left(\frac{X+Z}{\sqrt{2}} \right)^{\otimes 2} \right) \circ ((1-p) + pX^{\otimes 2}) \\
&= \left(\frac{\sin^2(t)}{2} + p - \frac{3}{2}p\sin^2(t) \right) X^{\otimes 2} + \frac{p\sin^2(t)}{2} Y^{\otimes 2} + \frac{(1-p)\sin^2(t)}{2} Z^{\otimes 2} + \dots
\end{aligned} \tag{S15}$$

When the state is further dephased (due to projective measurement in Z basis) such that $Z^{\otimes 2} \rightarrow 1$ is applied on the data qubits, the off-diagonal terms that anti-commute with $Z^{\otimes 2}$ are forced to vanish, and the noise reduces to

$$\mathcal{N} \rightarrow \left(\frac{\sin^2(t)}{2} + p\cos^2(t) \right) X^{\otimes 2} + \dots, \tag{S16}$$

therefore the effective coherent error becomes:

$$\frac{\sin^2(t_{\text{eff}})}{2} = \frac{\sin^2(t)}{2} + p\cos^2(t) \Rightarrow \sin^2(t_{\text{eff}}) = \sin^2(t) + p(1 + \cos(2t)), \tag{S17}$$

which is used in Fig. 4 in the main text.

C. Learnability and teleportation transition

Here we compare the learnability transition with the teleportation transition from decoded experimental data. In addition, we also show the data for odd code distances $d = 3, 5, 7$ here, in comparison with the even code distances (mainly discussed in the main text). In the teleportation protocol, Alice also does the measurements and gets her bitstrings from which she can in some sense learn the logical state information of the original surface code. Comparing to the teleportable transition from Bob's side, by tuning the entangling strength t , Alice also suffers a learnable transition, which can be captured by the entropy of the mixed logical state conditioned upon her measurement outcome.

S4. SURFACE CODE STATE PREPARATION

A. Circuit for preparing Alice in a logical state

The quantum circuit shown in Fig. S5 is designed to encode a set of physical qubits into the surface code logical state $|0_L\rangle$. The process begins with the physical qubits initialized in the product state $|0 \dots 0\rangle$. This initial state is stabilized by the set of single-qubit Pauli operators $\{Z_0, Z_1, Z_2, Z_3, Z_4, Z_5, Z_6, Z_7, Z_8\}$. Consequently, it is already an eigenstate of the logical Z_L operator and all Z -type stabilizers (A_s) with an eigenvalue of $+1$. However, it is not an eigenstate of the X -type stabilizers (B_f). The purpose of the encoding circuit is to evolve this initial state into a simultaneous eigenstate of all stabilizers with an eigenvalue of $+1$, including the X -type ones. This is accomplished by first applying Hadamard gates to a set of representative qubits, one for each X -type stabilizer B_f . A Hadamard gate transforms the corresponding Z_j operators into X_j operators. For example, if qubits Q_0, Q_2, Q_5 , and Q_6 are chosen as representatives, the stabilizers become $\{X_0, Z_1, X_2, Z_3, Z_4, X_5, X_6, Z_7, Z_8\}$. Next, a sequence of CNOT gates extends these single-qubit X operators into the desired multi-qubit X -type stabilizers. This leverages the CNOT gate's operator propagation properties: a CNOT from a control qubit Q_c to a target qubit Q_t transforms an operator X_c into $X_c X_t$, and it also transforms an operator Z_t into $Z_c Z_t$. By carefully scheduling the CNOT gates in layers, the stabilizer generators are progressively constructed. The evolution of the generator set after each CNOT layer in Fig. S5 is as follows:

$$\{X_0 X_3, Z_1, X_2, Z_0 Z_3, Z_4, X_5 X_8, X_6, Z_7, Z_5 Z_8\}, \tag{S18}$$

$$\{X_0 X_1 X_3, Z_0 Z_1, X_2, Z_0 Z_3, Z_4 Z_5, X_4 X_5 X_8, X_6 X_7, Z_6 Z_7, Z_5 Z_8\}, \tag{S19}$$

$$\{X_0 X_1 X_3 X_4, Z_0 Z_1 Z_2, X_1 X_2, Z_0 Z_3, Z_3 Z_4 Z_5, X_4 X_5 X_7 X_8, X_6 X_7, Z_6 Z_7 Z_8, Z_5 Z_8\}. \tag{S20}$$

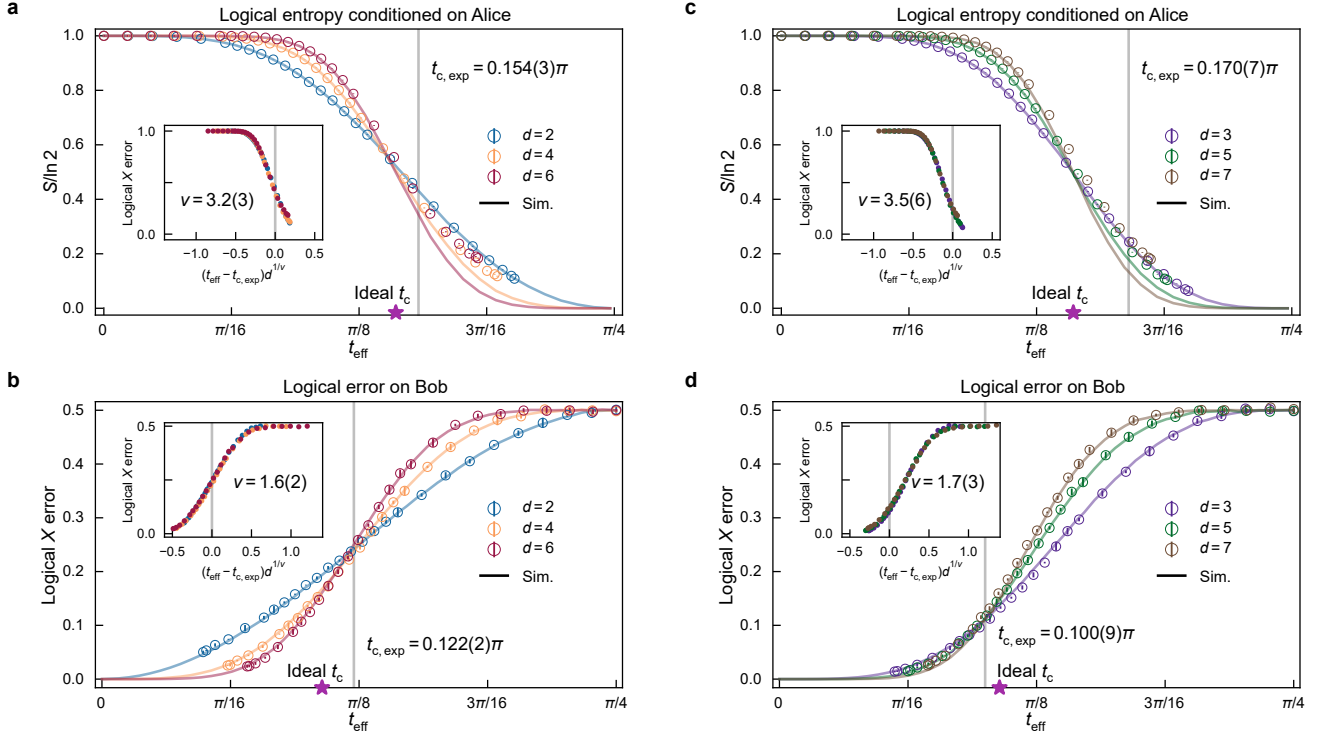


FIG. S4. **Alice's entropy vs Bob's logical error rate.** The transition point for Alice is 0.143π while for Bob is 0.107π , they are dual to each other such that $0.143\pi + 0.107\pi = 0.25\pi$. The mismatch between the Alice experiment data and numerical data is because entropy S is a nonlinear physical quantity and it is sensitive to the correlated errors in the experiment.

The final set of generators is equivalent to the stabilizer group defining the logical state $|0_L\rangle$ of the distance-3 surface code, which is

$$\{X_0X_1X_3X_4, X_1X_2, X_4X_5X_7X_8, X_6X_7, Z_0Z_3, Z_1Z_2Z_4Z_5, Z_3Z_4Z_6Z_7, Z_5Z_8, Z_6Z_7Z_8\}$$

The representative qubits and CNOT sequence are chosen to enable parallel gate implementation and minimize the total circuit depth, where the number of the two-qubit gate layers scales as $\lfloor (d+3)/2 \rfloor$ [4].

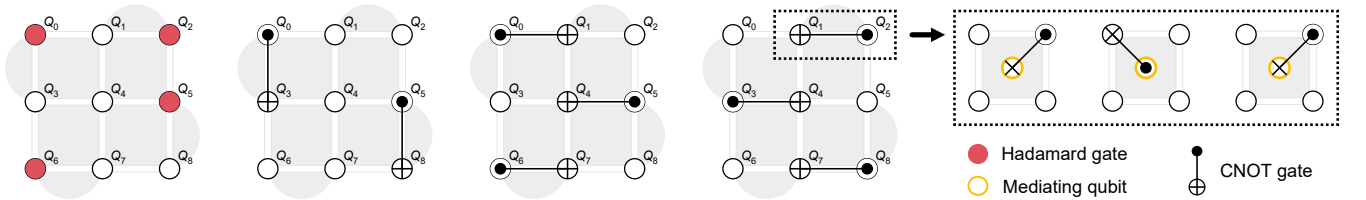


FIG. S5. **Circuit schematic for preparing a surface code in the logical state $|0_L\rangle$.** All physical qubits are initialized in the product state $|0 \cdots 0\rangle$. Next, Hadamard gates (red) are applied to representative qubits, which convert the corresponding single-qubit Pauli Z operators in the stabilizer group to single-qubit Pauli X operators. Subsequently, these operators are expanded into the required multi-qubit X -type and Z -type stabilizers through three layers of CNOT gates. The resulting state, in the absence of error, is the logical state $|0_L\rangle$. Because Alice's qubits are interleaved with Bob's, these CNOT gates are effectively implemented via mediating qubits (orange circles), as illustrated by the circuit in the dashed box.

However, because Alice's qubits are interleaved with Bob's qubits in order to implement teleportation between two surface codes, the stabilizers cannot be extended simply by applying two-qubit CNOT gates between adjacent qubits belonging solely to Alice. Instead, Bob's qubits must be used to mediate the propagation. Intuitively, this can be accomplished by a sequence of three layers of CNOT gates (Fig. S5). Consider three relevant stabilizers,

$$\{\cdots X_c I_m I_t \cdots, I_c Z_m I_t, \cdots I_c I_m Z_t \cdots\}, \quad (\text{S21})$$

which evolve under these three CNOT layers as

$$\{\cdots X_c X_m I_t \cdots, Z_c Z_m I_t, \cdots I_c I_m Z_t \cdots\}, \quad (\text{S22})$$

$$\{\cdots X_c X_m X_t \cdots, Z_c Z_m I_t, \cdots I_c Z_m Z_t \cdots\}, \quad (\text{S23})$$

$$\{\cdots X_c I_m X_t \cdots, I_c Z_m I_t, \cdots Z_c Z_m Z_t \cdots\}. \quad (\text{S24})$$

The final stabilizers can be rewritten as

$$\{\cdots X_c I_m X_t \cdots, I_c Z_m I_t, \cdots Z_c I_m Z_t \cdots\}. \quad (\text{S25})$$

Comparing the initial and final sets reveals two key findings. First, the mediating qubit Q_m is returned to its original state $|0\rangle$, which is identified by the persistence of the $I_c Z_m I_t$ stabilizer. Second, the overall effect is identical to that of a single CNOT gate with Q_c as the control and Q_t as the target. Consequently, the number of two-qubit gate layers still scales linearly as $3 \lfloor (d+3)/2 \rfloor$, which can be further reduced by reorganizing the CNOT gates to maximize their parallel implementation.

B. Circuit for preparing both Alice and Bob in a logical state

When evaluating the relative entropy, both Alice and Bob are prepared in a logical state. Similar to the procedure for encoding only Alice's qubits, all physical qubits are first initialized in the product state $|0 \cdots 0\rangle$. We then apply Hadamard gates to an appropriate subset of qubits so as to map a corresponding subset of stabilizers from the single-qubit Pauli Z operator to single-qubit Pauli X operator. The central step is then the systematic extension of these single-qubit Pauli X stabilizers to the nonlocal multi-qubit X -type stabilizers defining a surface code.

For concreteness, let us focus on the case of code distance of $d = 3$ (Fig. S6). We choose to propagate the support of the X -type stabilizers from left to right across the lattice. For larger distances, the number of required two-qubit gates can in principle be reduced by changing the direction and pattern of propagation, for instance by extending the stabilizers from the central region toward the boundaries. Here, using three layers of CNOT gates to effectively achieve a single CNOT gate between distant qubits in either Alice's or Bob's lattice proves infeasible. Instead, we introduce SWAP gates to dynamically transfer qubits between the two lattices.

The protocol proceeds as follows. First, two pairs of qubits—one from Alice's lattice and one from Bob's—are swapped. This brings the necessary qubits into proximity, allowing subsequent CNOT gates to extend the initial single-qubit X operators into two- and four-qubit X -type stabilizers. We note a circuit optimization here: the initial SWAP gates can be omitted by applying the initial Hadamard gates directly to the target qubits in their swapped positions. Subsequently, two swapped qubits are returned to their original lattices via additional SWAP operations. A new set of two qubit pairs is then swapped between Alice's and Bob's system to facilitate the next layers of CNOT gates. Finally, all qubits are restored to their respective lattices. At this stage, both Alice and Bob are prepared in the logical state $|0_L\rangle$.

C. Preparing logical states by measuring stabilizers

A more efficient state-preparation protocol leverages stabilizer measurements [5, 6], requiring only a constant-depth circuit. In this approach, the data qubits are first initialized in a specific product state, after which a single round of the surface-code cycle is applied to project the data qubits into the logical subspace. The resulting logical states may differ from those obtained via a unitary encoding circuit, as the stabilizers are randomly projected to either $+1$ (corresponding to the state $|0\rangle$) or -1 (corresponding to the state $|1\rangle$). Nevertheless, these states are logically equivalent up to a Pauli-frame update and therefore give the same logical expectation values in the absence of errors.

We implement this approach on our quantum processor, and then validate the generated states using the same method as in Fig. 5 of the main text, see Fig. S7. Here, we adopt the minimum-weight perfect matching (MWPM) decoder to generate logical outcomes. The experimental results demonstrate slightly lower logical error rates for all states, with the error asymmetry significantly alleviated. The fitted logical errors are $p_X = 0.022(1)$ and $p_Z = 0.017(1)$, compared to $p_X = 0.019(1)$ and $p_Z = 0.053(1)$ which are extracted from the experimental bitstrings of Fig. 5a in the main text.

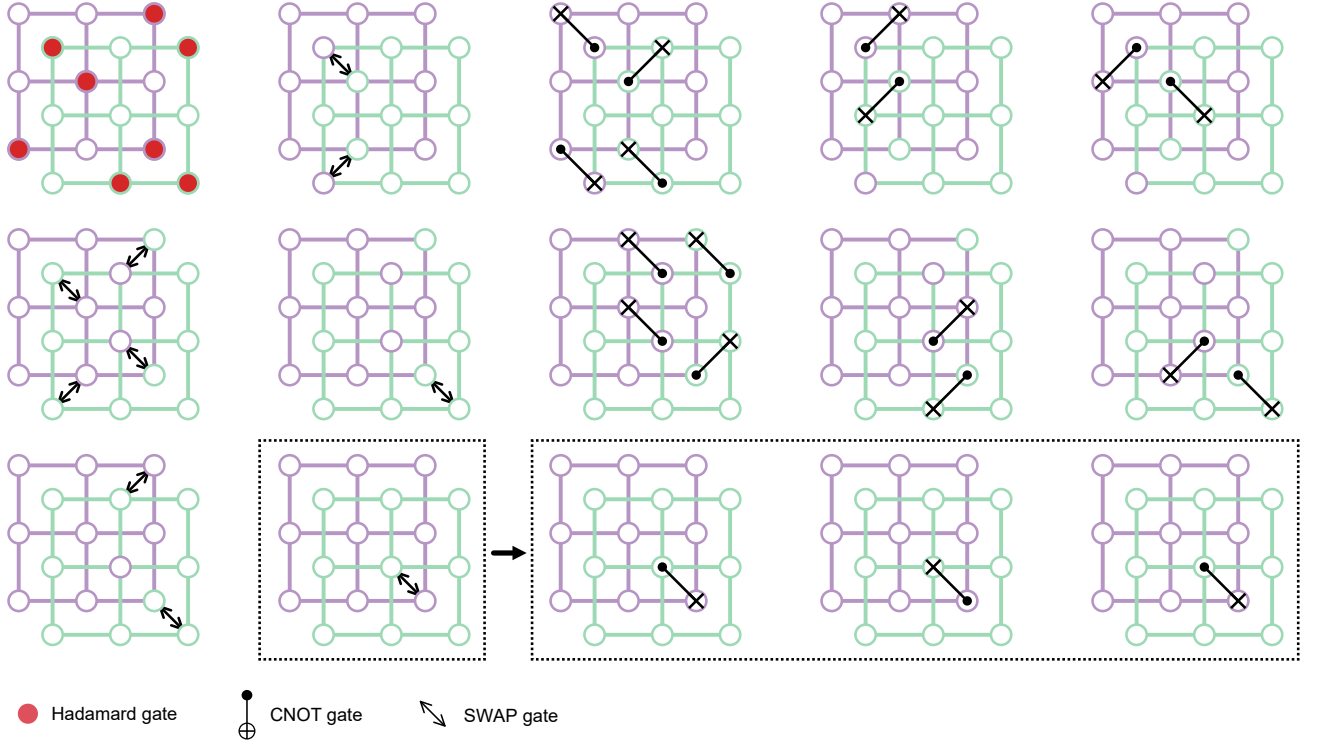


FIG. S6. **Circuit schematic for preparing two surface codes in the logical product state $|0_L\rangle \otimes |0_L\rangle$.** Starting from the product state $|0 \cdots 0\rangle$ for all physical qubits, Hadamard gates (red circles) are applied to the representative qubits selected from Alice’s lattice (purple) and Bob’s lattice (green). Subsequently, SWAP gates are employed to temporarily exchange qubits between Alice’s and Bob’s lattices, thereby bringing the relevant qubits into spatial proximity to enable the application of local CNOT gates. Through successive layers of these CNOT operations, single-qubit Pauli X or Z operators are progressively extended into the required multi-qubit X - or Z -type stabilizers. Following this, the exchanged qubits are returned to their original lattices via additional SWAP operations. The above SWAP-CNOT-SWAP sequence is repeated iteratively from left to right across the lattice, facilitating the propagation of entanglement throughout the system. This systematic process ultimately prepares Alice and Bob in the logical product state $|0_L\rangle \otimes |0_L\rangle$. A SWAP operation is implemented using three layers of CNOT gates, as illustrated in the circuit diagram within the dashed box.

S5. DECODING METHODS

Decoding the logical information from a noisy quantum code is generally an inference problem which allows various decoders, which may or may not yield the optimal threshold. While we have mainly showed the data from tensor network (TN)-based maximum likelihood (ML) decoder in the main text, here we also show the results from two alternative decoders: MWPM decoder and TN-based Bayesian decoder.

MWPM decoder is well-known to find the shortest path to pair up the syndromes which can be efficiently found by algorithm like Blossom algorithm, and here we use the PyMatching package [7]. It was shown [8] that it yields a threshold 0.103 as the zero temperature transition point of the random bond Ising model (RBIM), which is extremely close to the optimal threshold 0.109 as the finite temperature Nishimori transition point of the RBIM. The TN-based Bayesian decoder, on the other hand, outputs a probabilistic correction, that minimizes the “free energy” of the statistical model by taking into account the “entropy” as a key factor. Concretely, it exploits a mapping from the noisy quantum state to a classical statistical mechanics model. As illustrated in Fig. S8, the noisy quantum state subject to generic depolarizing errors is mapped to a random 8-vertex model [9]. In this representation, each vertex corresponds to a physical qubit and carries Ising variables $\tau^x = \pm 1$ and $\tau^z = \pm 1$ that are determined by the X -basis and Z -basis measurements. The Ising variable $s = \pm 1$ and $s' = \pm 1$ at the face center captures the stabilizer, whose flip can fluctuate an error string to its gauge equivalent counterpart that shares the same syndrome measurement or quantum number. Therefore $\tau^x s' s' = -1$ together describes an e -string and $\tau^z s s = -1$ describes an m -string. The

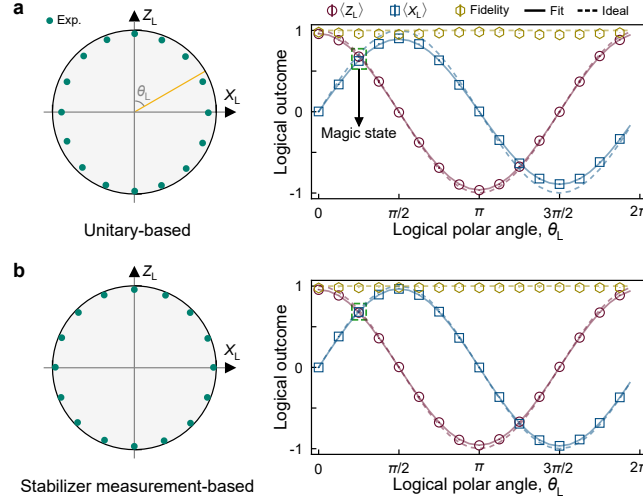


FIG. S7. **Comparison of two logical state preparation methods on Alice ($d = 3$) with the MWPM decoder.** **a**, Logical measurement results for the unitary-based state preparation method. The logical outcomes are derived from the same bitstrings used in Fig. 5a, but are decoded here using the MWPM decoder instead of the ML decoder. The reconstructed states projected onto the XZ -plane exhibit an elliptical distortion, reflecting an asymmetry in the logical error rates. The average state fidelity is 0.964(11). **b**, Logical measurement results for the stabilizer measurement-based state preparation method ($d = 3$). In contrast to the unitary method, the reconstructed states on the XZ -plane display a more circular distribution, reflecting a significant alleviation of error asymmetry. The average state fidelity is 0.981(2), demonstrating an improvement over the unitary-based approach.

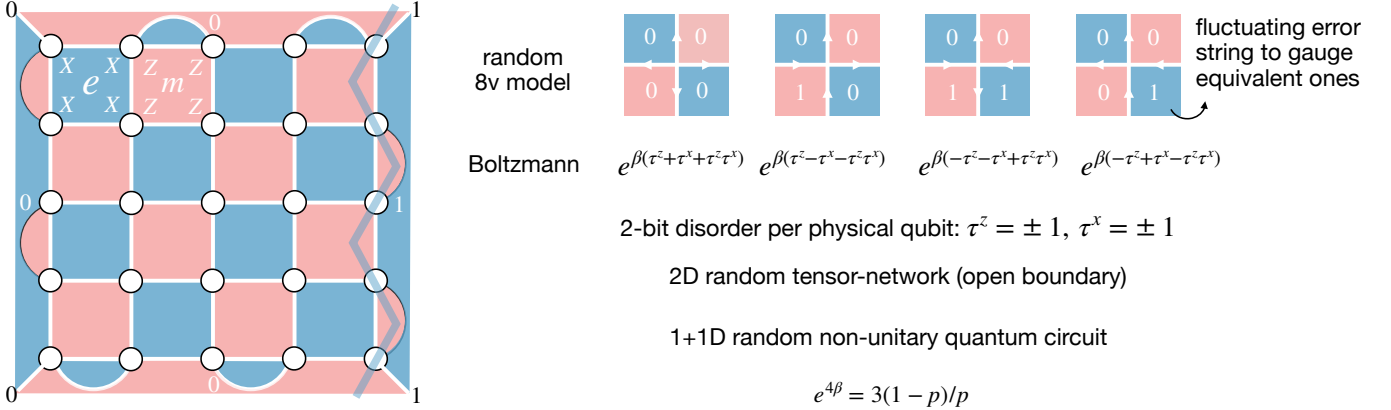


FIG. S8. **Random 8-vertex model describing the statistical ensemble of depolarizing error strings hitting the surface code**, taking a 2-step procedure: noisy quantum state $\rightarrow$ bilayer Ising model on faces $\rightarrow$ 8-vertex model. The two logical operator strings are equivalent to inserting the non-contractible domain walls for the blue and pink Ising layers, respectively. The non-contractible domain walls can be moved to the boundary, reducing to simply the boundary variable being 1 (otherwise we fix the boundary variable to be 0 since there is no stabilizer here that can fluctuate the logical error string). Thus the bilayer Ising model is placed in fixed boundary condition. Using the standard map from the bilayer Ising model to the vertex model, the face centered Ising variable is dual to the arrows on the bond, where each vertex expresses a weight given the state of the arrows. One can view the 8-vertex network as a 2D tensor network or a 1+1D quantum circuit, where the arrow direction is viewed as a qubit. The boundary tensor needs some care since the Ising variable is fixed. When serving as a decoder for the experiment, the experimental data bitstring can be input as τ (up to a gauge transformation) that creates the syndromes. The depolarizing error here can be easily generalized to anisotropic but still diagonal Pauli errors with independent p_X, p_Y, p_Z .

partition function [9] characterizing gauge equivalent error chains under depolarizing noise is

$$\mathcal{Z} = \sum_{ss'} e^{\beta(\tau^x s'_k s'_l + \tau^z s_i s_j + \tau^x \tau^z s_i s_j s'_k s'_l)} , \quad (\text{S26})$$

where $e^{4\beta} = \frac{3(1-p)}{p}$ with p representing the individual Pauli X , Y or Z noise. In the main text, we consider the simple version with only bit-flip noise, in this case, the random 8-vertex model reduce to RBIM [8]:

$$\mathcal{Z}_{\text{RBIM}} = \sum_{\langle i,j \rangle} e^{\beta' \tau^z s_i s_j}, \quad (\text{S27})$$

with $e^{2\beta'} = \frac{1-p_X}{p_X}$, p_X represents the bit-flip probability. In practice, a quantum processor gains the data bitstring, which is fed into the tensor network decoder shown in Fig. S8 as τ variable, the decoder computing the corresponding partition function and yields the decoded logical information.

In Fig. S9, we present the decoding results obtained using the MWPM decoder and TN-based Bayesian decoder, applied to the measured bitstrings shown in Fig. 3 of the main text (even code distances). For the MWPM decoder, the extracted thresholds are $0.116(2)\pi$ at $\theta = \pi/2$ and $0.164(2)\pi$ at $\theta = \pi/4$; these values appear slightly overestimated, which is also observed in the ML decoder. In contrast, the TN-based Bayesian decoder yields thresholds of $0.107(1)\pi$ at $\theta = \pi/2$ and $0.152(2)\pi$ at $\theta = \pi/4$, which agree very closely with the asymptotic theoretical thresholds.

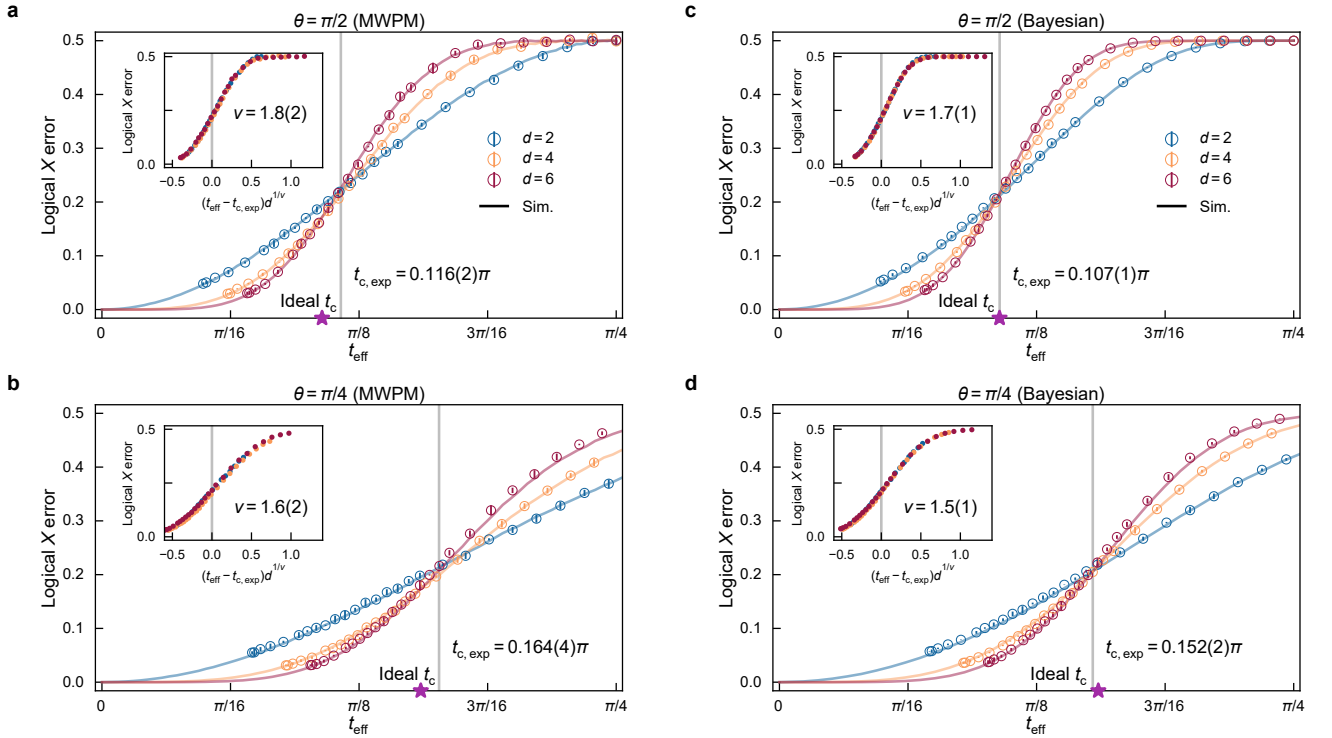


FIG. S9. **Results for the teleportation transition using MWPM and Bayesian decoders.** For each panel, the critical threshold is extracted from a finite-size scaling analysis shown in the inset, and is indicated by the vertical grey solid line. The stars show the asymptotic thresholds of $t_c(\theta = \pi/2) = 0.1072128(2)\pi$ and $t_c(\theta = \pi/4) = 0.1548014(3)\pi$.

S6. INTEGRATION OF QUANTUM ERROR CORRECTION WITH THE LOGICAL TELEPORTATION PROTOCOLS

Outlooking the fault tolerant distributed quantum computation, we should be able to integrate quantum error correction with our logical teleportation protocols. The typical error correction scheme relies on the entanglement between the data qubits and the syndrome qubits, which are then projectively measured to learn about the noise. The teleportation relies on the entanglement between Alice's data qubits and Bob's data qubits, where Alice's qubits are also projectively measured to enable the teleportation. In an ideal situation for *physically remote* teleportation, we need four groups of qubits: (Alice's data qubits, Alice's syndrome qubits, Bob's data qubits, and Bob's syndrome qubits), roughly in need of $4d^2$ qubits, where d is the code distance.

Here we can design a proof of principle compact simulation setup to demonstrate the integration of quantum error correction with the logical teleportation protocols, with only $2d^2$ qubits based on the existing qubit connectivity layout, but in need of fast and reliable mid-circuit measurements and qubit reset techniques. The schematic is shown in Fig. S10. First we can use Bob's qubits as syndrome qubits for Alice, one cycle of which was already tested in our appendix Section S4C *Preparing logical states by measuring stabilizers*. One can collect the measurement record and perform error correction as feedback or in post-processing manner. After several rounds of syndrome measurements, we can teleport the surface code state from Alice to Bob. Then we reset and reuse Alice's qubits as syndrome qubits for Bob.

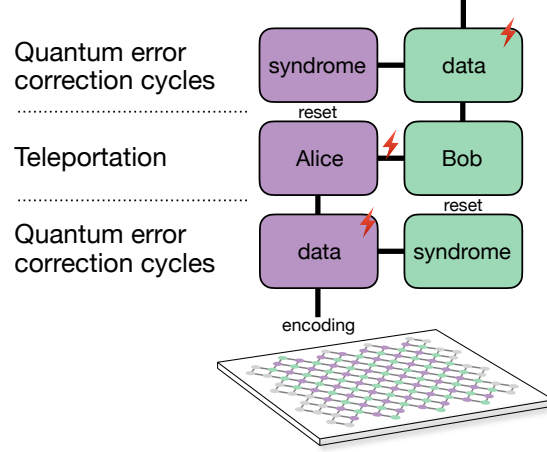


FIG. S10. **Schematic illustration of the integration of quantum error correction with the logical teleportation protocols** on the same qubit layout, with mid-circuit measurements and qubit reset.

-
- [1] S. Bravyi, S. Sheldon, A. Kandala, D. C. McKay, and J. M. Gambetta, Mitigating measurement errors in multiqubit experiments, *Phys. Rev. A* **103**, 042605 (2021).
 - [2] C. Gidney, Stim: a fast stabilizer circuit simulator, *Quantum* **5**, 497 (2021).
 - [3] R. Fan, Y. Bao, E. Altman, and A. Vishwanath, Diagnostics of mixed-state topological order and breakdown of quantum memory, *PRX Quantum* **5**, 020343 (2024).
 - [4] K. J. Satzinger *et al.*, Realizing topologically ordered states on a quantum processor, *Science* **374**, 1237 (2021).
 - [5] Y. Li, A magic state's fidelity can be superior to the operations that created it, *New J. Phys.* **17**, 23037 (2015).
 - [6] Y. Ye *et al.*, Logical magic state preparation with fidelity beyond the distillation threshold on a superconducting quantum processor, *Phys. Rev. Lett.* **131**, 210603 (2023).
 - [7] O. Higgott and C. Gidney, Sparse Blossom: correcting a million errors per core second with minimum-weight matching, *Quantum* **9**, 1600 (2025).
 - [8] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, Topological quantum memory, *J. Math. Phys.* **43**, 4452 (2002).
 - [9] H. Bombin, R. S. Andrist, M. Ohzeki, H. G. Katzgraber, and M. A. Martin-Delgado, Strong resilience of topological codes to depolarization, *Phys. Rev. X* **2**, 021004 (2012).