

Efficient Magic State Distillation by Zero-Level Distillation

Tomohiro Itogawa^{1,*}, Yugo Takada¹, Yutaka Hirano¹ and Keisuke Fujii^{1,2,3,4}

¹Graduate School of Engineering Science, [Osaka University](#), 1-3 Machikaneyama, Toyonaka, Osaka 560-8531, Japan

²Center for Quantum Information and Quantum Biology, [Osaka University](#), 1-2 Machikaneyama, Toyonaka 560-0043, Japan

³[RIKEN Center for Quantum Computing \(RQC\)](#), Hirosawa 2-1, Wako, Saitama 351-0198, Japan

⁴Fujitsu Quantum Computing Joint Research Division, Center for Quantum Information and Quantum Biology, [Osaka University](#), 1-2 Machikaneyama, Toyonaka, Osaka 565-8531, Japan



(Received 21 November 2024; revised 28 February 2025; accepted 22 May 2025; published 20 June 2025)

Magic state distillation (MSD) is an essential element for universal fault-tolerant quantum computing, which distills a high-fidelity magic state from noisy magic states using ideal (error-corrected) Clifford operations. For ideal Clifford operations, it needs to be performed on the logical qubits and hence incurs a large spatiotemporal overhead, which is one of the major bottlenecks for the realization of fault-tolerant quantum computers (FTQCs). Here we propose zero-level distillation, which prepares a high-fidelity logical magic state at the physical level, namely *zero level*, using physical qubits and nearest-neighbor two-qubit gates on a square lattice. We develop a zero-level distillation circuit and show that distillation can be made even more efficient than the conventional sophisticated approaches with logical level distillations. The key idea involves the Knill *et al.*-type distillation using the Steane code and its careful mapping to the square-lattice architecture with error detection. The distilled magic state on the Steane-code state is then teleported or converted to surface codes. We numerically find that the error rate of the logical magic state scales as approximately $100 \times p^2$ in terms of the physical error rate p . For example, with a physical error rate of $p = 10^{-4}$ (10^{-3}), the logical error rate is reduced to $p_L = 10^{-6}$ (10^{-4}), resulting in an improvement of 2 (1) orders of magnitude. This contributes to reducing both space and time overhead for early FTQC as well as full-fledged FTQC combined with conventional multilevel distillation protocols.

DOI: [10.1103/thxx-njr6](https://doi.org/10.1103/thxx-njr6)

I. INTRODUCTION

Quantum computers are expected to provide advantages in solving problems that are intractable for classical computers such as prime factorization [1], linear system solver [2], and quantum chemistry [3]. Significant experimental efforts have been dedicated to the realization of quantum computers based on various physical systems. Among these, the superconducting system is one of the most promising candidates, and systems with 50–100 qubits have already been experimentally demonstrated [4,5]. These quantum computers are called noisy intermediate-scale quantum computers (NISQ) [6], whose noise level is still high and the number of qubits is still limited, are currently unable to run sophisticated quantum algorithms with theoretically proven quantum speedup.

Although NISQ-aware algorithms are being developed [7], an ultimate solution to these problems is to protect quantum information through quantum error correction [8] to realize a fault-tolerant quantum computer (FTQC) [9].

Surface codes [10,11] are one of the most promising approaches for a fault-tolerant quantum computer using superconducting qubits, since they can be implemented on a two-dimensional square lattice and have high noise resilience [12,13]. In FTQC, the entire computation has to be performed fault tolerantly, with quantum information encoded in logical qubits. While Clifford gates can be implemented relatively easily in a fault-tolerant manner, non-Clifford gates, such as the T gate, are hard to execute fault tolerantly [14]. Therefore, magic state distillation (MSD) [15] is employed to prepare a high-fidelity magic state $T|+\rangle$ from noisy ones, which is hence used to implement the T gate via gate teleportation [16].

While MSD is a crucial operation for achieving universal quantum computation, it requires a large number of qubits, which is an obstacle to the realization of FTQC [17]. This is because in most MSD protocols distillation is performed by using logical qubits with concatenating a QEC code, which features transversal implementations of

*Contact author: u348506d@ecs.osaka-u.ac.jp

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.

the H or T gates. In order to mitigate this, physical-level distillation protocols have been proposed with error detection [18–20]. Physical-level distillation has great potential to reduce the physical overhead, and its effectiveness has been demonstrated in an experiment with trapped ions [21]. However, this physical-level approach has not been used for resource estimation because it employs all-to-all or complicated gate connectivity, and a concrete implementation on an architecture compatible with surface codes is lacking.

In this work, we propose *zero-level distillation* to prepare a high-fidelity logical magic state using physical qubits and nearest-neighbor two-qubit gates on a square lattice. Here zero level refers to performing distillation entirely at the physical level instead of distillation at a higher level using logical qubits. More precisely, in zero-level distillation, a magic state is initially encoded into the Steane code non-fault-tolerantly. Then this noisy magic state is verified using the Hadamard test of the logical H gate following the Knill *et al.*-type protocol [22]. The logical magic state is teleported to the rotated or planar surface code. In addition to the teleportation-based approach, we also develop a fault-tolerant code conversion from the Steane code to the surface code, which allows us to reduce the number of qubits further while the depth is increased.

Throughout these processes, the operations employed are single-qubit gates and measurements and nearest-neighbor two-qubit gates. Furthermore, any single-point error is detected by syndrome measurements and the Hadamard test, and hence the logical error rate scales as $\mathcal{O}(p^2)$ with respect to the physical error rate p . Since QEC codes are not needed for fault-tolerant logical Clifford gates, the spatial overhead is much smaller than that of conventional MSD protocols.

It might be thought that the advantage of the proposed zero-level distillation is limited to be the case with very low physical error rate because of the square-lattice constraint. However, this is not the case if the distillation circuit is carefully designed. For performance analysis, we fully simulate zero-level distillation circuits consisting of approximately 40 qubits, reducing the number of qubits required for simulation to 23 qubits. As a result, zero-level distillation reduces the logical error rate of a magic state p_L to $p_L \simeq 100 \times p^2$ by using only a physical depth of 25 for the rotated surface code (23 for the planar surface code). For example, in the case of $p = 10^{-4}$ ($p = 10^{-3}$), the logical error rate results in $p_L = 10^{-6}$ ($p_L = 10^{-4}$), while the success rate is fairly high, 70% (95%). Although the $100p^2$ scaling is slightly worse than $35p^3$ for the conventional method, it offers practical advantages in terms of the overall spatiotemporal overhead. Specifically, in the context of the early-FTQC architecture [23], the zero-level distillation facilitates a significant enhancement, increasing the reliability of non-Clifford gate operations by 2 orders of magnitude, in contrast to a scenario without MSD. For

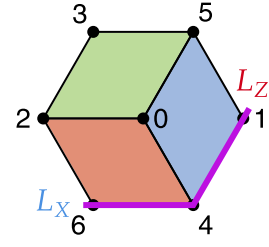


FIG. 1. The Steane code. A qubit is located at each vertex. The red, blue, and green faces represent stabilizers.

full-fledged FTQC, zero-level distillation combined with conventional multilevel distillation [24] allows us to save the number of physical qubits significantly to achieve a given accuracy by reducing the number of levels.

The rest of the paper is organized as follows. In Sec. II, we provide a preliminary explanation of the Steane code and MSD. In Sec. III, we provide a detailed description of our proposal, zero-level distillation. In Sec. IV, we explain the protocol to convert the Steane-code state to the rotated surface code directly. In Sec. V, we present the outcomes of our numerical simulation. In Sec. VI, we discuss the implications of zero-level distillation to early-FTQC and full-fledged FTQC. Then, Sec. VII is devoted to a conclusion.

II. PRELIMINARY

In this section, we provide a preliminary explanation of the Steane code and MSD based on the transversality of the Hadamard gate on it.

The Steane code is a $[[7, 1, 3]]$ stabilizer code, which can correct an arbitrary single-qubit Pauli error or detect an arbitrary two-qubit Pauli error. Stabilizer generators are defined by $XIXIXI$, $XXIIXI$, $XIXXIXI$, $ZIZIZI$, $ZZIIZI$, $ZIZZIZI$, each corresponds to one of the three colored faces in Fig. 1. On the Steane code, all Clifford gates can be implemented transversally, while non-Clifford gates such as the T gate cannot. Nonetheless, leveraging the transversality of the Hadamard H gate, a special resource state, referred to as the magic state, can be prepared fault tolerantly as follows.

A magic state $|A\rangle$ is defined as $|A\rangle \equiv e^{-i\frac{\pi}{8}Y}|+\rangle$, which is an eigenstate of H . Leveraging the transversality of H gate and controlled-Hadamard gates $\Lambda(H)$, we implement a Hadamard test of the logical Hadamard operator followed by decoding, as shown in Fig. 2 [22,25], where the physical controlled-Hadamard gate $\Lambda(H)$ is constructed by $A^\dagger \Lambda(X) A$ using the CNOT gate $\Lambda(X)$ and the non-Clifford gate A . Suppose that all Clifford gates are ideal, and only non-Clifford gates are subject to errors with a probability p , the above Hadamard test of the logical Hadamard operator and decoding process detect up to two errors. This provides the magic state with error rate $O(p^3)$. In order to make

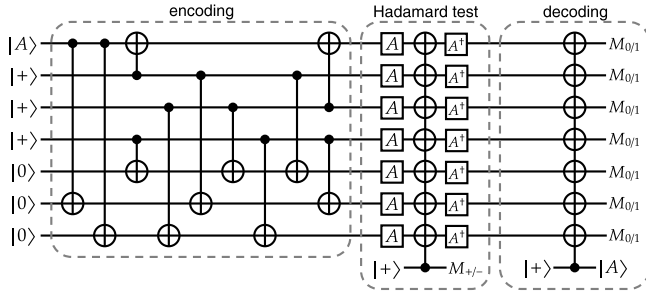


FIG. 2. MSD circuits based on transversality of the H gate. The encoding circuit creates a logical magic state encoded with the Steane code. The Hadamard test distills the magic state by measuring $H^{\otimes 7}$. The decoding circuit is based on the one-bit teleportation.

Clifford gates ideal, each qubit in Fig. 2 is further encoded into a logical qubit, and error-corrected Clifford gates are employed for distillation. However, this approach costs a large amount of physical qubits and operations due to the concatenation of two QEC codes.

A lower-cost distillation method with flag qubits has been proposed by Goto [18]. This protocol works with noisy Clifford gates by carefully designing the circuit with flag qubits so that any single-point error during the Clifford gates does not compromise the entire distillation process. Since this approach eliminates the need for logical qubits in Clifford gates, it can substantially reduce the number of qubits needed for MSD. However, the original proposal in Ref. [18] relies on the Steane code and all-to-all gate connectivity, which makes it difficult to apply for FTQC with the surface code on a nearest-neighbor architecture.

III. ZERO-LEVEL DISTILLATION

In order to reduce the overhead of magic state distillation, we propose *zero-level distillation*. This

protocol prepares a logical magic state on the surface code without using multiple logical qubits. We achieve physical-level distillation using the Steane code by carefully designing a fault-tolerant distillation circuit with fewer physical qubits and nearest-neighbor two-qubit gates on a square lattice. Then, the logical magic state is teleported from the Steane code to the surface code, which can be viewed as lattice surgery between color and surface codes [26]. (In the next section, we will also provide an alternative approach based on a code conversion from the Steane code to the surface code without teleportation.) This combination of two QEC codes allows us to prepare the logical magic state using fewer physical qubits with high fidelity on a square lattice as detailed below.

Zero-level distillation consists of three key processes: non-fault-tolerant magic state encoding, postselection via the Hadamard test, and teleportation-based injection using lattice surgery. The detailed steps are as follows (see the schematic circuit diagram in Fig. 3):

- (i) Encode a magic state in the Steane code non-fault-tolerantly. In parallel, a seven-qubit cat state is also prepared.
- (ii) Execute distillation by the Hadamard test using the cat state. If the parity of the measurement outcome is even, the output state is accepted. In parallel, encode $|+\rangle_L$ with the rotated surface code.
- (iii) Merge and split the magic state and $|+\rangle_L$, and perform the projection by the logical ZZ operator via the lattice surgery.
- (iv) Measure the Z stabilizers on the Steane code and the surface code.
- (v) Measure qubits on the Steane code directly in X basis to complete teleportation.

Figure 3 shows the circuit for encoding a noisy logical magic state and preparing the cat state corresponding to

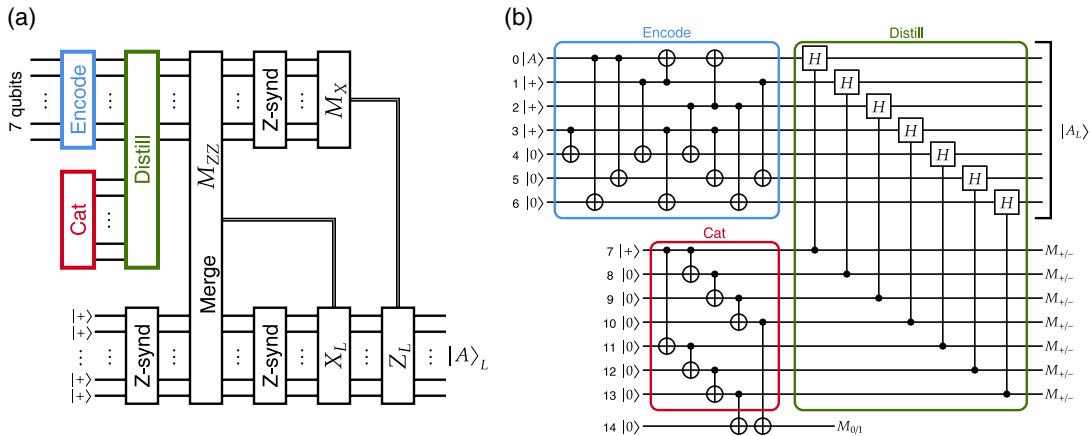


FIG. 3. (a) The zero-level distillation circuit. (b) The detailed circuits for encoding of noisy magic state, preparation of cat state, and distillation. The circuit in the blue box encodes a magic state encoded with the Steane code non-fault-tolerantly. The circuit in the red box prepares a cat state. The circuit in the green box is the distillation circuit utilizing the Hadamard test.

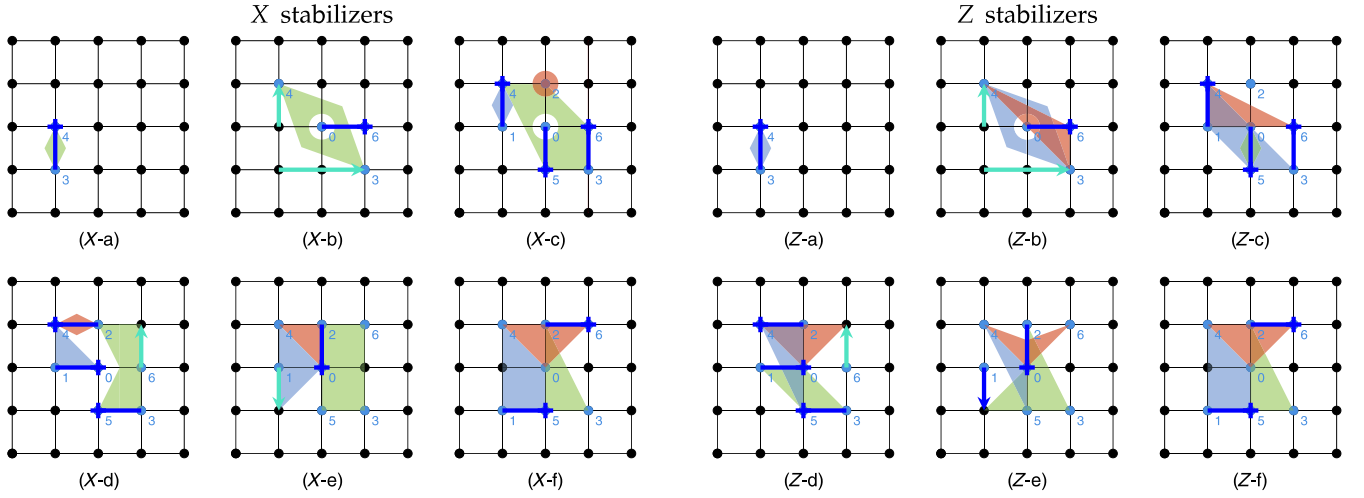


FIG. 4. Qubit arrangement and stabilizer deformation during encoding a magic state. The left figure represents the X -stabilizer deformation, and the right figure illustrates the Z -stabilizer deformation. Each dot represents a qubit: the blue dots and numbers correspond to the locations of the qubits shown in Fig. 3, and the light blue arrows indicate the transfer of qubits using one-bit teleportations. The blue bold edges indicate the CNOT gates, where the + symbol on one side represents the target qubit.

step (i). The blue box in Fig. 3 shows the circuit for non-fault-tolerant encoding of a magic state using the Steane code. As shown in Fig. 4, the circuit can be constructed by using nearest-neighbor two-qubit gates on a square lattice.

The red box in Fig. 3 shows the circuit for preparing a cat state. While in previous work [18], magic state distillation is performed using two ancilla qubits, a seven-qubit cat state is more suitable in our situation where the qubit connectivity is highly limited. Therefore, we use a seven-qubit cat state $\frac{1}{\sqrt{2}}(|0\rangle^{\otimes 7} + |1\rangle^{\otimes 7})$ as ancilla qubits. This allows data and ancilla qubits to be adjacent on a square lattice as shown in Figs. 4 and 5.

Note that in Figs. 4 and 5, some qubits have to be moved so that CNOT gates between neighboring qubits can be performed. Instead of using the swap operation, we employed one-bit teleportation to move a qubit to the neighboring site. This is simply because the swap operation requires three CNOT gates, which results in a greater circuit depth. However, if a swap gate can be implemented quickly and reliably, such as the i SWAP gate, it can be replaced with such a swap operation. In step (ii), a Hadamard test for the logical Hadamard gate is performed using the cat state. The green box in Fig. 3 shows the distillation circuit with the Hadamard test, where the controlled-Hadamard gate $\Lambda(H)$ is implemented as $A^\dagger \Lambda(X) A$. Figure 6 (right) shows the details of the arrangement of qubits during distillation. Since the data and ancilla qubits are positioned adjacent to each other, CNOT gates can be applied directly between them. After the logical controlled-Hadamard gate, the ancilla qubits from 7 to 13 in Fig. 3 are measured in the X basis. If the parity of the measurement outcomes is odd indicating any error, the distillation process is rejected.

Otherwise, the output is accepted and proceeds to the next step. Next, some qubits are repositioned for the lattice surgery as shown in Fig. 7(b). Simultaneously, the $|+\rangle_L$ state is encoded using the rotated surface code by preparing $|+\rangle$ states and measuring the Z stabilizers as shown in Fig. 7(a).

In step (iii), Steane and rotated surface codes are merged as shown in Fig. 7(c), and then split via the lattice surgery. This results in a projection onto $L_Z^{\text{Steane}} \otimes L_Z^{\text{surface}}$ as its eigenvalue can be obtained from the product of Z -type operators at the boundary, indicated by the purple area of Fig. 7(c). During the lattice surgery, the Z stabilizers must be measured twice to detect measurement errors. If the measurement outcomes differ, the event is rejected, and the protocol must be restarted from the beginning.

In step (iv), Z stabilizers on the Steane code and the rotated surface code are measured. Figure 8 illustrates the syndrome measurement circuits for the Steane code [27], where the ancilla qubits in Fig. 8 correspond to red, blue, or green dots in Fig. 7(d). The measurements detect two-qubit errors introduced during merging, as well as errors during magic state encoding or $|+\rangle_L$ -state encoding.

In step (v), each qubit of the Steane code is directly measured in the X basis to obtain the eigenvalues of the X stabilizers and logical operators as shown in Fig. 7(d). Note that these eigenvalues should be interpreted appropriately according to the measurement outcome of the X stabilizer at the boundary for splitting as usually done in the lattice surgery. This completes the teleportation of the distilled magic state from the Steane code to the rotated surface code. As usual, the Pauli frame of the logical qubit is updated based on the measurement outcome. Throughout

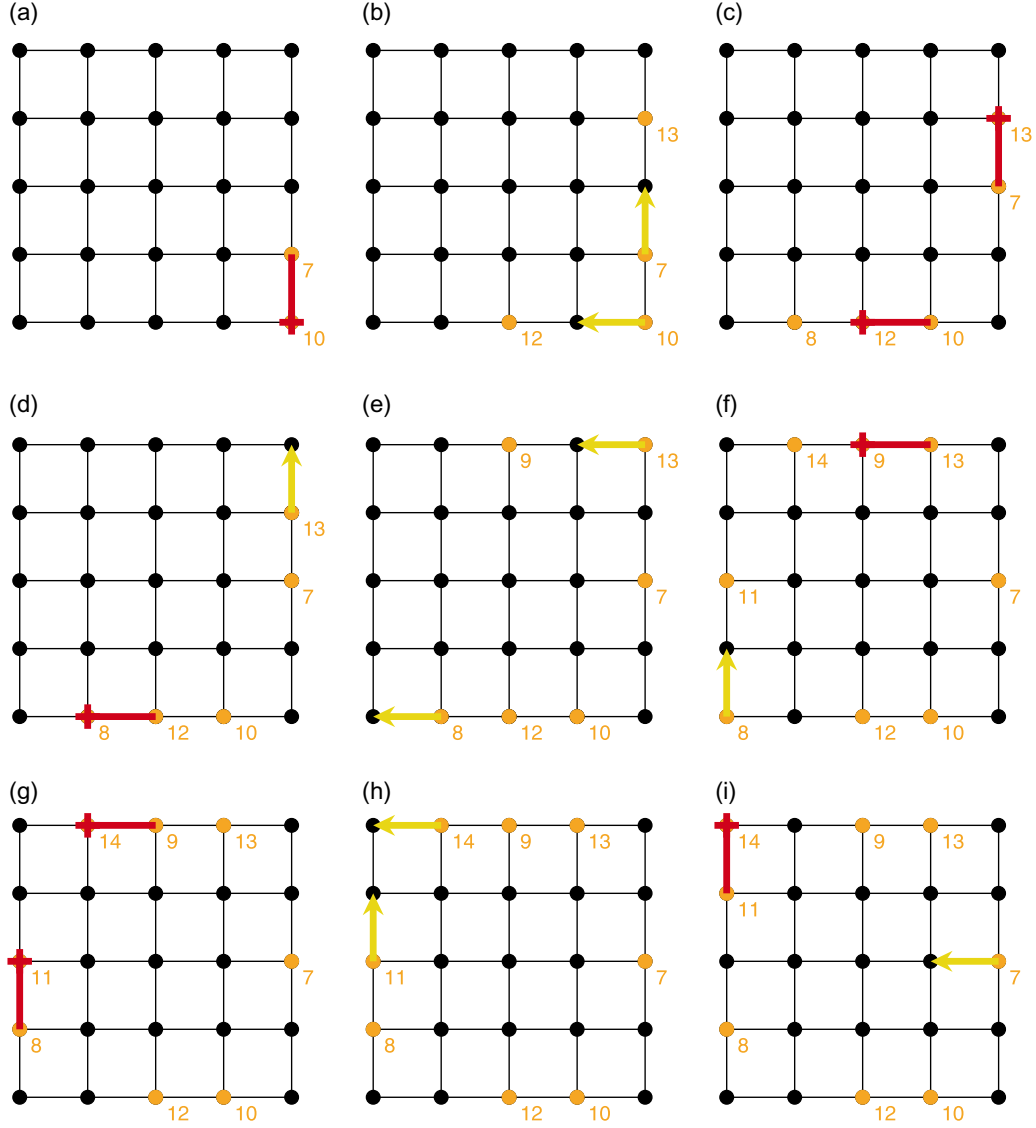


FIG. 5. Qubits arrangement during encoding a cat state. Each dot represents a qubit: the orange dots and numbers correspond to the locations of the qubits in Fig. 3. The yellow arrows indicate the transfer of qubits using one-bit teleportations. The red bold edges indicate the CNOT gates where + symbol on one side represents the target qubit.

the entire procedure, a physical depth of 25 is used. While we have demonstrated the proposed protocol using the rotated surface code, an implementation with the planar surface code is also provided in Appendix A.

The above zero-level distillation generates a magic state encoded in a $d = 3$ surface code, but it is desirable that the output state is kept on the surface code with a larger code distance to avoid error accumulation. Therefore, the surface code is further expanded fault tolerantly as shown in Fig. 9. In this process, $|0\rangle$ and $|+\rangle$ ancilla qubits are properly prepared [23] and stabilizer measurements are performed.

IV. DIRECT CODE CONVERSION FROM STEANE TO SURFACE CODES

We propose an alternative protocol to prepare the logical magic state on the rotated surface code by converting the Steane-code state to the rotated surface code directly, without relying on teleportation. This approach eliminates the need to generate two logical states and hence consumes fewer qubits compared to the teleportation-based approach described above. As a drawback, the circuit depth is increased to 42, roughly double that of the teleportation method. This leads to a higher logical error rate caused by additional idling noise.

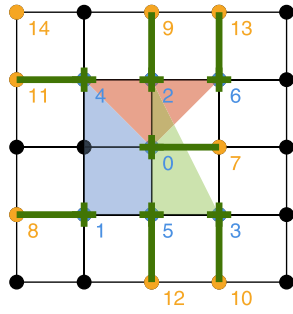


FIG. 6. Qubit arrangement during the Hadamard test. The blue and orange dots and numbers indicate the locations of the qubits shown in Fig. 3. A magic state and a cat state are encoded in the blue and orange qubits, respectively. The green edges indicate the CNOT gates, where + symbol on one side represents the target qubit.

First, as shown in Fig. 10(a), the magic state is encoded as described in Sec. III. Subsequently, as shown in Figs. 10(b)–10(l), we perform one-bit teleportations to move the qubits appropriately and stabilizer measurements to detect errors, and to form the stabilizer operators for the rotated surface code. Specifically, Figs. 10(c)–10(e) show Z -, X -, and Z -stabilizer measurements, respectively. If any of the syndrome values is odd in these three stabilizer measurements, the distillation process must be restarted from the beginning.

Next, we deform the Steane code to a six-qubit code of distance two while preserving the logical state and fault tolerance. As shown in Fig. 10(f), qubit 6 is measured directly in the X basis. The six-qubit code consists of two weight-four X and Z stabilizers and one weight-three X

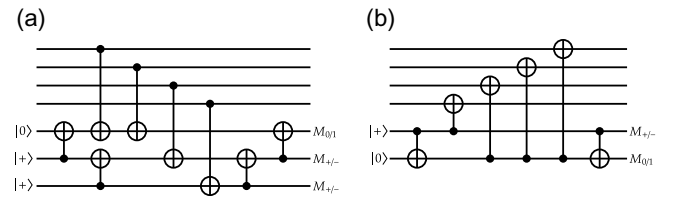


FIG. 8. (a) The syndrome measurement circuit for the Steane code using three ancilla qubits. (b) The syndrome measurement circuit using two ancilla qubits. They correspond to red and blue-green in Fig. 7(d).

stabilizer originating from the Steane code as shown in Fig. 10(g). Similar to the Steane code, the logical Z operator is composed of three qubits, $Z_1Z_3Z_5$. On the other hand, the logical X operator is composed of two qubits, X_1X_4 . In Fig. 10(h), certain qubits are prepared in the $|+\rangle$ states, which subsequently form part of an X stabilizer in the rotated surface code.

Finally, the six-qubit code is deformed to the rotated surface code by measuring appropriate stabilizers with obtaining the syndrome values inherited from the Steane code. In Fig. 10(i), two Z -type operators are measured; one is for the stabilizer of the rotated surface code and the other is to convert the X -type stabilizer of the Steane code (red face) into X stabilizers of the rotated surface code. The measurement outcomes of these two Z -type operators have to coincide since they are linked by the blue and green stabilizers of the Steane code. Therefore, if the measurement outcomes disagree indicating any error, the distillation process is rejected. Similarly, in Fig. 10(j), two X -type operators are measured to form the X stabilizers

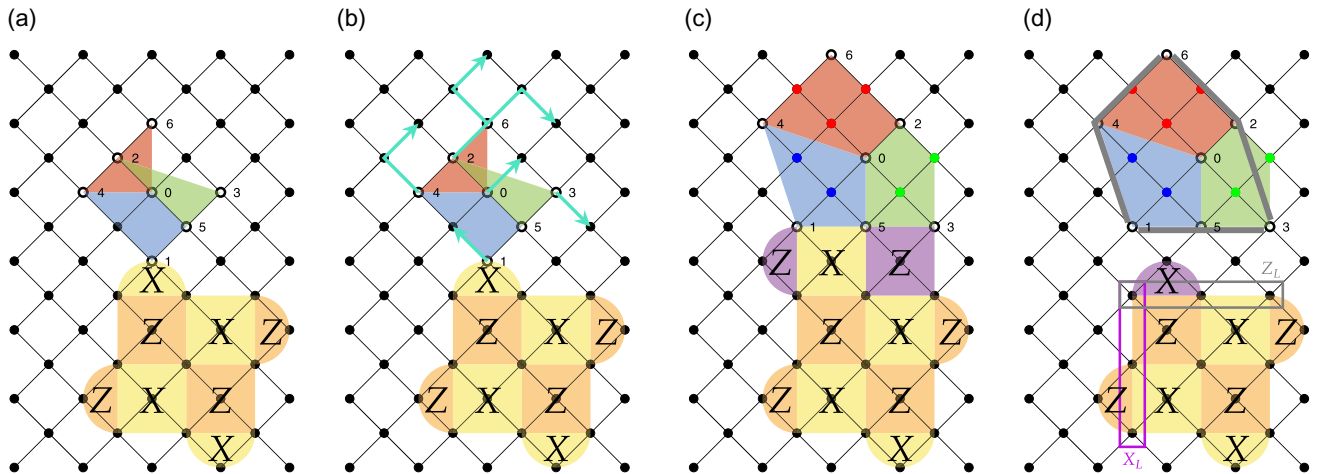


FIG. 7. Qubit arrangement and measurements during teleportation from the Steane code to the rotated surface code. The white dots and numbers indicate the locations of the qubits shown in Fig. 3, where the distilled magic state is encoded in the white qubits. The light blue arrows indicate the transfer of qubits using teleportation. The red, blue, and green faces indicate the stabilizers of the Steane code, and the orange and yellow faces indicate the stabilizers of the rotated surface code. The purple faces indicate the Z stabilizers for lattice surgery. The red, blue, and green dots are the ancilla qubits for the stabilizer measurements. The gray lines indicate logical operators of the Steane code. The purple box and the gray box indicate logical operators of the rotated surface code.

for the rotated surface code. If their measurement outcomes do not coincide, the distillation process is rejected. This measurement removes the Z stabilizer on the blue face, and only the X stabilizer is left (colored yellow). In Fig. 10(k), the X stabilizers are measured to be formed, which is repeated twice to detect measurement errors. In Fig. 10(l) two Z -type stabilizers are measured and formed, which are repeated twice to detect measurement errors. This measurement removes the X stabilizer on the green face, and only the Z stabilizer is left (colored orange). This completes the code conversion, and the magic state is now encoded in the rotated surface code. A physical depth of 42 is used to complete the code conversion, including the magic state distillation part on the Steane code. Throughout this process, the Pauli frame should be updated based on the measurement outcomes as usual.

V. NUMERICAL SIMULATION

We perform numerical simulations to verify the fault tolerance of zero-level distillation and to estimate the resulting logical error rate and success rate. Since our protocol includes non-Clifford gates, we employ full-vector simulation using Qulacs [28]. While the original circuits require approximately 40 qubits for the rotated surface code (50 for the planar surface code), we perform numerical simulations using only approximately 20 qubits by reusing the ancilla qubits without changing the structure of the original circuit [29]. In the case of the code conversion, 15 qubits are enough for the entire simulation.

Regarding the noise model, we employ a standard depolarizing noise model; each single-qubit and two-qubit gate is followed by single-qubit and two-qubit depolarizing noise with a probability p , respectively, where an idling process is also regarded as an identity gate and hence is followed by noise. Note that while this noise model has been widely used to evaluate the performance of existing fault-tolerant schemes, there is another noise model that is more realistic for superconducting qubit hardware such as SI1000 [30]. In such a noise model, it would be better to implement the qubit transport with a direct swap operation to avoid measurements as mentioned before. Also, we expect that applying SI1000 would lead to an improved logical error rate, since the zero-level distillation circuit includes quite a few idling operations, which have less impact with such an error model with much smaller idling noise.

In order to estimate the logical error rate p_L , the obtained magic state on the rotated surface code is virtually projected to the code space, and fidelity between the obtained magic state and the ideal magic state is calculated. If the fidelity is not a unity, we count it as a logical error. The reason why we project the state to the ideal code space is to estimate the logical error rate by excluding potentially detectable errors. These detectable errors can be

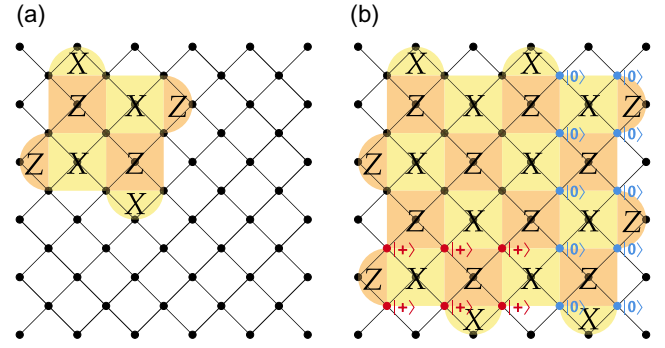


FIG. 9. An expansion of the code distance from $d=3$ to $d=5$ in the rotated surface code. (a) Before expansion. (b) The output magic state on the $d=3$ surface code is located in the corner and other data qubits are initialized to $|0\rangle$ or $|+\rangle$ ancilla qubits. Then syndrome measurements are done, where error correction is performed without postselection.

postselected in the further process in the leading order by sacrificing a small amount of success rate.

The logical error rate p_L is estimated for various physical error rates p , where the numbers of samples is chosen to be between 10^6 and 10^7 . The resultant logical error rate is shown in Fig. 11 for the cases of teleportation-based approaches for the planar (red) and rotated (blue) surface codes, and the code-conversion approach (purple). We can see that the logical error rate scales as $p_L = a \times p^2$ and the coefficient is 93.4 for the planar surface code and 106 for the rotated surface code. In the case of the code conversion, the coefficient is given by 199. Compared to the rotated surface code, the planar surface code has a slightly lower logical error rate due to less shuttling of qubits. The intersection of $p_L \simeq 100p^2$ and $p_L = p$ is located around $p = 10^{-2}$, which indicates that the logical error rate is improved by 1 order of magnitude at $p = 10^{-3}$ and by 2 orders of magnitude at $p = 10^{-4}$. Compared to the teleportation-based approach, the code conversion provides a slightly higher logical error rate because of its large circuit depth.

The success rate of distillation is plotted as a function of the physical error rate p in Fig. 12 for the teleportation-based approaches of the planar surface code (red) and the rotated surface code (blue), and the code conversion (purple). The success rate is 70% when $p = 10^{-3}$ and 95% when $p = 10^{-4}$, indicating that distillation succeeds with a high probability. Compared to the planar surface code, the rotated surface code has a slightly higher success rate since it uses fewer physical qubits. The success rate of the code conversion is almost the same as that of the planar surface code since it uses fewer physical qubits and gates.

We also perform numerical simulations to verify that surface-code expansion works fault tolerantly and to estimate the logical error rate and success rate in such a case. For this purpose, a full-vector simulation using Qulacs cannot be used because of the large number of qubits. Instead,

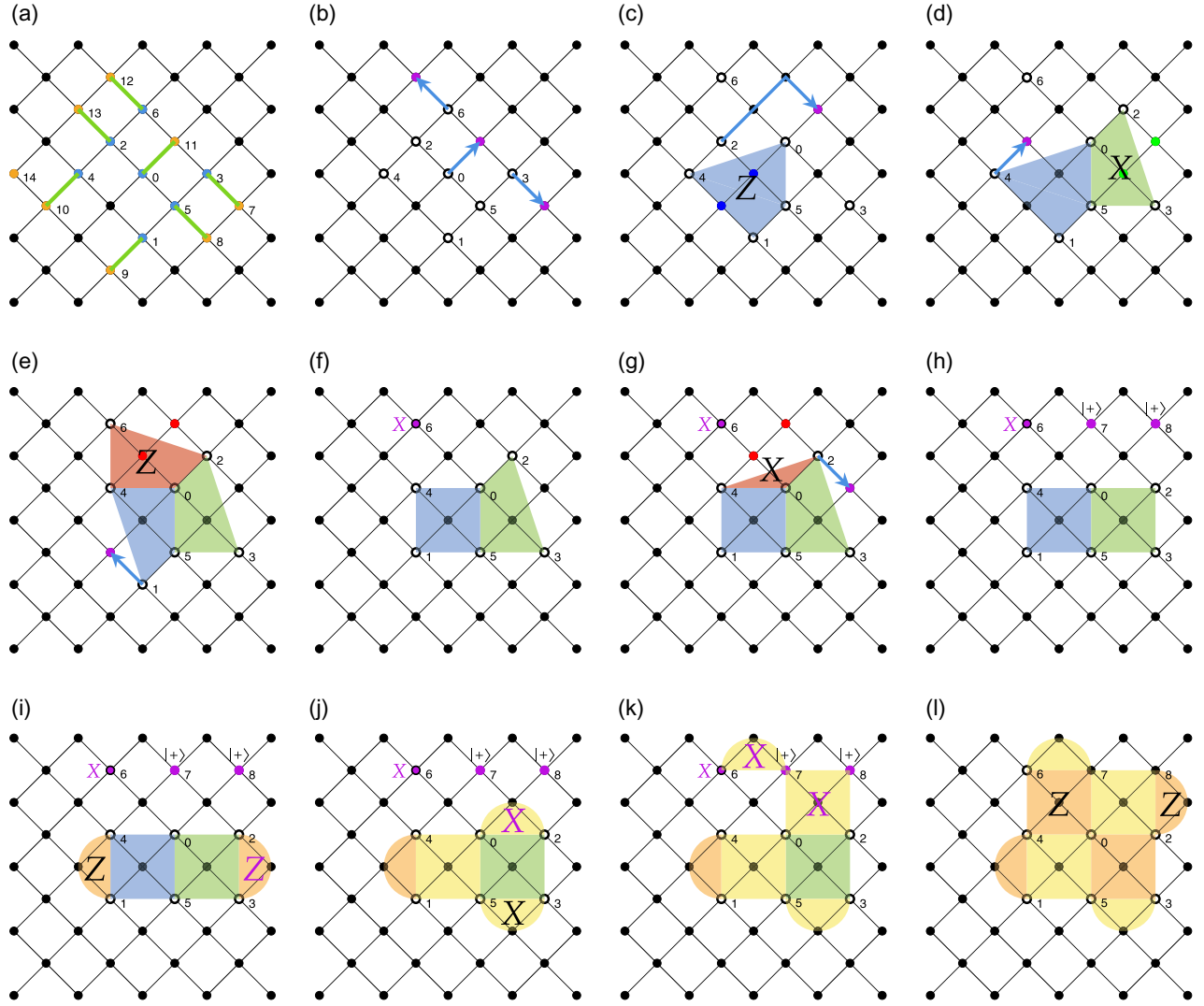


FIG. 10. Qubit arrangement and measurements during conversion from the Steane code to the rotated surface code. The blue and orange dots and numbers correspond to the locations of the qubits shown in Fig. 3. The blue arrows indicate the transfer of qubits using one-bit teleportations. The red, blue, and green faces indicate the stabilizers of the Steane code, and the orange and yellow faces indicate the stabilizers of the rotated surface code. The red, blue, and green dots are the ancilla qubits for the stabilizer measurements of the Steane code.

we use the stabilizer simulator Stim [31]. We replace all $e^{-i\frac{\pi}{8}Y}$ gates with $e^{-i\frac{\pi}{4}Y}$ gates and perform $e^{-i\frac{\pi}{4}Y}$ distillation instead of $e^{-i\frac{\pi}{8}Y}$ distillation. Since this change does not affect the circuit configuration, it is reasonable to estimate the logical error rate. In order to verify this, we compare the results obtained by the full vector simulation with Qulacs and Clifford simulation with Stim for the $d = 3$ case without the expansion. Figure 13 shows the difference in the logical error rate between two simulation approaches for $e^{-i\frac{\pi}{4}Y}$ and $e^{-i\frac{\pi}{8}Y}$ distillations. We find that replacing $\pi/8$ with $\pi/4$ slightly decreases the logical error rate, but the success rate results in the same, as shown in Fig. 14.

Figure 15 shows the results of the logical error rate when the code distance of an output state is extended. The cyan

dots indicate the logical error rate when the code distance is not extended, and other points (blue, purple, orange, red) indicate the data with the code distances extended from $d = 3$ to $d = 5$ or $d = 7$. The cyan, blue, and purple points are calculated using error detection when projecting to the code space. On the other hand, the orange and red points are calculated using error correction. Figure 16 shows the results of the success rate when the code distance of an output state is extended. As Fig. 15, five data points are plotted for each code distance and projection mode. These results clearly demonstrate that the proposed method maintains a low logical error rate while slightly reducing the success rate, even when the surface code is expanded to mitigate the accumulation of logical errors after zero-level distillation.

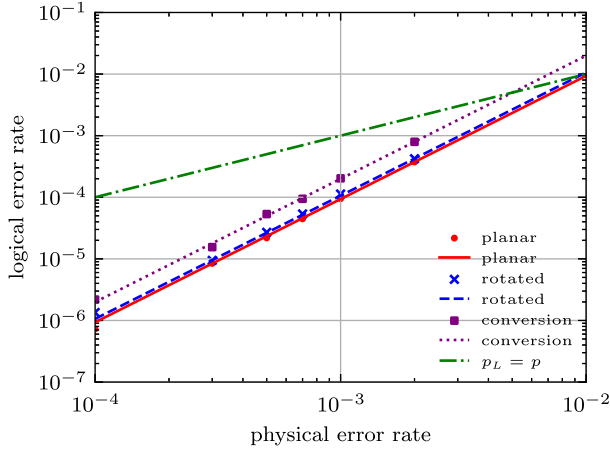


FIG. 11. The logical error rate p_L plotted as a function of the physical error rate p for the teleportation-based approaches of the planar (red circle) and rotated (blue cross) surface code, and the code conversion (purple square). The red, blue, and purple lines scale quadratically. The green line indicates $p_L = p$.

VI. QUALITATIVE IMPLICATIONS OF ZERO-LEVEL DISTILLATION

The potential implications of zero-level distillation include both early-FTQC [23] and full-fledged FTQC. In early-FTQC, the limited availability of physical qubits prevents the execution of conventional multilevel distillation protocols [24]. Therefore, zero-level distillation, which operates with a very small number of qubits, spatial overhead of almost one logical qubit, is well suited for early FTQC, despite its worse scaling $100p^2$ compared to $35p^3$ of conventional methods. Zero-level distillation can generate magic states with a logical error rate of 10^{-6} under the physical error-rate assumption of 10^{-4} , which would be

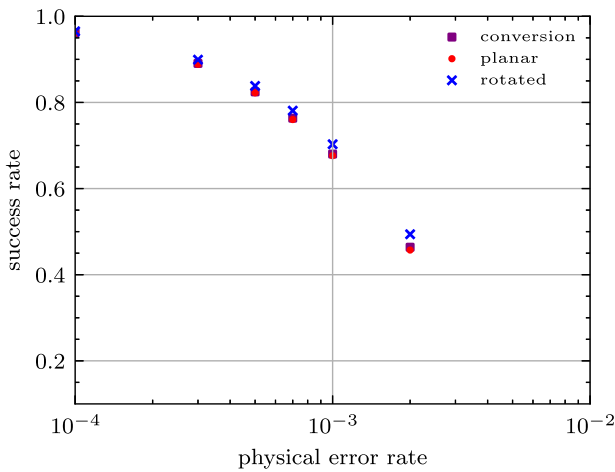


FIG. 12. The success rate plotted as a function of the physical error rate p for the teleportation-based approaches of the planar (red circle) and rotated (blue cross) surface code, and the code conversion (purple square).

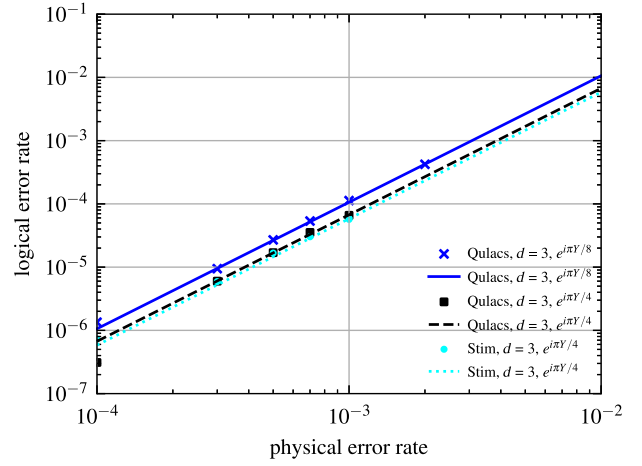


FIG. 13. The logical error rate p_L plotted as a function of the physical error rate p for the $e^{-i\frac{\pi}{8}Y}$ distillation in Qulacs (blue cross), the $e^{-i\frac{\pi}{4}Y}$ distillation in Qulacs (black square), and the $e^{-i\frac{\pi}{4}Y}$ distillation in Stim (cyan dot). The blue, black, and cyan lines scale quadratically.

sufficient to achieve Megaquop [32] in early FTQC. When applying continuous rotation gates with an approximation accuracy of 10^{-6} , roughly speaking, a few tens of T gates are required [33]. By employing zero-level distillation, it becomes possible to perform around a few 10^4 continuous rotation gate operations with fully protected Clifford gates, thereby expanding the range of algorithms beyond NISQ.

Regarding the scalability to larger code distances for full-fledged FTQC applications, two studies, (0+1)-level distillation [34] and magic state cultivation [35], have already been discussed since the first version of this paper appeared. (0+1)-level distillation utilizes the magic

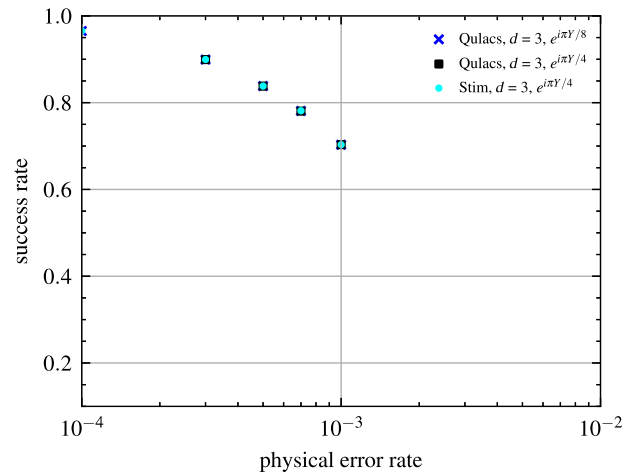


FIG. 14. The success rate plotted as a function of the physical error rate p for the $e^{-i\frac{\pi}{8}Y}$ distillation in Qulacs (blue cross), the $e^{-i\frac{\pi}{4}Y}$ distillation in Qulacs (black square), and the $e^{-i\frac{\pi}{4}Y}$ distillation in Stim (cyan dot).

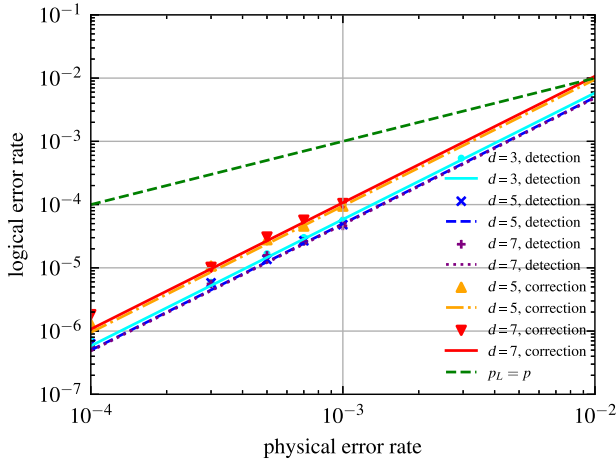


FIG. 15. The logical error rate p_L plotted as a function of the physical error rate p . Five data are plotted for each code distance and projection mode. The detection and correction modes mean performing error detection and correction, respectively, when expanding the code. Cyan, blue, purple, orange, and red lines scale quadratically. The green line indicates $p_L = p$.

states generated by zero-level distillation as input states for conventional level-1 distillation protocols. Combining zero-level distillation, whose logical error rate scales as $O(p^2)$, with conventional methods scaling as $O(p^3)$ results in an overall scaling of $O(p^6)$. Zero-level distillation requires almost no additional qubits compared to conventional methods, resulting in a minimal increase in spatial overhead. Notably, Ref. [34] has demonstrated an improvement in the logical error rate by up to 6 orders of magnitude using this method, while consuming the same spatiotemporal overhead 2.5×10^5 qubitcycles. Additionally, it is shown that the spatiotemporal overhead required

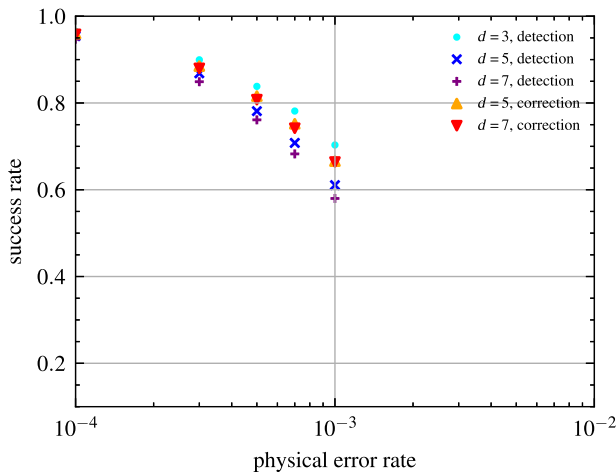


FIG. 16. The success rate plotted as a function of the physical error rate p for each code distance and projection mode. The detection and correction modes mean performing error detection and correction, respectively, when expanding the code.

to achieve a logical error rate of 10^{-16} can be reduced to approximately one third [34].

On the other hand, magic state cultivation is inspired by the concept of zero-level distillation as mentioned in Ref. [35]. It employs larger color codes to perform physical-level distillation even under connectivity constraints. This method successfully suppresses the logical error rate to $O(p^5)$ by introducing various innovative techniques, such as the “double-check” method. As a result, it can generate magic states with a given logical error rate while reducing the spacetime overhead by 2 orders of magnitude [35].

VII. CONCLUSION

We proposed zero-level distillation, which efficiently distills and prepares the logical magic state encoded in surface codes without requiring multiple logical qubits. All operations required can be implemented on the square lattice connectivity and the number of required physical qubits is substantially reduced and spatial overhead for one or two logical patches is sufficient.

According to the numerical simulation, zero-level distillation with teleportation successfully reduces the logical error rate p_L of a logical magic state to $p_L = 100 \times p^2$. For example, when $p = 10^{-3}$ and $p_L = 10^{-4}$, the logical error rates result in $p = 10^{-4}$ and $p_L = 10^{-6}$, respectively, indicating 1 and 2 orders of magnitude improvement. The success rate is reasonably high even when $p = 10^{-3}$. The depth of the zero-level distillation circuit is only 25, hence it is compatible with the conventional multilevel distillation routines [24]. In addition, we also developed zero-level distillation based on the code conversion to further reduce the number of physical qubits employed.

ACKNOWLEDGMENTS

This work is supported by MEXT Quantum Leap Flagship Program (MEXT Q-LEAP) Grant No. JPMXS0120319794, JST COI-NEXT Grant No. JPMJPF2014, and JST Moonshot R&D Grant No. JPMJMS2061.

DATA AVAILABILITY

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

APPENDIX: ZERO-LEVEL DISTILLATION FOR THE PLANAR SURFACE CODE

The magic state distilled on the Steane code can also be teleported to the planar surface code as well as the rotated surface code. In this case, the $|+\rangle_L$ state encoded with the

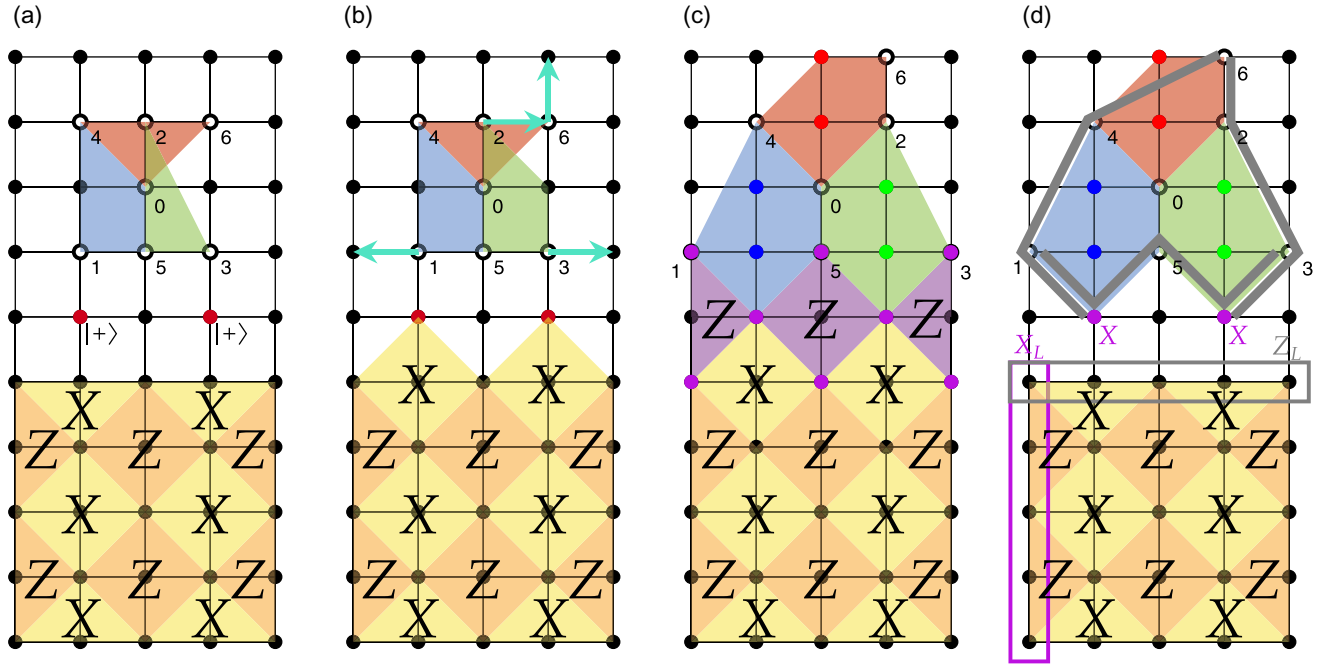


FIG. 17. Qubits arrangement and measurements during teleportation from the Steane code to the planar surface code. The white dots and numbers correspond to the locations of the qubits shown in Fig. 3, where the distilled magic state is encoded in the white qubits. The light blue arrows indicate the transfer of qubits using one-bit teleportations. The red, blue, and green faces indicate the stabilizers of the Steane code, and the orange and yellow faces indicate the stabilizers of the planar surface code. The red, blue, and green dots are the ancilla qubits for the stabilizer measurements of the Steane code. The purple faces indicate the Z-stabilizers at the boundary for the lattice surgery. The gray lines indicate logical operators of the Steane code. The purple box and the gray box indicate logical operators of the planar surface code.

planar surface code and $|+\rangle$ ancilla physical qubits are prepared in parallel with the Hadamard test as shown in Fig. 17(a). After the Hadamard test, some data qubits are moved as shown in Fig. 17(b). Then, as shown by the purple area of Fig. 17(c), $L_Z^{\text{Steane}} \otimes L_Z^{\text{surface}}$ is measured as a lattice surgery. During this lattice surgery, the Z stabilizers have to be measured twice to detect measurement errors. As shown by the purple dots in Fig. 17(d), ancilla qubits on the boundary are measured in the X basis, which breaks the Z stabilizers on the boundary, and two code states are split.

As shown in Fig. 17(d), the Z stabilizers are measured on both Steane and planar surface codes. Finally, the qubits on the Steane code are directly measured in the X basis to calculate the parities of X stabilizers and logical operators. Throughout all operations from the beginning, a physical depth of 23 is used. Note that the Pauli frame has to be updated appropriately throughout these measurements as usual.

[1] Peter W. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, *SIAM J. Comput.* **26**, 1484 (1997).

- [2] Aram W. Harrow, Avinatan Hassidim, and Seth Lloyd, Quantum algorithm for linear systems of equations, *Phys. Rev. Lett.* **103**, 150502 (2009).
- [3] Alán Aspuru-Guzik, Anthony D. Dutoi, Peter J. Love, and Martin Head-Gordon, Simulated quantum computation of molecular energies, *Science* **309**, 1704 (2005).
- [4] A. Morvan *et al.*, Phase transition in random circuit sampling, (2023).
- [5] Youngseok Kim, Andrew Eddins, Sajant Anand, Ken Xuan Wei, Ewout van den Berg, Sami Rosenblatt, Hasan Nayfeh, Yantao Wu, Michael Zaletel, Kristan Temme, and Abhinav Kandala, Evidence for the utility of quantum computing before fault tolerance, *Nature* **618**, 500 (2023).
- [6] John Preskill, Quantum computing in the NISQ era and beyond, *Quantum* **2**, 79 (2018).
- [7] M. Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C. Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R. McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, and Patrick J. Coles, Variational quantum algorithms, *Nat. Rev. Phys.* **3**, 625 (2021).
- [8] Peter W. Shor, Scheme for reducing decoherence in quantum computer memory, *Phys. Rev. A* **52**, R2493 (1995).
- [9] Keisuke Fujii, Quantum computation with topological codes: From qubit to topological fault-tolerance, (2015).

- [10] A. Yu. Kitaev, Fault-tolerant quantum computation by anyons, *Ann. Phys. (N. Y)* **303**, 2 (2003).
- [11] Sergey Bravyi and Robert Raussendorf, Measurement-based quantum computation with the toric code states, *Phys. Rev. A* **76**, 022304 (2007).
- [12] Austin G. Fowler, Matteo Mariantoni, John M. Martinis, and Andrew N. Cleland, Surface codes: Towards practical large-scale quantum computation, *Phys. Rev. A* **86**, 032324 (2012).
- [13] Robert Raussendorf and Hans J. Briegel, A one-way quantum computer, *Phys. Rev. Lett.* **86**, 5188 (2001).
- [14] Emanuel Knill and Bryan Eastin, Restrictions on transversal encoded quantum gate sets, *Phys. Rev. Lett.* **102**, 110502 (2009).
- [15] Sergey Bravyi and Alexei Kitaev, Universal quantum computation with ideal Clifford gates and noisy ancillas, *Phys. Rev. A* **71**, 022316 (2005).
- [16] Daniel Gottesman and Isaac L. Chuang, Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations, *Nature* **402**, 390 (1999).
- [17] Craig Gidney and Martin Ekerå, How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits, *Quantum* **5**, 433 (2021).
- [18] Hayato Goto, Minimizing resource overheads for fault-tolerant preparation of encoded states of the steane code, *Sci. Rep.* **6**, 19578 (2016).
- [19] Christopher Chamberland and Andrew W. Cross, Fault-tolerant magic state preparation with flag qubits, *Quantum* **3**, 143 (2019).
- [20] Christopher Chamberland and Kyungjoo Noh, Very low overhead fault-tolerant magic state preparation using redundant ancilla encoding and flag qubits, *Npj Quantum Inf.* **6**, 91 (2020).
- [21] Lukas Postler, Sascha Heußen, Ivan Pogorelov, Manuel Risper, Thomas Feldker, Michael Meth, Christian D Marciniak, Roman Stricker, Martin Ringbauer, and Rainer Blatt *et al.*, Demonstration of fault-tolerant universal quantum gate operations, *Nature* **605**, 675 (2022).
- [22] Emanuel Knill, Raymond Laflamme, and Wojciech H. Zurek, Resilient quantum computation, *Science* **279**, 342 (1998).
- [23] Yutaro Akahoshi, Kazunori Maruyama, Hirotaka Oshima, Shintaro Sato, and Keisuke Fujii, Partially fault-tolerant quantum computing architecture with error-corrected Clifford gates and space-time efficient analog rotations, (2023).
- [24] Daniel Litinski, Magic state distillation: Not as costly as you think, *Quantum* **3**, 205 (2019).
- [25] Emanuel Knill, Raymond Laflamme, and Wojciech H. Zurek, Resilient quantum computation: Error models and thresholds, *Proc. R. Soc. London. Ser. A: Math. Phys. Eng. Sci.* **454**, 365 (1998).
- [26] Hendrik Poulsen Nautrup, Nicolai Friis, and Hans J. Briegel, Fault-tolerant interface between quantum memories and quantum processors, *Nat. Commun.* **8**, 1321 (2017).
- [27] Lingling Lao and Carmen G. Almudever, Fault-tolerant quantum error correction on near-term quantum processors using flag and bridge qubits, *Phys. Rev. A* **101**, 032333 (2020).
- [28] Yasunari Suzuki, Yoshiaki Kawase, Yuya Masumura, Yuria Hiraga, Masahiro Nakadai, Jiabao Chen, Ken M. Nakanishi, Kosuke Mitarai, Ryosuke Imai, Shiro Tamiya, Takahiro Yamamoto, Tennin Yan, Toru Kawakubo, Yuya O. Nakagawa, Yohei Ibe, Youyuan Zhang, Hirotsugu Yamashita, Hikaru Yoshimura, Akihiro Hayashi, and Keisuke Fujii, Qulacs: A fast and versatile quantum circuit simulator for research purpose, *Quantum* **5**, 559 (2021).
- [29] Mitsuki Katsuda, Kosuke Mitarai, and Keisuke Fujii, Simulation and performance analysis of quantum error correction with a rotated surface code under a realistic noise model, *Phys. Rev. Res.* **6**, 013024 (2024).
- [30] Craig Gidney, Michael Newman, Austin Fowler, and Michael Broughton, A fault-tolerant honeycomb memory, *Quantum* **5**, 605 (2021).
- [31] Craig Gidney, Stim: a fast stabilizer circuit simulator, *Quantum* **5**, 497 (2021).
- [32] John Preskill, Beyond NISQ: The megasclop machine, *ACM Trans. Quantum Comput.* **6**, 18 (2025).
- [33] Neil J. Ross and Peter Selinger, Optimal ancilla-free Clifford+T approximation of z-rotations, *Quantum Info. Comput.* **16**, 901 (2016).
- [34] Yutaka Hirano, Tomohiro Itogawa, and Keisuke Fujii, in *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)*, (IEEE, Montréal, 2024), Vol. 1, p. 843.
- [35] Craig Gidney, Noah Shutty, and Cody Jones, Magic state cultivation: Growing T states as cheap as CNOT gates, [arXiv:2409.17595](https://arxiv.org/abs/2409.17595).