Experimental Demonstration of High-Fidelity Logical Magic States from Code Switching

Lucas Daguerre,[1] Robin Blume-Kohout,[2] Natalie C. Brown,[3] David Hayes,[3] and Isaac H. Kim[4]
[1]*Department of Physics and Astronomy, University of California, Davis, California 95616, USA*
[2]*Quantum Performance Laboratory, Sandia National Laboratories, Albuquerque, New Mexico 87185, USA*
[3]*Quantinuum, 303 South Technology Court, Broomfield, Colorado 80021, USA*
[4]*Department of Computer Science, University of California, Davis, California 95616, USA*

(Received 12 August 2025; accepted 18 September 2025; published 14 October 2025)

Preparation of high-fidelity logical magic states has remained as a necessary but daunting step towards building a large-scale fault-tolerant quantum computer. One approach is to fault-tolerantly prepare a magic state in one code and then switch to another, a method known as code switching. We experimentally demonstrate this protocol on an ion-trap quantum processor, yielding a logical magic state encoded in an error-correcting code with state-of-the-art logical fidelity. Our experiment is based on the first demonstration of code switching between color codes, from the fifteen-qubit quantum Reed-Muller code to the seven-qubit Steane code. We prepare an encoded magic state in the Steane code with 82.58% probability, with an infidelity of at most $5.1(2.7) \times 10^{-4}$. The reported infidelity is lower than the leading infidelity of the physical operations utilized in the protocol by a factor of at least 2.7, indicating the quantum processor is below the pseudothreshold. Furthermore, we create two copies of the magic state in the same quantum processor and perform a logical Bell basis measurement for a sample-efficient certification of the encoded magic state. The high-fidelity magic state can be combined with the already-demonstrated fault-tolerant Clifford gates, state preparation, and measurement of the 2D color code, completing a universal set of fault-tolerant computational primitives with logical error rates equal or better than the physical two-qubit error rate.

DOI: 10.1103/dck4-x9c2 Subject Areas: Quantum Physics, Quantum Information

I. INTRODUCTION

Since the discovery of Shor's factoring algorithm [1], it has been well known that quantum computers can solve certain computational problems exponentially faster than the best known classical algorithms [1,2]. However, the applications for which a quantum computer is expected to provide a substantial advantage tend to require a large number of physical operations [3]. Existing quantum computers are noisy, and taming the noise to the level low enough to run these algorithms remains as an outstanding open challenge.

Theoretically, the threshold theorem asserts that a noisy quantum computer can simulate a noiseless quantum computer with a moderate overhead, if the physical error rate is sufficiently low [4]. Bridging the gap between the requirement of the threshold theorem and what is experimentally feasible has been one of the central efforts in quantum computing that has spanned multiple decades of work. A large body of recent literature demonstrated different aspects of fault-tolerant quantum computation, e.g., logical state preparation, memory, and gates [5–12]. In particular, recently Ref. [13] demonstrated that the fidelity of logical controlled-NOT gates and state preparation are equal or better than their physical counterparts in a trapped ion quantum processor. However there was one logical operation whose fidelity was worse than that of the involved physical operations: T gate, or equivalently, the preparation of a logical magic state [14]. This is a necessary ingredient for building a universal set of fault-tolerant gates.

Historically, preparation of a high-fidelity magic state has been one of the most costly and challenging parts of fault-tolerant quantum computation. A common approach has been to inject a noisy magic state into a code, which is then distilled [14]. Injection of a magic state is not fault tolerant, so to reduce the error rate to a level comparable to the other logical operations, it is necessary to carry out magic state distillation [14], which can lead to a considerable amount of extra complexity and overhead; see Ref. [15] and the references therein.

An alternative approach is to use code switching [16–18]. In this approach, the magic state is prepared in a code that admits a transversal logical non-Clifford gate,

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

TABLE I. Comparison of experimental magic state fidelities for quantum codes with distance $d \leq 3$, including acceptance rare (AR) and experimental platform. The magic states correspond to $|\bar{T}\rangle = (1/\sqrt{2})(|\bar{0}\rangle + e^{i\pi/4}|\bar{1}\rangle)$, $|\bar{H}\rangle = \cos(\pi/8)|\bar{0}\rangle + \sin(\pi/8)|\bar{1}\rangle$, $|\bar{M}\rangle = \cos(\theta/2)|\bar{0}\rangle + e^{i\pi/4}\cos(\theta/2)|\bar{1}\rangle$ with $\theta = \cos^{-1}(1/\sqrt{3})$, and $|\overline{CZ}\rangle = (1/\sqrt{3})(|\overline{00}\rangle + |\overline{10}\rangle + |\overline{01}\rangle)$. Our reported value 0.99949(27) is a lower bound for the fidelity (see Sec. IV). Results marked with (PS) indicate that postselection is further utilized in the logical state tomography measurements. Using postselected logical state tomography measurements, our reported fidelity lower bound is 0.99985(11) with 81.26% of acceptance rate.

Year	Reference	State	Fidelity	AR (%)	Codes	Platform
2025	This work	$\|\bar{T}\rangle$	0.99949^{+27}_{-27}	82.58	Steane, $[[15,1,3]]$	Trapped ion
2024	(PS)Lacroix *et al.* [31]	$\|\bar{T}\rangle$	0.9992^{+3}_{-15}	75.2	Steane	Superconducting qubits
		$\|\bar{H}\rangle$	0.9959^{+38}_{-37}	74.6		
	Rodriguez *et al.* [30]	$\|\bar{M}\rangle$	0.994^{+3}_{-4}	1	Steane, $[[5,1,3]]$	Neutral atoms
	(PS)Pogolerov *et al.* [19]	$\|\bar{T}\rangle$	0.963^{+4}_{-4}	19	Steane, $[[10,1,2]]$	Trapped ion
	Kim *et al.* [29]	$\|\bar{H}\rangle$	0.8806^{+2}_{-2}	36.3	Surface code	Superconducting qubits
2023	Ye *et al.* [28]	$\|\bar{T}\rangle$	0.8771^{+9}_{-9}	73.41	Surface code	Superconducting qubits
		$\|\bar{H}\rangle$	0.9090^{+9}_{-9}	73.41		
	Gupta *et al.* [27]	$\|\overline{CZ}\rangle$	0.9877^{+11}_{-11}	17	Surface code	Superconducting qubits
2021	Anderson *et al.* [5]	$\|\bar{T}\rangle$	0.978^{+6}_{-6}	100	Steane	Trapped ion
	Postler *et al.* [7]	$\|\bar{H}\rangle$	0.994^{+5}_{-14}	13.7	Steane	Trapped ion
	Marques *et al.* [26]	$\|\bar{T}\rangle$	0.966	<25	Surface code	Superconducting qubits
2020	Egan *et al.* [25]	$\|\bar{H}\rangle$	0.972^{+12}_{-12}	100	Bacon-Shor	Trapped ion

such as the T gate. Then the code is switched to another code that admits a complementary set of fault-tolerant logical gates; see Ref. [19] for a recent experiment. Code switching had been deemed less practical than magic state distillation in the past [18], but a more recent work [20] suggests that it can be improved further than previously thought possible.

In this paper, we experimentally demonstrate that code switching can be used to prepare a high-fidelity magic state encoded in a quantum error-correcting code, using a modest amount of resources. To that end, we present the first experimental demonstration of magic state preparation that employs a code switching between color codes [21,22], from the $[[15,1,3]]$ quantum Reed-Muller code [23] to the $[[7,1,3]]$ Steane code [24]. We report the state-of-the-art logical infidelity of at most $5.1(2.7) \times 10^{-4}$, using error-corrected logical quantum state tomography. Indeed, this infidelity is lower than all the reported results in the literature [5,7,19,25–31], even the ones whose infidelity is estimated using postselected logical quantum state tomography [31], while only using 28 qubits in total. In a similar line of research, another experiment done by some of the authors, which uses an error-detecting code and postselected logical quantum state tomography, will appear elsewhere [32].

Aside from achieving the state-of-the-art infidelity for the logical magic state encoded in a quantum error-correcting code, this experiment also shows that the error rates of the existing quantum processor is below the pseudothreshold of our protocol. The dominant source of errors are the two-qubit error $1.05(8) \times 10^{-3}$ and the state preparation and measurement (SPAM) error $1.38(12) \times 10^{-3}$, which are higher than the infidelity of the reported magic state by a factor of 2 and 2.7, respectively. Thus, continued improvement in physical layer fidelities should improve the infidelity of the magic state at an even faster rate.

Because of the high fidelity of our magic state, the number of samples needed to certify its fidelity becomes rather large. In order to reduce this complexity, we perform a Bell basis measurement over two copies of the encoded magic state. It was recently shown that such a measurement can lead to an asymptotic improvement in the number of samples needed, assuming the two copies are identical [33]. We use a simple variant of this method with a more relaxed set of assumptions, yielding a rigorous upper bound on the infidelity. A comparison of our results with other experiments in the literature can be found in Table I.

II. MAGIC STATE PREPARATION PROTOCOL

We prepare a magic state $|\bar{T}\rangle = (1/\sqrt{2})(|\bar{0}\rangle + e^{i\pi/4}|\bar{1}\rangle)$ using the code-switching protocol of Ref. [20]. This protocol utilizes two different quantum error-correcting codes: the Steane code $[[7,1,3]]$ [24] and the quantum Reed-Muller $[[15,1,3]]$ (qRM) code [23]. These codes require $n = 7$ and $n = 15$ physical qubits, respectively, and encode a single logical qubit $k = 1$. They are also distance $d = 3$ instances of the two-dimensional (2D) [21] and three-dimensional (3D) [22] color code families,

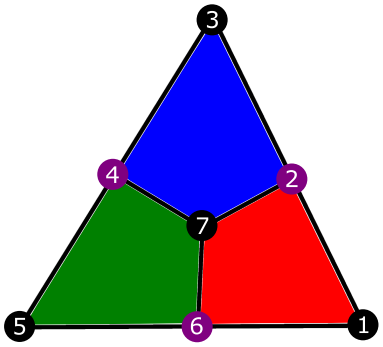


FIG. 1. Schematic representation of the [[7, 1, 3]] Steane code in a 2D lattice including labels for the physical qubits.

respectively. Indeed, they are capable of correcting arbitrary single-qubit errors and detecting arbitrary two-qubit errors. While the Steane code supports the full Clifford group transversally, it does not admit any transversal non-Clifford gate [34,35]. In contrast, the qRM code supports a transversal non-Clifford $\bar{T} = \mathrm{diag}(1, e^{i\pi/4})$ gate [22]. Because of this, the code-switching protocol of Ref. [20], described at the logical level in Fig. 3, allows the preparation of a magic state $|\bar{T}\rangle = \bar{T}|\overline{+}\rangle$ in the Steane code. Hence, the magic state can be used to implement a logical $\bar{T}$ gate to an arbitrary logical state encoded in the Steane code via gate teleportation [36], effectively circumventing the restrictions imposed by [34,35]. For completeness, we provide additional information about the color codes in Appendix C, as well as a geometrical realization of their embedding in 2D and 3D lattices in Figs. 1 and 2, respectively.

The protocol of Fig. 3 comprises of several different logical components. Regarding state initialization, a Steane code block is prepared in the logical $|\bar{0}\rangle$ state, and a qRM code block is prepared in the logical $|\overline{+}\rangle$ state. Their initialization is done using verification flag-based gadgets

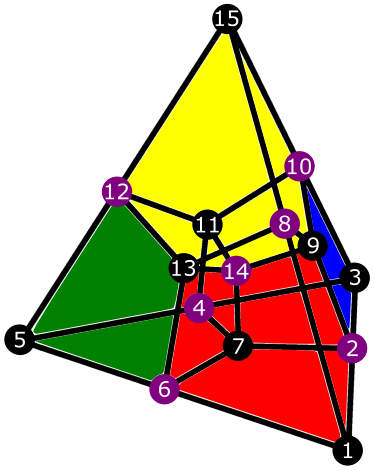


FIG. 2. Schematic representation of the [[15, 1, 3]] qRM code in a 3D tetrahedral lattice including labels for the physical qubits.

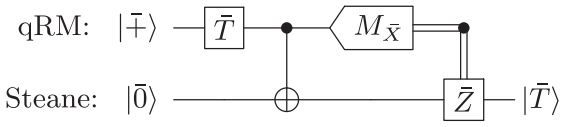


FIG. 3. The magic state preparation protocol at the logical level [20]. A qRM (quantum Reed-Muller) and a Steane code block are prepared in $|\overline{+}\rangle$ and $|\bar{0}\rangle$, respectively. The transversal logical $\bar{T}$ gate is applied, creating a logical magic state $|\bar{T}\rangle = \bar{T}|\overline{+}\rangle$ in the qRM code. The magic state is then teleported to the Steane code block by applying a logical $\overline{\mathrm{CNOT}}$ gate, measuring the logical $\bar{X}$ operator and applying a logical $\bar{Z}$ correction if $\bar{X} = -1$. The output of the protocol is a logical magic state $|\bar{T}\rangle$ in the Steane code block.

[20,37–39], described in Appendix C. Although this process is often viewed as postselection, we instead refer to this as a *preselection* [12] gadget. Because the magic state preparation protocol can be repeated until success, without affecting the other parts of the computation.

After initializing the two code blocks, a magic state is prepared in the qRM code block using the native transversal logical $\bar{T}$ gate given by

$$\bar{T} = T_1 T_2^\dagger T_3 T_4^\dagger T_5 T_6^\dagger T_7 T_8^\dagger T_9 T_{10}^\dagger T_{11} T_{12}^\dagger T_{13} T_{14}^\dagger T_{15}. \quad (1)$$

Once the magic state is prepared in the qRM code block, its logical information is transferred to the Steane code block using a logical one-bit teleportation gadget [36]. Correspondingly, an entangling logical $\overline{\mathrm{CNOT}}$ gate between both code blocks is applied [40,41]

$$\overline{\mathrm{CNOT}} = \bigotimes_{i=1}^{7} \mathrm{CNOT}_i, \quad (2)$$

which consist of physical CNOT gates between qubits at the bottom face of the qRM code block acting as controls, and the Steane code block acting as targets, see Fig. 4. Finally, the logical $\bar{X}$ operator is destructively measured in the qRM code block, where further pre-selection is employed (see Appendix B). In cases where the measurement outcome $\bar{X} = -1$, a logical $\bar{Z}$ correction has to be applied, which in practice it can be taken into account in software postprocessing without being physically applied (see Appendix B).

III. LOGICAL MAGIC STATE CERTIFICATION

In order to quantify the quality of the noisy magic state $\bar{\rho}$ produced in the experiments, we perform quantum state tomography at the logical level to estimate the fidelity $F(\bar{\rho}, |\bar{T}\rangle\langle\bar{T}|)$. The fidelity can be simplified to the following expression [see (D3) in Appendix D]:

$$F(\bar{\rho}, |\bar{T}\rangle\langle\bar{T}|) = \frac{1}{2} + \frac{1}{2\sqrt{2}}(\langle\bar{X}\rangle_{\bar{\rho}} + \langle\bar{Y}\rangle_{\bar{\rho}}), \quad (3)$$

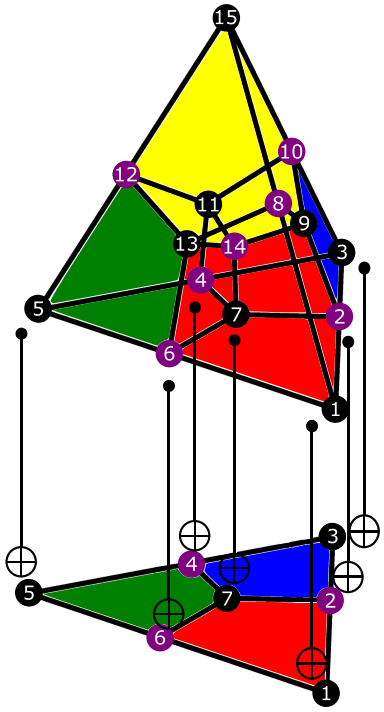


FIG. 4. Logical $\overline{\text{CNOT}}$ gate between the Steane and qRM code blocks, consisting of physical CNOT gates acting on the base of the tetrahedron as controls and on the Steane code block as targets.

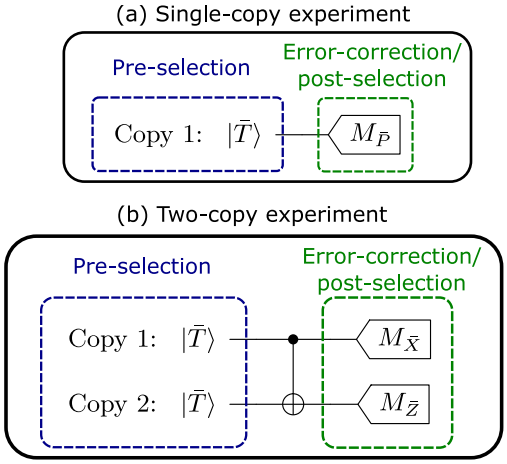


FIG. 5. Single-copy (a) and two-copy experiments (b), respectively. In both experiments, the magic states are prepared using preselection as in Fig. 3. In (a) the logical $\bar{P}$ operators with $\bar{P} \in \{\bar{X}, \bar{Y}, \bar{Z}\}$ are measured, whereas in (b) the logical $\bar{X}_1 \otimes \bar{Z}_2$ operator is measured. The overlap with the singlet ϵ (4) is given by the fraction of measurement outcomes such that $(m_{\bar{X}}, m_{\bar{Z}}) = (-1, -1)$. In both cases, the logical operators can be determined either with an ideal round of error correction or postselection (error detection).

where $\langle \bar{P} \rangle_{\bar{\rho}} = \text{Tr}(\bar{\rho}\,\bar{P})$ is the expected value of $\bar{P} \in \{\bar{X}, \bar{Y}\}$. We refer to the method of determining $\langle \bar{P} \rangle_{\bar{\rho}} = \text{Tr}(\bar{\rho}\,\bar{P})$ for $\bar{P} \in \{\bar{X}, \bar{Y}, \bar{Z}\}$ as the *single-copy experiment*, see Fig. 5(a). These expected values can be estimated by applying physical measurements in the X, Y, or Z eigenbasis and applying error correction or postselection (error detection) to the measurement outcomes. For more details, see Appendix B.

A drawback of determining the fidelity solely from (3) is that neither $\bar{X}$ nor $\bar{Y}$ possess the target magic state as an eigenstate. Thus, their variance is of order unity. If we obtain N_{shots} independent shots to determine the expected values, the fidelity can be bounded only up to $F \geq 1 - O(1/\sqrt{N_{\text{shots}}})$. This is unpractical if the infidelity $1 - F$ is small, which is precisely the case at hand.

In order to lower the sample complexity, we employ an alternative method that can achieve $F \geq 1 - O(1/N_{\text{shots}})$. This involves performing an additional experiment using two copies of the state, which we refer to as the *two-copy experiment*. The main idea is to perform a simplified version of the SWAP test [42] to estimate the purity of the state; see Fig. 5(b) for the circuit diagram. This process lets us determine the purity of an experimental state by quantifying the overlap with the singlet (antisymmetric) state $|\bar{\Phi}_1\rangle = (1/\sqrt{2})(|\overline{01}\rangle - |\overline{10}\rangle)$,

$$\epsilon \equiv \langle \bar{\Phi}_1 | \bar{\rho}_1 \otimes \bar{\rho}_2 | \bar{\Phi}_1 \rangle = \langle \overline{-1} | \overline{\text{CNOT}_s} (\bar{\rho}_1 \otimes \bar{\rho}_2) \overline{\text{CNOT}_s} | \overline{-1} \rangle, \tag{4}$$

where $|\overline{-1}\rangle \equiv |\bar{-}\rangle_1 |\bar{1}\rangle_2$. Here the logical $\overline{\text{CNOT}_s}$ gate is implemented between two copies of the Steane code, unlike the logical $\overline{\text{CNOT}}$ (2) gate; see Appendix C 1 for more details. For the experimental details of the two-copy experiment, see Appendix B.

Combining the results from one- and two-copy experiments, we can obtain a rigorous lower bound on the fidelity:

$$F(\bar{\rho}, \bar{\sigma}_T) \geq 1 - \left(\epsilon + \frac{\delta_X^2 + \delta_Y^2 + \delta_Z^2}{4} \right). \tag{5}$$

The values $\vec{\delta} = (\delta_X, \delta_Y, \delta_Z)$ correspond to the difference of the vectors associated to the Bloch sphere representation of the experimental state $\bar{\rho} = (1/2)(I + \vec{v} \cdot \vec{\sigma})$ and the target logical pure state $\bar{\sigma}_T = (1/2)(I + \vec{u} \cdot \vec{\sigma})$,

$$\vec{v} = \vec{u} + \vec{\delta}, \tag{6}$$

where $\vec{\sigma}$ is the vectorial representation of the Pauli matrices. Note that $\vec{v} = (\langle \bar{X} \rangle_{\bar{\rho}}, \langle \bar{Y} \rangle_{\bar{\rho}}, \langle \bar{Z} \rangle_{\bar{\rho}})$ and for magic $|\bar{T}\rangle\langle\bar{T}|$ states $\vec{u} = \{(1/\sqrt{2}), (1/\sqrt{2}), 0\}$; see Appendix D for a derivation. The formula (5) holds regardless of the logical pure target state $\bar{\sigma}_T$, and achieves the desired $F \geq 1 - O(1/N_{\text{shots}})$ scaling.

We remark that a recent work also proposed a method with the same scaling in sample complexity [33]. Our approach differs from that work in two aspects. First, unlike Ref. [33], our result can be extended to the setup in which the states are not necessarily identical, only assuming they are in a product state. In this case, Eq. (5) applies to at least one of the two copies. For the worst-case fidelity between the two copies, we obtain a similar bound, albeit with a worse constant appearing in front of ϵ. Second, Ref. [33] proposes to twirl both copies of the magic states before applying the entangling $\overline{\mathrm{CNOT}}_s$ gate. We forgo the twirling operation and instead perform the single-copy experiments. For future work, it will be interesting to compare the estimated fidelity using these two different approaches.

IV. EXPERIMENTAL RESULTS

The logical state certification experiments were carried out on the Quantinuum H2-1 trapped-ion quantum processor [43,44] using 28 qubits total (22 data and 6 ancilla qubits) for the single-copy experiments, and 56 qubits total (44 data and 12 ancilla qubits) for the two-copy experiments. The previous numbers assume midcircuit measurement and qubit reuse since the single-copy experiments require 31 qubits (22 data and 9 ancilla qubits), whereas the two-copy experiments require 62 qubits (44 data and 18 ancilla qubits). A virtue of the trapped-ion quantum processor is its all-to-all connectivity, capable of implementing requisite non-local gates with little additional error. Technical information about the Quantinuum H2-1 quantum processor can be found in Appendix A.

The fidelity lower bound for the magic state $|\bar{T}\rangle$ is computed using (5). The result gives for an error-corrected magic state an infidelity $1-F \leq 5.1(2.7)\times 10^{-4}$. With a total of $N_{\mathrm{shots}} = 10000$ shots for each Pauli basis, the average acceptance rate across al single-copy experiments is 82.58%. The dominant contribution to the fidelity bound comes from the two-copy experiments since $\epsilon = 4.56(2.63)\times 10^{-4}$. We also perform $N_{\mathrm{shots}} = 10000$ shots of the two-copy experiments. The acceptance rate is 65.73% and 3 logical $|\overline{-1}\rangle$ states are detected in the experiments. The magnitude of ϵ, which quantifies the purity of the state, is 1 order of magnitude larger than the single-copy experiments contribution to (5), which is $\frac{1}{4}(\delta_X^2+\delta_Y^2+\delta_Z^2) = 5.24(5.60)\times 10^{-5}$. A summary of the experimental results can be also found in Table II in Appendix B.

If we employ postselection (error detection) in the decoding step, the infidelity bound becomes $1-F \leq 1.5(1.1)\times 10^{-4}$, with an acceptance rate of 81.26%. The two-copy experiment yields $\epsilon = 9.28(9.28)\times 10^{-5}$ with an acceptance rate of 61.82% (no logical $|\overline{-1}\rangle$ states are detected). The contribution to the fidelity bound due to ϵ is still dominant, because it is almost twice as large as the single-copy experiment contribution $\frac{1}{4}(\delta_X^2+\delta_Y^2+\delta_Z^2) = 5.76(5.93)\times 10^{-5}$. Thus, the infidelity may be estimated to be even lower if we take more samples. We leave such studies for future work.

The reported infidelity of $5.1(2.7)\times 10^{-4}$ is below the leading error rate of the involved physical operations. For comparison, the two-qubit error rate is $p_2 = 1.05(8)\times 10^{-3}$ and the SPAM error rate $1.38(12)\times 10^{-3}$. Thus these error rates are below the pseudothreshold of our scheme. Moreover, preparation of the encoded magic state and its characterization has an infidelity expected to be lower than the infidelity of the same process at the physical level. [Such a process is essentially limited by the SPAM error of $1.38(12)\times 10^{-3}$.] Lastly, the reported infidelity is the new state of the art, even compared against the other methods in the literature that use postselection; see Table I. Thus our experiment serves as a new frontier for preparing high-fidelity magic states in experiments, an essential ingredient of fault-tolerant quantum computation.

V. DISCUSSION

In this paper, we experimentally demonstrate one of the key building blocks of fault-tolerant quantum computation: preparation of high-fidelity encoded magic states. We use code switching from the 15-qubit qRM code to the 7-qubit Steane code [16–18], employing a verified preselection protocol [20]. The reported infidelity of the logical magic state is the new state-of-the-art result for magic states encoded in a quantum error-correcting code. Moreover, this infidelity is lower than the leading physical error rate (dominated by SPAM) by a factor of 2.7, indicating that the quantum processor is below the pseudothreshold of the protocol.

Combined with the fault-tolerant Clifford gates, state preparation and measurement using the same code [13], we complete a universal set of logical operations whose logical error rates are all equal or better than the physical two-qubit error rate.

It is notable that the state-of-the-art fidelity was achieved using a rather modest number of qubits (22 data and 6 ancilla qubits). The modest resource requirement suggests that the realistic cost of magic-state distillation may be reduced substantially in the near future. For instance, if we use the standard 15-to-1 magic state distillation protocol [14], the error rate of the magic state can be reduced from p to $\sim 35p^3$. Plugging in the infidelity of our magic state, we obtain an expected magic state infidelity of at most 4.6×10^{-9}. This is a number low enough to study the challenging 2D Hubbard model [45] and it is even lower than the inverse of the number of T gates to run quantum algorithms for FeMoco [46–48]. We anticipate this finding to affect future studies of resource estimate for quantum algorithms on a realistic quantum computer.

Our work also highlights the importance of verified preselection in designing fault-tolerant quantum computation

protocols [12,20,38,49,50]. Certain parts of quantum computations can be repeated probabilistically until success, without affecting other parts of the computation. Our work shows that there is a large design space for improving the existing protocols this way, even in the near term.

One might suggest the next frontier is to demonstrate a logical $\bar{T}$ gate with an infidelity lower than the physical T-gate error rate. Because the physical single-qubit error rate of H2-1 quantum processor is $1.9(4) \times 10^{-5}$, the magic state created in this experiment is not sufficient for this goal. If the logical error rate is dominated by two-qubit gate errors p, and we assume the logical error rate is proportional to $\propto p^2$ in the $d = 3$ protocol, the two-qubit gate error would need to improve by approximately $\sqrt{(5.1 \times 10^{-4})/(1.9 \times 10^{-5})} \sim 5\text{x}$ for this goal. Improving the two-qubit gate fidelity by 5x is not inconceivable, as trapped-ions have achieved gate errors near this level using microwave gates [51]; however, such gates have yet to be demonstrated in a system of many interacting qubits.

Extending our approach to a higher-distance instance of the color code may prove to be the most straightforward approach once more qubits are available in next-generation devices. In such a case, because the number of physical qubits scales as $\Theta(d^3)$, and the postselection rate is going to be exponentially suppressed (although strategies with a partial level of postselection could be engineered), we expect the role of our protocol to be slightly different than magic state distillation, since the size of the 3D color code should be capped. Indeed, our protocol can be seen as a source of high-fidelity magic states for 2D color codes at some finite distance, which will need to be further complemented with a procedure that increases the code distance, together with at least one round of magic state distillation.

Furthermore, we advise some healthy skepticism in the definition of the next frontier regarding a logical $\bar{T}$-gate breaking even the physical T-gate error rate, as one can imagine this goal being easier to achieve in a system with a slightly worse single-qubit gate error (since the logical error could still be dominated by the two-qubit gate error), and it seems troublesome to define milestones that are easier to achieve with lower performance machines. Another approach to gauging progress is that once all logical operations have error rates below the dominant source of noise in the physical layer (which we have now achieved), the next frontier is to demonstrate algorithmic break-even through lower circuit infidelities, where the circuits use a universal gate set in a standardized way. Of course, defining such a family of circuits is a complex task and beyond the scope of this work. Finally, extending our approach to other code families that admit a transversal logical non-Clifford gate such as $\overline{\text{CCZ}}$ [52], could be another interesting direction to pursue in the future.

ACKNOWLEDGMENTS

We thank Ciarán Ryan-Anderson for discussions. This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

L. D. and I. K. devised the project and designed the protocol implementation. N. B. and D. H. helped with adapting the protocol to the hardware. L. D. implemented the code for circuit generation. L. D. and N. B. executed experiments on the quantum computer and collected data. L. D. Analyzed the data. R. B.-K. devised the sample-efficient magic state certification protocol. L. D. improved bounds for the magic state fidelity bound. L. D. and I. K. contributed to the writing of the manuscript. I. K. led the overall project as the lead principal investigator. All authors contributed to the technical discussions and editing of the manuscript.

The authors declare no other competing interests.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available. The data are available from the authors upon reasonable request.

APPENDIX A: EXPERIMENTAL SETUP

We used Quantinuum's H2-1 trapped-ion processor based on a QCCD architecture [44] to perform the experiments in this article. Regarding the quantum processor, the most recent randomized benchmarking experiments averaged over all four gate zones showed fidelities of $p_1 = 1.9(4) \times 10^{-5}$ and $p_2 = 1.05(8) \times 10^{-3}$ for single-qubit and two-qubit gates, respectively. SPAM errors are dominated by $|1\rangle$ state measurement errors $p_{|1\rangle} = 1.38(12) \times 10^{-3}$, over $|0\rangle$ state measurement errors $p_{|0\rangle} = 6.0(8) \times 10^{-4}$. Crosstalk errors from midcircuit SPAM operations are $p_{\text{ct}} = 6.6(8) \times 10^{-6}$.

Memory errors per qubit were benchmarked in Ref. [44] using random circuits, resulting in an estimate for memory error per circuit layer (∼60 ms) of $p_{\text{idle}} = 2.0(2) \times 10^{-4}$. However, the circuits used in this work have a very different structure, with some qubits idling for multiple circuit layers. To mitigate coherent accumulation of memory error, we apply dynamical decoupling pulses approximately every 30 ms, ideally keeping the memory error in each layer of gates $\leq 2.0(2) \times 10^{-4}$ while adding a small amount of error from the decoupling pulses $\sim 1.9(4) \times 10^{-5}$. Confirming these effects is difficult in simulation due to the incompatibility of stabilizer simulations with coherent

error accumulation, and a comprehensive experimental characterization of individual error sources' impacts is beyond the scope of this work.

The experiments were prepared using the OPENQASM2.0 programming language, designed to represent quantum circuits. The quantum circuits are then converted to the `Circuit` class using the `pytket` module, so they can be submitted to the H-series compiler stack where they are converted into machine level instructions. In this process, no optimization is utilized. The quantum circuits are also designed such that no classically controlled operations are applied, thus minimizing the latency time, though we note that previous experiments [5] have shown that real-time application of corrections can be done without significant error.

APPENDIX B: EXPERIMENTAL METHODS

In this Appendix we provide additional information about the data analysis. A summary of the experimental results and numerical emulations is shown in Table II.

In the single-copy experiments (Fig. 5) the output of a given shot is a binary string with 31 entries. Out of those 31 bits, 7 of them correspond to Steane code data qubits, 15 to qRM code data qubits, 4 to Z-type qRM code stabilizers, and 5 to flags (2 from state initialization, and 3 from the Z-type stabilizer circuit measurements). We pre-select and analyze only those shots such that the 4Z-type stabilizers and 5 flags are $+1$. Moreover, from the 15 entries corresponding to the X-type destructive measurements of the qRM code data qubits, we extract the qRM X-type syndrome using the definitions shown in Eq. (C4). In case any of these X-type syndromes are nontrivial, we discard the shot. Finally, the information from the remaining 7 Steane code data qubits is analyzed to extract logical Pauli basis expected values $\langle\bar{P}\rangle_{\bar{\rho}}$ for $\bar{P}\in\{\bar{X},\bar{Y},\bar{Z}\}$. First, the Steane code syndrome is extracted using (C1), followed by a correction which is applied based on a look-up table decoder (in cases where postselection is considered, only the instances with trivial syndromes are analyzed). Then, the logical operator $\bar{P}$ of the single-copy experiment is computed using the definitions shown in Eq. (C2). Since the logical $\bar{Z}$ correction of Fig. 3 is not physically applied, when the logical $\bar{X}^{\mathrm{qRM}}=-1$ in the qRM code, the resultant logical operator $\bar{P}$ is flipped if $\bar{P}=\bar{X}$ or $\bar{Y}$.

Similarly, the data analysis of the two-copy experiment of Fig. 5 entails decoding a 62 binary string for every shot. Only shots where the qRM X and Z-type stabilizers and flags are $+1$ for both copies are kept. Thus, the acceptance rate of the two-copy experiment is roughly the square of the acceptance rate of the single-copy experiment. To extract the logical $\bar{X}_1\otimes\bar{Z}_2$ operator, similar to the decoding process of the single-copy experiment, $\bar{X}_1$ and $\bar{Z}_2$ are obtained for the respective Steane code copy. Since the logical $\bar{Z}$ corrections of the two copies of the protocol of Fig. 3 are not physically applied, they are considered in postprocessing. Because $\overline{\mathrm{CNOT}_s}(\bar{Z}_1^{a_1}\otimes\bar{Z}_2^{a_2})\overline{\mathrm{CNOT}_s}=\bar{Z}_1^{a_1+a_2}\otimes\bar{Z}_2^{a_2}$ [where $a_i=(1-\bar{X}_i^{\mathrm{qRM}})/2$ indicates whether a correction has to be applied to the ith copy], the logical measurement of the first copy $\bar{X}_1$ has to be flipped when $a_1+a_2\equiv 1 \bmod 2$.

The values reported in Table II correspond to the mean over all the postselected shots N_{post}, and the acceptance rate is the number of post-selected shots over the total number of shots N_{shots}. The reported acceptance rate for the fidelity in Tables I and II is the average over the acceptance rates of the single-copy experiments.

The (1σ) confidence error bars are computed as $\Delta\langle\bar{P}\rangle_{\bar{\rho}}=\sqrt{\mathrm{Var}(\bar{P})/N_{\mathrm{post}}}$ for the single-copy experiments, where $\mathrm{Var}(\bar{P})$ is the sample variance. For the two-copy experiment, the (1σ) error bar is computed using a binomial distribution $\Delta\epsilon=\sqrt{p_f(1-p_f)/N_{\mathrm{post}}}$, where $p_f=N_{\overline{|-1\rangle}}/N_{\mathrm{post}}$ is the fraction of detected logical $|\overline{-1}\rangle$ states $N_{\overline{|-1\rangle}}$. In a case where $N_{\overline{|-1\rangle}}=0$, the (1σ) confidence interval is $[0,1.147/N_{\mathrm{post}})$ [53], so we choose $p_f=\Delta\epsilon=1.147/2N_{\mathrm{post}}$. Finally, the (1σ) error bar for the fidelity lower bound for F (5) is computed assuming uncorrelated error propagation.

TABLE II. Experimental results for the single- and two-copy experiments analyzed using an ideal round of error correction or postselection (error detection). The fidelity lower bound is computed using (5). The fidelity acceptance rate is the average over the acceptance rates of the single-copy experiments.

	Error correction		Postselection	
Quantity	Value	Acceptance rate (%)	Value	Acceptance rate (%)
$F(\bar{\rho},\lvert\bar{T}\rangle\langle\bar{T}\rvert)$	0.99949(27)	82.58	0.99985(11)	81.26
ϵ	$4.56(2.63)\times10^{-4}$	65.73	$9.28(9.28)\times10^{-5}$	61.83
$\langle\bar{X}\rangle_{\bar{\rho}}$	0.6993(78)	83.56	0.6987(78)	82.66
$\langle\bar{Y}\rangle_{\bar{\rho}}$	0.7193(77)	81.24	0.7197(77)	79.72
$\langle\bar{Z}\rangle_{\bar{\rho}}$	0.0000(109)	82.96	−0.0007(110)	81.42

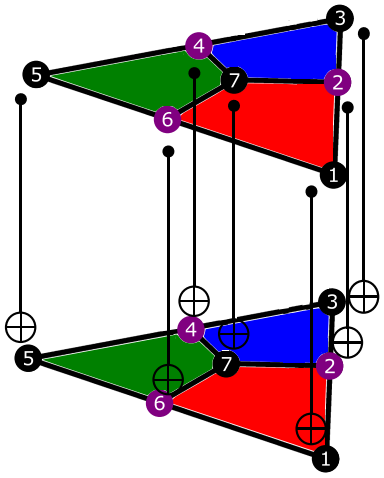


FIG. 6. Logical $\overline{\text{CNOT}_s}$ gate between Steane codes, consisting of physical CNOT gates between corresponding physical qubits.

APPENDIX C: COLOR CODES

In this Appendix we describe properties of the Steane $[[7,1,3]]$ code (Appendix C 1) and the quantum Reed-Muller $[[15,1,3]]$ code (qRM) (Appendix C 2).

1. Steane code

The stabilizer group S_{2D} of the Steane $[[7,1,3]]$ code is generated by

$$\begin{aligned} p_1^Z &= Z_1Z_2Z_6Z_7, \quad & p_1^X &= X_1X_2X_6X_7, \\ p_2^Z &= Z_2Z_3Z_4Z_7, \quad & p_2^X &= X_2X_3X_4X_7, \\ p_3^Z &= Z_4Z_5Z_6Z_7, \quad & p_3^X &= X_4X_5X_6X_7. \end{aligned} \tag{C1}$$

These stabilizer generators correspond to weight-4 2-cells or plaquettes, as it can be observed from Fig. 1. Note that since the Steane code is a self-dual code and the stabilizer weights are 4, the Y-type stabilizers are $p_i^Y = p_i^X p_i^Z$ for $i = 1, 2, 3$. Minimum weight representatives of the Pauli logical operators $\bar{X}$, $\bar{Y}$, and $\bar{Z}$ are

$$\bar{X} = X_1X_2X_3, \quad \bar{Y} = -Y_1Y_2Y_3, \quad \bar{Z} = Z_1Z_2Z_3. \tag{C2}$$

The logical $|\bar{0}\rangle$ state is thus a $+1$ eigenstate of $\bar{Z}|\bar{0}\rangle = |\bar{0}\rangle$ as well as of the stabilizer group generators. In the protocol of Fig. 3, the $|\bar{0}\rangle$ is initialized using the technique of [38], which is shown in Fig. 7(a).

The logical $\overline{\text{CNOT}_s}$ gate utilized in the two-copy experiment of Fig. 5(b) is shown in Fig. 6.

2. Quantum Reed-Muller code

The stabilizer group S_{3D} of the quantum Reed-Muller [[15,1,3]] (qRM) code is generated by Z-type generators

$$\begin{aligned} p_1 &= Z_1Z_2Z_6Z_7, \quad & p_{10} &= Z_6Z_7Z_{13}Z_{14}, \\ p_2 &= Z_2Z_3Z_4Z_7, \quad & p_{11} &= Z_2Z_7Z_9Z_{14}, \\ p_3 &= Z_4Z_5Z_6Z_7, \quad & p_{12} &= Z_4Z_7Z_{11}Z_{14}, \\ p_4 &= Z_1Z_6Z_8Z_{13}, \quad & p_{13} &= Z_8Z_{12}Z_{13}Z_{15}, \\ p_5 &= Z_1Z_2Z_8Z_9, \quad & p_{14} &= Z_8Z_9Z_{10}Z_{15}, \\ p_6 &= Z_2Z_3Z_9Z_{10}, \quad & p_{15} &= Z_{10}Z_{11}Z_{12}Z_{15}, \\ p_7 &= Z_3Z_4Z_{10}Z_{11}, \quad & p_{16} &= Z_8Z_9Z_{13}Z_{14}, \\ p_8 &= Z_4Z_5Z_{11}Z_{12}, \quad & p_{17} &= Z_9Z_{10}Z_{11}Z_{14}, \\ p_9 &= Z_5Z_6Z_{12}Z_{13}, \quad & p_{18} &= Z_{11}Z_{12}Z_{13}Z_{14}, \end{aligned} \tag{C3}$$

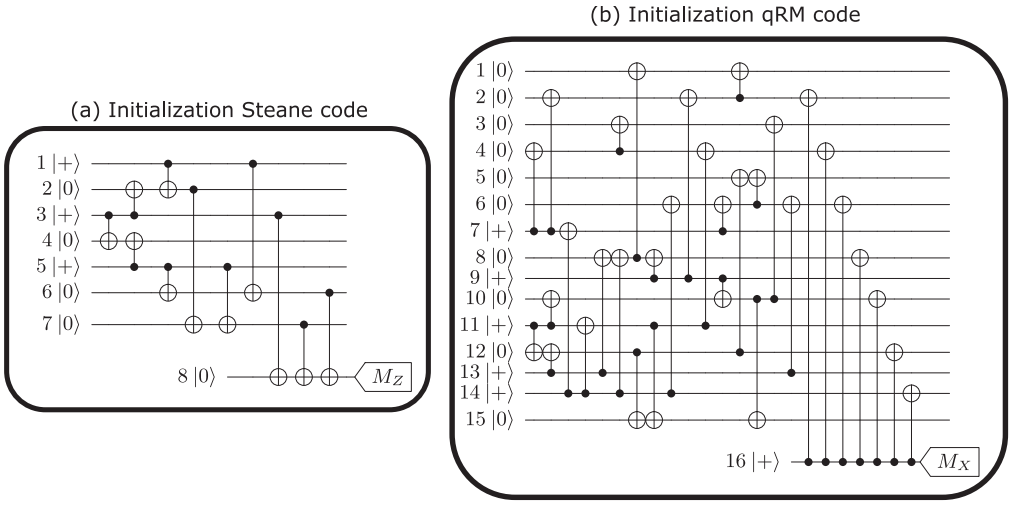


FIG. 7. (a) Flag-based state initialization for the logical $|\bar{0}\rangle$ state in the Steane code and (b) for the logical $|\overline{+}\rangle$ state in the qRM code. The initialization consist of a nonfault tolerant encoding followed by a verification step, which detects propagated uncorrectable X errors for (a) and Z errors for (b), respectively.

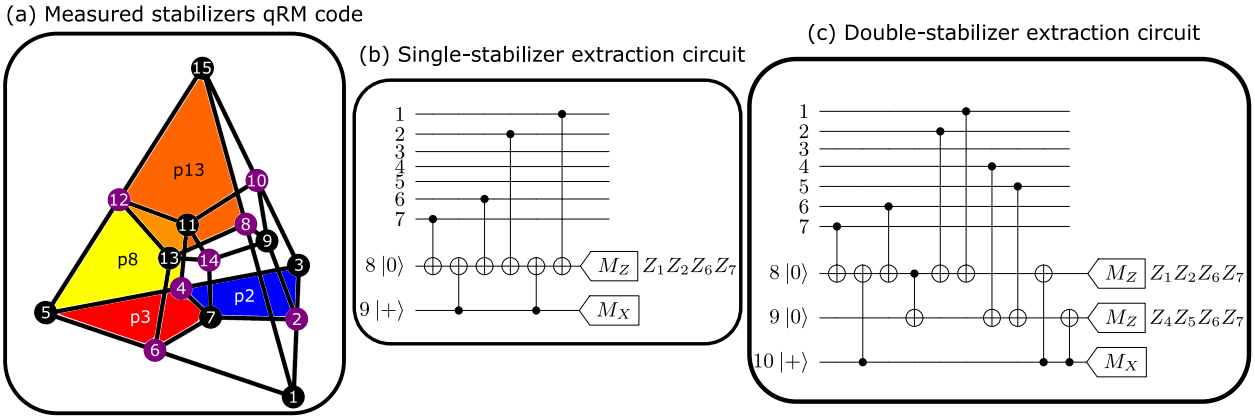


FIG. 8. (a) Set of Z-type plaquette stabilizers of the qRM code $\{p_2, p_3, p_8, p_{13}\}$, which can detect high-weight uncorrectable X errors originating from the encoding of the logical $|\overline{+}\rangle$ state, Fig. 7(b). (b) Flag-based single-stabilizer extraction circuit, used for measuring the Z-type plaquettes p_8 and p_{13}. (3) Parallel-flag-based double-stabilizer extraction circuit, used for simultaneously measuring the Z-type plaquettes $\{p_2, p_3\}$.

as well as by X-type generators,

$$\begin{aligned} c_1 &= X_1X_2X_6X_7X_8X_9X_{13}X_{14}, \\ c_2 &= X_4X_5X_6X_7X_{11}X_{12}X_{13}X_{14}, \\ c_3 &= X_2X_3X_4X_7X_9X_{10}X_{11}X_{14}, \\ c_4 &= X_8X_9X_{10}X_{11}X_{12}X_{13}X_{14}X_{15}. \end{aligned} \tag{C4}$$

From the total of 18 weight-4 plaquette Z-type stabilizer generators, only 10 of them are independent, for instance $\{p_1, p_2, p_3, p_7, p_8, p_9, p_{13}, p_{16}, p_{17}, p_{18}\}$ [20]. The X-type generators correspond to weight-8 3-cells, as it can be observed from Fig. 2. Minimum weight representatives of the Pauli logical operators $\bar{X}$ and $\bar{Z}$ are

$$\bar{X} = X_1X_2X_3X_4X_5X_6X_7, \quad \bar{Z} = Z_1Z_2Z_3. \tag{C5}$$

The logical $|\overline{+}\rangle$ state is thus a $+1$ eigenstate of $\bar{X}|\overline{+}\rangle = |\overline{+}\rangle$ as well as of the stabilizer group generators. In the protocol of Fig. 3, the $|\overline{+}\rangle$ is initialized using the technique of [39], which is shown in Fig. 7(b).

Since uncorrectable higher-weight X-type errors are not detected in the initialization of the $|\overline{+}\rangle$ of Fig. 7(b), they can lead to a logical failure. Hence, a set of independent Z-type stabilizers should be measured before applying the logical $\bar{T}$ gates in the qRM code. However, if we are only interested in detecting uncorrectable weight-2 or higher X-type errors, it turns out that 4 stabilizers is enough [20]. Indeed, we measure the following stabilizers before applying the logical $\bar{T}$ gate $\{p_2, p_3, p_8, p_{13}\}$, see Fig. 8(a). The plaquettes p_{13} and p_8 are measured first, utilizing the flag-based stabilizer extraction circuit of [37] [see Fig. 8(b)], and finally p_2 and p_3 are simultaneously measured using the parallel flag-based stabilizer extraction circuit of [37], see Fig. 8(c).

APPENDIX D: LOGICAL STATE CERTIFICATION METHOD

In this Appendix we prove a lower bound for the fidelity $F(\bar{\rho}, \bar{\sigma}_T) = \mathrm{Tr}(\bar{\rho}\bar{\sigma}_T)$, between an experimental state $\bar{\rho}$ and a target ideal pure state $\bar{\sigma}_T$, both with Bloch sphere decompositions

$$\bar{\rho} = \frac{I + \vec{v}\cdot\vec{\sigma}}{2}, \qquad \bar{\sigma}_T = \frac{I + \vec{u}\cdot\vec{\sigma}}{2}, \qquad |\vec{u}|^2 = 1. \tag{D1}$$

We define the difference between their vector representation $\vec{\delta}$ as

$$\vec{v} = \vec{u} + \vec{\delta}. \tag{D2}$$

The vector $\vec{v} = (\langle\bar{X}\rangle_{\bar{\rho}}, \langle\bar{Y}\rangle_{\bar{\rho}}, \langle\bar{Z}\rangle_{\bar{\rho}})$. For a magic $|\bar{T}\rangle\langle\bar{T}|$ state $\vec{u} = \{(1/\sqrt{2}), (1/\sqrt{2}), 0\}$. The fidelity $F(\bar{\rho}, \bar{\sigma}_T)$ is

$$F(\bar{\rho}, \bar{\sigma}_T) = \frac{1}{2}(1 + \vec{u}\cdot\vec{v}) = 1 + \frac{1}{2}\vec{u}\cdot\vec{\delta}. \tag{D3}$$

On the other hand, the trace of the state squared is

$$\mathrm{Tr}(\bar{\rho}^2) = \frac{1}{2}(1 + |\vec{v}|^2) = 1 + \frac{1}{2}|\vec{\delta}|^2 + \vec{u}\cdot\vec{\delta}. \tag{D4}$$

Thus, combining (D3) and (D4) yields

$$\mathrm{Tr}(\bar{\rho}^2) = 1 + \frac{1}{2}|\vec{\delta}|^2 + 2(F(\bar{\rho}, \bar{\sigma}_T) - 1). \tag{D5}$$

The values of $\vec{\delta}$ can be obtained from the single-copy experiment. However, the trace of the state squared has to be estimated from an experiment involving at least two copies, see Sec. III. Indeed, similarly to the SWAP test [42], which determines the purity of a state, the projection onto

the singlet (antisymmetric) state $|\bar{\Phi}_1\rangle = (1/\sqrt{2})(|\overline{01}\rangle - |\overline{10}\rangle)$ is equal to

$$\begin{aligned}\mathrm{Tr}(\bar{\rho}_1\bar{\rho}_2) &= \mathrm{Tr}[\overline{\mathrm{SWAP}}(\bar{\rho}_1 \otimes \bar{\rho}_2)] \\ &= 1 - 2\langle\bar{\Phi}_1|\bar{\rho}_1 \otimes \bar{\rho}_2|\bar{\Phi}_1\rangle. \end{aligned} \tag{D6}$$

Note that we do not assume that both copies $\bar{\rho}_1$ and $\bar{\rho}_2$ are identical, even though they were produced using the same quantum circuits. The projection ϵ is defined as

$$\begin{aligned}\epsilon &\equiv \langle\bar{\Phi}_1|\bar{\rho}_1 \otimes \bar{\rho}_2|\bar{\Phi}_1\rangle \\ &= \langle\overline{-1}|\overline{\mathrm{CNOT}_s}(\bar{\rho}_1 \otimes \bar{\rho}_2)\overline{\mathrm{CNOT}_s}|\overline{-1}\rangle, \end{aligned} \tag{D7}$$

where $|\overline{-1}\rangle \equiv |\bar{-}\rangle_1|\bar{1}\rangle_2$. Because of the Cauchy-Schwarz inequality

$$\sqrt{\mathrm{Tr}(\bar{\rho}_1^2)\mathrm{Tr}(\bar{\rho}_2^2)} \geq \mathrm{Tr}(\bar{\rho}_1\bar{\rho}_2) = 1 - 2\epsilon. \tag{D8}$$

Moreover, because of the AM-GM inequality

$$\frac{\mathrm{Tr}(\bar{\rho}_1^2) + \mathrm{Tr}(\bar{\rho}_2^2)}{2} \geq \sqrt{\mathrm{Tr}(\bar{\rho}_1^2)\mathrm{Tr}(\bar{\rho}_2^2)} \geq 1 - 2\epsilon. \tag{D9}$$

Then, $1 - 2\epsilon$ is a lower bound for the maximum (or average) trace of the experimental state squared

$$\max\{\mathrm{Tr}(\bar{\rho}_i^2)\} \geq 1 - 2\epsilon. \tag{D10}$$

Finally, combining (D5) and (D10) we prove the lower bound on the fidelity for at least one copy

$$F(\bar{\rho}, \bar{\sigma}_T) \geq 1 - \left(\epsilon + \frac{\delta_X^2 + \delta_Y^2 + \delta_Z^2}{4}\right). \tag{D11}$$

[1] P. Shor, *Algorithms for quantum computation: Discrete logarithms and factoring*, in *Proceedings 35th Annual Symposium on Foundations of Computer Science, Santa Fe, NM, USA* (1994), pp. 124–134, 10.1109/SFCS.1994.365700.

[2] S. Lloyd, *Universal quantum simulators*, Science **273**, 1073 (1996).

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

[4] D. Aharonov and M. Ben-Or, *Fault-tolerant quantum computation with constant error*, in *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, STOC '97 (Association for Computing Machinery, New York, USA, 1997), pp. 176–188, 10.1145/258533.258579.

[5] C. Ryan-Anderson *et al.*, *Realization of real-time fault-tolerant quantum error correction*, Phys. Rev. X **11**, 041058 (2021).

[6] A. Erhard *et al.*, *Entangling logical qubits with lattice surgery*, Nature (London) **589**, 220 (2021).

[7] L. Postler *et al.*, *Demonstration of fault-tolerant universal quantum gate operations*, Nature (London) **605**, 675 (2022).

[8] D. Bluvstein *et al.*, *Logical quantum processor based on reconfigurable atom arrays*, Nature (London) **626**, 58 (2024).

[9] R. Acharya *et al.*, *Suppressing quantum errors by scaling a surface code logical qubit*, Nature (London) **614**, 678 (2023).

[10] B. W. Reichardt *et al.*, *Demonstration of quantum computation and error correction with a tesseract code*, arXiv:2409.04628.

[11] B. W. Reichardt *et al.*, *Logical computation demonstrated with a neutral atom quantum processor*, arXiv:2411.11822.

[12] A. Paetznick *et al.*, *Demonstration of logical qubits and repeated error correction with better-than-physical error rates*, arXiv:2404.02280.

[13] K. Mayer *et al.*, *Benchmarking logical three-qubit quantum Fourier transform encoded in the Steane code on a trapped-ion quantum computer*, arXiv:2404.08616.

[14] S. Bravyi and A. Kitaev, *Universal quantum computation with ideal Clifford gates and noisy ancillas*, Phys. Rev. A **71**, 022316 (2005).

[15] D. Litinski, *Magic state distillation: Not as costly as you think*, Quantum **3**, 205 (2019).

[16] 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).

[17] H. Bombín, *Dimensional jump in quantum error correction*, New J. Phys. **18**, 043038 (2016).

[18] M. E. Beverland, A. Kubica, and K. M. Svore, *Cost of universality: A comparative study of the overhead of state distillation and code switching with color codes*, PRX Quantum **2**, 020341 (2021).

[19] I. Pogorelov, F. Butt, L. Postler, C. D. Marciniak, P. Schindler, M. Müller, and T. Monz, *Experimental fault-tolerant code switching*, Nat. Phys. **21**, 298 (2025).

[20] L. Daguerre and I. H. Kim, *Code switching revisited: Low-overhead magic state preparation using color codes*, Phys. Rev. Res. **7**, 023080 (2025).

[21] H. Bombin and M. A. Martin-Delgado, *Topological quantum distillation*, Phys. Rev. Lett. **97**, 180501 (2006).

[22] H. Bombin and M. A. Martin-Delgado, *Topological computation without braiding*, Phys. Rev. Lett. **98**, 160502 (2007).

[23] A. Steane, *Quantum Reed-Muller codes*, arXiv:quant-ph/9608026.

[24] A. Steane, *Multiple particle interference and quantum error correction*, Proc. R. Soc. A **452**, 2551 (1996).

[25] L. Egan *et al.*, *Fault-tolerant control of an error-corrected qubit*, Nature (London) **598**, 281 (2021).

[26] J. F. Marques *et al.*, *Logical-qubit operations in an error-detecting surface code*, Nat. Phys. **18**, 80 (2022).

[27] R. S. Gupta *et al.*, *Encoding a magic state with beyond break-even fidelity*, Nature (London) **625**, 259 (2024).

[28] 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).

[29] Y. Kim, M. Sevior, and M. Usman, *Magic state injection on IBM quantum processors above the distillation threshold*, arXiv:2412.01446.

[30] P. S. Rodriguez *et al.*, *Experimental demonstration of logical magic state distillation*, arXiv:2412.15165.
[31] N. Lacroix *et al.*, *Scaling and logic in the color code on a superconducting quantum processor*, arXiv:2412.14256.
[32] S. Dasu, S. Burton, K. Mayer, D. Amaro, J. A. Gerber, K. Gilmore, D. Gresh, A. C. Potter, and D. Hayes, *Breaking even with magic: demonstration of a high-fidelity logical non-Clifford gate*, arXiv:2506.14688.
[33] S.-u. Lee, M. Yuan, S. Chen, K. Tsubouchi, and L. Jiang, *Efficient benchmarking of logical magic state*, arXiv:2505.09687.
[34] B. Eastin and E. Knill, *Restrictions on transversal encoded quantum gate sets*, Phys. Rev. Lett. **102**, 110502 (2009).
[35] S. Bravyi and R. Konig, *Classification of topologically protected gates for local stabilizer codes*, Phys. Rev. Lett. **110**, 170503 (2013).
[36] X. Zhou, D. W. Leung, and I. L. Chuang, *Methodology for quantum logic gate construction*, Phys. Rev. A **62**, 052316 (2000).
[37] R. Chao and B. W. Reichardt, *Quantum error correction with only two extra qubits*, Phys. Rev. Lett. **121**, 050502 (Aug, 2018).
[38] H. Goto, *Minimizing resource overheads for fault-tolerant preparation of encoded states of the Steane code*, Sci. Rep. **6**, 19578 (2016).
[39] F. Butt, S. Heußen, M. Rispler, and M. Müller, *Fault-tolerant code-switching protocols for near-term quantum processors*, PRX Quantum **5**, 020345 (2024).
[40] M. Sullivan, *Code conversion with the quantum Golay code for a universal transversal gate set*, Phys. Rev. A **109**, 042416 (2024).
[41] S. Heußen and J. Hilder, *Efficient fault-tolerant code switching via one-way transversal* CNOT *gates*, arXiv:2409.13465.
[42] A. Barenco, A. Berthiaume, D. Deutsch, A. Ekert, R. Jozsa, and C. Macchiavello, *Stabilization of quantum computations by symmetrization*, SIAM J. Comput. **26**, 1541 (2006).
[43] D. J. Wineland, C. Monroe, W. M. Itano, D. Leibfried, B. E. King, and D. M. Meekhof, *Experimental issues in coherent quantum state manipulation of trapped atomic ions*, J. Res. Natl. Inst. Stand. Technol. **103**, 259 (1998).
[44] S. A. Moses *et al.*, *A race-track trapped-ion quantum processor*, Phys. Rev. X **13**, 041052 (2023).
[45] E. T. Campbell, *Early fault-tolerant simulations of the Hubbard model*, Quantum Sci. Technol. **7**, 015007 (2022).
[46] M. Reiher, N. Wiebe, K. M. Svore, D. Wecker, and M. Troyer, *Elucidating reaction mechanisms on quantum computers*, Proc. Natl. Acad. Sci. U.S.A. **114**, 7555 (2017).
[47] Z. Li, J. Li, N. S. Dattani, C. J. Umrigar, and G. K.-L. Chan, *The electronic complexity of the ground-state of the* FeMo *cofactor of nitrogenase as relevant to quantum simulations*, J. Chem. Phys. **150**, 024302 (2019).
[48] G. H. Low, R. King, D. W. Berry, Q. Han, A. E. DePrince, A. White, R. Babbush, R. D. Somma, and N. C. Rubin, *Fast quantum simulation of electronic structure by spectrum amplification*, arXiv:2502.15882.
[49] T. Itogawa, Y. Takada, Y. Hirano, and K. Fujii, *Even more efficient magic state distillation by zero-level distillation*, arXiv:2403.03991.
[50] C. Gidney, N. Shutty, and C. Jones, *Magic state cultivation: Growing T states as cheap as* CNOT *gates*, arXiv:2409.17595.
[51] C. M. Löschnauer *et al.*, *Scalable, high-fidelity all-electronic control of trapped-ion qubits*, arXiv:2407.07694.
[52] A. Kubica, B. Yoshida, and F. Pastawski, *Unfolding the color code*, New J. Phys. **17**, 083026 (2015).
[53] J. A. Hanley and A. Lippman-Hand, *If nothing goes wrong, is everything all right? Interpreting zero numerators*, JAMA, J. Am. Med. Assoc. **249**, 1743 (1983).