Loss-tolerant linear optical quantum computing under nonideal fusions using multiphoton qubits

Seok-Hyung Lee,1 Srikrishna Omkar,2 Yong Siah Teo,1 and Hyunseok Jeong1

1 Department of Physics and Astronomy, Seoul National University, Seoul 08826, Republic of Korea
2 ORCA Computing, Toronto M6P3T1, Canada

Cluster states, which are entangled states used for measurement-based linear optical quantum computing (LOQC), are typically constructed by combining many small resource states through fusions. However, such fusions are never ideal because of theoretical limitations as well as environmental factors such as photon losses. We propose a methodology for handling nonideal fusions, which enables more realistic error simulations. Accordingly, using only single-photon polarization qubits, we show that an excessively high fusion success rate is required to implement the scheme in linear optical systems. We then propose an alternative scheme using the parity state encoding, where loss thresholds over 2% are achievable using about 390 three-qubit Greenberger-Horne-Zeilinger states per data qubit and high thresholds near 10% can be reached if they are sufficiently supplied. Hence, this scheme is a competitive candidate for near-term universal fault-tolerant LOQC.

I. INTRODUCTION

Optical qubits with dual-rail encoding are a promising candidate for quantum computing (QC) with the advantages of long decoherence time, ease of single-qubit manipulation via linear optical circuits, and ease of detecting photon losses by counting photon numbers [1]. A representative way to achieve universal QC in linear optical systems is measurement-based quantum computing (MBQC) [2, 3] processed by single-qubit measurements on a multi-qubit cluster state. In particular, a family of cluster states called Raussendorf-Harrington-Goyal (RHG) lattices [4–6] permits universal fault-tolerant quantum computing [7, 8], although magic state distillation [10] are required for non-Clifford gates.

The generation of RHG lattices, which is a major challenge to realize fault-tolerant MBQC, can be done by entangling multiple small resource states with fusions of types I and II [11]. Both types of fusions are not ideal in linear optical systems because of theoretical limitations as well as environmental factors such as photon losses. In general, fault-tolerant fusion failure rates should be less than ∼14.5% < 50% if the adaptive scheme in Ref. [12] is used, which is only possible with additional resources [13]. There exist several types of approaches to overcome this shortcoming. Some examples are (i) different types of encoding strategies with coherent states [14, 16], hybrid qubits [15, 17], and multiphoton qubits [18, 19] that significantly improve error thresholds and resource overheads [19], (ii) ancillary photons to boost the success rate of a type-II fusion to 75% [20, 21], which enables MBQC with the renormalization method [22], (iii) redundant structures added to resource states to replace a single fusion by multiple fusion attempts [23], and (iv) use of squeezing for teleportation channels [24] or inline-processes [25, 26].

Previous studies frequently treated fusion failure with edge removals [27, 30] or edge-connected qubit removals [12, 16, 19, 22]. However, the effect of failed fusions on the remaining state should be handled more carefully for a strictly realistic analysis. In this Letter, we study how failed fusions corrupt stabilizers and how errors arising from such corruption are propagated during the generation of cluster states. It is shown that this effect is equivalent to assigning error rates on some qubits in the final lattice. It is important that such information is heralded by the fusion outcomes, which not only enables much more realistic error simulations, but is also useful for decoding syndromes.

Based on the proposed methodology, we investigate the fault-tolerance of two MBQC schemes. We first consider using only single-photon polarization qubits with fusions assisted by ancillary photons. We estimate loss thresholds for different fusion failure rates and show that the maximum allowable fusion failure rate is too low for feasible linear-optical implementation; that is, precise high-resolution photon-number resolving detectors (PNRDs) and ancillary states that are hard to generate with linear optics are required. We then propose an alternative scheme using the parity state encoding [31] and concatenated Bell-state measurement [32]. The scheme requires only on-off detectors or PNRDs resolving up to two photons and all resource states can be generated with linear optics. We numerically calculate its loss thresholds, which shows that the scheme is a competitive candidate for near-term universal fault-tolerant linear-optical QC.

II. PRELIMINARIES

We denote the four Bell states by $|\Phi^{\pm}\rangle := |0\rangle|0\rangle \pm |1\rangle|1\rangle$ and $|\Psi^{\pm}\rangle := |0\rangle|1\rangle \pm |1\rangle|0\rangle$ (normalization coefficients are omitted unless otherwise necessary) and call “$\pm$” its sign and “$\Phi$” or “$\Psi$” its letter. An ideal Bell-state measurement (BSM) entails the measurements of $X \otimes X$ and $Z \otimes Z$ on two qubits, whose outcomes are
addressed as its sign and letter outcomes, respectively. We use the polarization of photons as the degree of freedom for dual-rail encoding and denote the horizontally (vertically) polarized state by $|\text{H (V)}\rangle$.

For a given graph $G = (V, E)$ where a qubit is placed on each vertex, a cluster state $|G\rangle$ is defined as the state stabilized by $S_v := X_v \prod_{v' \in N(v)} Z_{v'}$ (that is, $S_v |G\rangle = |G\rangle$) for each vertex $v$, where $X_v$ and $Z_v$ are respectively Pauli-X and Z operators on the qubit at $v$ and $N(v)$ is the set of the vertices connected with $v$. Among the two types of fusions, we only consider type II since type I may convert photon losses into unheralded computational errors [23]. A type-II fusion is done by measuring $X_{v_1} Z_{v_2}$ and $Z_{v_1} X_{v_2}$ for a pair of qubits $(v_1, v_2)$ satisfying that $\{v_1\} \cup N(v_1)$ and $\{v_2\} \cup N(v_2)$ are disjoint and neither $N(v_1)$ nor $N(v_2)$ is empty. In practice, this operation is realized by applying the Hadamard ($H$) gate on either $v_1$ or $v_2$ and then performing a BSM on them. The effect of a type-II fusion is to connect (disconnect) every possible pair of disconnected (connected) qubits, one from $N(v_1)$ and the other from $N(v_2)$, and apply several Pauli-Z operators determined by the letter ($m_{\text{lett}}$) and sign ($m_{\text{sign}}$) outcomes of the BSM. These Pauli-Z operators are compensated by updating the Pauli frame [33] with a classical computer. This effect can be checked by tracking the transformation of stabilizers. An example of a fusion is shown in Fig. 2 with the transformation of two stabilizers.

We consider errors of qubits in the “vacuum” [4] that are measured in the $X$-basis, thus $X$-errors do not affect the results. Henceforth, every error mentioned from now on is a $Z$-error.

III. NONIDEAL FUSIONS

Let us consider the example of Fig. 1. If the qubits are single-photon polarization ones, the BSM can discriminate only two Bell states (say, $|\Phi^+\rangle$) among the four without additional resources. The intact final state is obtained only when the BSM succeeds. When the BSM fails (which is heralded), $m_{\text{lett}}$ is determined while $m_{\text{sign}}$ is left ambiguous. Since the marginal state of qubits 0 and 0′ is maximally mixed before the fusion (see Appendix A), we simply randomly assign the value of $m_{\text{sign}}$. Then, the operator $m_{\text{sign}} X_1 Z_1 Z_2$, which is a stabilizer of the intact final state, gives $\pm 1$ randomly when measured. Whereas, the other two stabilizers $m_{\text{lett}} Z_1 X_1$ and $m_{\text{lett}} Z_1 X_2$ are left undamaged. The key point is that this scenario is equivalent to a 50% chance of an erroneous qubit 1 in the intact final state. Both interpretations give the same statistics if the stabilizers of the intact final state are measured, thus every process in MBQC, which is described with the stabilizer formalism, works in the same way in both of them.

In general, a nonideal BSM gives one of multiple possible outcomes (for the above case, say, “$\Phi^+$,” “$\Phi^-$,” and “failure”), and one of the four Bell states that gives the highest posterior probability under the outcome is selected, assuming that their prior probabilities are equal. From the posterior probabilities, the probability $q_{\text{sign}}$ ($q_{\text{lett}}$) that the selected sign (letter) is wrong can be obtained. Then, in the case of Fig. 1, this is equivalent to qubit 1 having an error with probability $q_{\text{sign}}$ and qubits 1′ and 2′ having correlated errors with probability $q_{\text{lett}}$ in the intact final state. We term a qubit with a nonzero error rate “deficient.”

Additionally, if a qubit participated in a fusion has an error, this error propagates to the qubits in the “opposite” side. For example, an erroneous qubit 0 induces an error in the $X_0 Z_0$ measurement, which is equivalent to erroneous qubits 1′ and 2′.

IV. BUILDING A LATTICE

An RHG lattice is built with linear three-qubit cluster states called microclusters [27]. The process is composed of two steps: In step 1, a central microcluster and two side microclusters are merged by two fusions to form a five-qubit cluster state called a star cluster, as shown in Fig. 2(a). The qubit with degree four in a star cluster is called a central qubit and the other four qubits are called side qubits. In step 2, the side qubits of star clusters are fused with each other, as shown in Fig. 2(b). Eventually, only the central qubits constitute the lattice and they are measured in appropriate bases.

The locations of the $H$ gates in fusions (called $H$-configuration) may be chosen arbitrarily, but we here consider two specific configurations and call them “HIC” and “HIS.” In the HIS (HIS) configuration, the $H$ gates in step 1 are applied on the qubits in central (side) mi-
Central qubit. For the HIS configuration, propagated to the nearby central qubits after step 2. A side qubits deficient and the errors in the side qubits are used. A step-1 fusion may make one central qubit or two ideal step-1 and 2 fusions when the HIC configuration is the processes that central qubits become deficient by non-

qubits deficient as described previously. Figure 3 shows rations [see Fig. 2(b)].

FIG. 2. (a) Steps 1 and (b) 2 of the process to build an RHG lattice through fusions (orange squares). In (a), center and side microclusters are fused to form star clusters. In (b), star clusters are fused to form an RHG lattice. The “C” (“S”) marks in (a) and the red dots in (b) indicate the qubits on which the $H$ gates are applied in the HIC (HIS) configuration.

Nonideal fusions during lattice building render some qubits deficient as described previously. Figure 3 shows the processes that central qubits become deficient by non-

ideal step-1 and 2 fusions when the HIC configuration is used. A step-1 fusion may make one central qubit or two side qubits deficient and the errors in the side qubits are propagated to the nearby central qubits after step 2. A step-2 fusion may make one or two central qubits deficient. For the HIS configuration, $q_{\text{sign}}$ and $q_{\text{lett}}$ should be swapped accordingly. Note that the sign and letter errors of a fusion always affect different types (“primal” or “dual” [4]) of qubits. Thus, the correlation between these two errors, if any, can be neglected if the primal and dual lattices are considered separately.

V. NOISE MODEL

We suppose a noise model where each photon suffers an independent loss with probability $\eta$, which arises from imperfections throughout the protocol: three-photon Greenberger-Horne-Zeilinger (3-GHZ) states (which are initial resource states), delay lines, optical switches, and photodetectors. Indeed, photon-loss rates may vary from qubit to qubit, thus $\eta$ can be regarded as a maximal loss rate in the scheme. We further assume that noises that cannot be modelled with photon losses such as dark counts are negligible.

Note that both nonideal fusions and photon losses in central qubits are sources of deficiency. We use single-

photon polarization encoding for central qubits in both the scheme discussed below. They are measured in Pauli bases by using polarizing beam splitters and on-off detectors (except singular qubits for state injection [11], which is out of the scope of this Letter). When a photon loss is detected, we select the measurement outcome randomly and regard that it has a 50% chance of having an error.

VI. SCHEME WITHOUT LOGICAL ENCODING

We first investigate using single-photon polarization encoding for all qubits with fusions assisted by ancillary photons. A fusion detects a loss with probability $1 - (1 - \eta)^2$, where $q_{\text{lett}} = q_{\text{sign}} = 1/2$. If otherwise, it fails with probability $p_f$, where $q_{\text{lett}} = 0$ and $q_{\text{sign}} = 1/2$. These two cases make some central qubits deficient, which can be tracked by the methodology proposed above. The HIC configuration is used to make the failure of a step-1 fusion affect only one central qubit [see Fig. 3(a)].

3-GHZ states are used as basic building blocks to construct a cluster state. Two schemes are considered: All generated star clusters are used for step 2 or star clusters where all the step-1 fusions succeed are post-selected. The former scheme is more resource-efficient than the latter: Three 3-GHZ states are required per central qubit without post-selection, while averagely $2[4+(1-\eta)^2]/(1-\eta)^2 \approx 10$ of them are required otherwise. Moreover, the latter requires additional switches.

To obtain the loss thresholds of the schemes, we simulate the logical identity gate with the length of $4d+1$ unit cells along the simulating time axis, where $d$ is the code distance. For decoding syndromes, we use the weighted minimum-weight perfect matching (MWP) decoder via PyMatching package [34] to exploit the error rates of individual qubits. The loss thresholds are calculated by...
finding the intersections of logical error rates for $d = 9$ and $d = 11$. See Appendix C for more details.

The obtained loss thresholds for various fusion failure rates are presented in Fig. 4. Theoretical estimations using the percolation threshold [33,34] are also presented; see Appendix B. It shows that, even if $\eta$ is only 1%, $p_f$ should be less than about 10%. To achieve such small fusion failure rates, almost accurate PNRDs resolving up to 16 photons and ancillary multi-photon states hard to generate with linear optics are required [21]. Moreover, the above simulation does not consider the imperfection of ancillary states and additional PNRDs; if they are considered, the loss thresholds may be even lower.

Note that the obtained threshold of $p_f$ is about 13% for the scheme with post-selection, which is slightly lower than the value 14.5% of the adaptive scheme in Ref. [12]. However, our scheme has an advantage that the measurement bases of central qubits do not need to be selected in an adaptive manner.

VII. SCHEME WITH THE PARITY STATE CODES

We now introduce an alternative scheme: Fusion success rates are boosted by encoding side qubits with the parity state codes [31] and employing the concatenated BSM (CBSM) scheme [32]. On-off detectors or PNRDs resolving up to two photons are enough for the scheme and all resource states can be generated linear-optically. This scheme resolves the crucial shortcomings of the aforementioned scheme at the cost of increasing resource overheads.

The $(n,m)$ parity state encoding defines a basis as $|0_L\rangle := |+(m)\rangle^\otimes n$ and $|1_L\rangle := |-(m)\rangle^\otimes n$, where $|\pm(m)\rangle := (|\uparrow\rangle + |\downarrow\rangle)^\otimes m \pm (|\uparrow\rangle - |\downarrow\rangle)^\otimes m$. The Hilbert space has a hierarchical structure composed of three levels: the lattice level with basis $\{|0_L\rangle, |1_L\rangle\}$, the block level with basis $\{|\pm(m)\rangle\}$, and the physical level with basis $\{|\uparrow\rangle, |\downarrow\rangle\}$. In the CBSM scheme, a BSM of a level is decomposed into multiple BSMs of the level below. Our CBSM scheme is similar with the one in Ref. [32] except the following two points: (i) A physical-level BSM may or may not discriminate a photon loss and a failure. PNRDs resolving up to two photons are required to make such discrimination possible, while on-off detectors are enough for the other case. (ii) The letter outcome of a lattice-level BSM is obtained by a weighted majority vote of the letter outcomes of block-level BSMs. See Appendix C for details on the CBSM scheme and also Appendix D for the calculation of its error rates $(q_{\text{sign}}, q_{\text{lett}})$.

For a practical reason, we consider generating $post-H$ microclusters (that is, the states obtained by applying several lattice-level $H$ gates on microclusters) directly, instead of generating microclusters first and then applying the $H$ gates for the fusions. Each central (side) post-$H$ microcluster can be rewritten as a physical-level cluster state with $2nm + 1$ $(3nm)$ photons up to some physical-level $H$ gates, whose structure depends on the $H$-configuration. It can be generated by fusing multiple 3-GHZ states appropriately. See Appendix F for their physical-level graph structures and the detailed instructions of their generation.

We now evaluate the photon loss threshold and the resource overhead of the scheme. For Monte-Carlo error simulations, we consider various settings of the model on the type of photon detectors (on-off detectors or PNRDs resolving up to two photons), the $H$-configuration, and the post-selection of star clusters. Like the case without logical encoding, we simulate the logical identity gate with the weighted MWPM decoder; see Appendix C. The resource overhead is quantified by the expected number $N_{\text{GHZ}}^*$ of 3-GHZ states required per center qubit. Note that $\sim 6N_{\text{GHZ}}^*T(d - 1)^3$ of 3-GHZ states are needed per logical qubit, where $T$ is the length along the simulated time axis which is typically $O(d)$ for a logical gate. The process to generate $post-H$ microclusters significantly affect $N_{\text{GHZ}}^*$, thus it should be optimized carefully; see Appendix F for details.

Figure 5(a) presents the loss thresholds $\eta_{\text{th}}$ for various parameters: the encoding size $(n,m)$, the type of detectors (PNRD resolving up to two photons or on-off detector), the post-selection (PS) of star clusters, and the $H$-configuration. It shows that loss thresholds near 10% is achievable for suitable values of $n$ and $m$. Note that a large encoding size does not always mean a high threshold since the physical- and block-level repetitions of the encoding respectively suppress one of the sign and letter errors but assists the other.

We also obtain the resource overheads $N_{\text{GHZ}}^*$ for various parameters, which are shown in Fig. 5(b) with the loss thresholds. Since $N_{\text{GHZ}}^*$ depends on the photon loss rate, it is assumed that the scheme is operated at the loss rate of $\eta_{\text{th}}/2$. When PNRDs resolving up to two photons are used, loss thresholds over 2%, 3%, or 4% are possi-
VIII. DISCUSSION

We described how nonideal type-II fusions herald deficiency on qubits in a cluster state used for linear optical measurement-based quantum computing (MBQC). We applied this method on evaluating the loss-tolerance of two computing schemes: using single-photon qubits with fusions assisted by ancillary photons and using logical qubits encoded with the parity-state codes. The former scheme requires relatively few three-photon Greenberger-Horne-Zeilinger (3-GHZ) states (~10 per data qubit if star clusters are post-selected), but precise high-resolution photon-number resolving detectors (PNRDs) and ancillary states hard to generate with linear optics are demanded. On the other hand, although the latter scheme requires many 3-GHZ states (~390 per data qubit to achieve a photon loss threshold of 2%) and
switching circuits, it is fully performed with linear optics and uses only only-off detectors or PNRDs resolving up to two photons. Moreover, this method has the potential to reach very high loss thresholds near 10% as the technology for generating 3-GHZ states evolves. Therefore, it is a competitive candidate for near-term universal fault-tolerant linear optical MBQC.

One may apply our method of handling nonideal fusions on other encoding schemes or decoding methods (such as the union-find decoder [37]) for improvement of fault-tolerance or resource overheads. It will be worth investigating whether the renormalization method in Ref. 22 can lower the required fusion success rate of the scheme without logical encoding despite errors accumulating during the scheme. More comprehensive consideration of component-wise noises including both heralded photon losses and unheralded errors (such as dark counts on photodetectors) may make the simulation results even more realistic and practical. Lastly, our methods may be generalized for other MBQC schemes such as the color-code-based one 38 or the fusion-based quantum computing scheme 39 that is attracting attention recently.

Appendix A: Proof of the statement on marginal states of a cluster state

We here prove the statement: For a cluster state $|G\rangle_V$ with a graph $G = (V, E)$ and given two vertices $a, b \in V$, if $\{a\} \cup N(a)$ and $\{b\} \cup N(b)$ are disjoint and neither $N(a)$ nor $N(b)$ is empty where $N(v)$ for a vertex $v \in V$ is the set of vertices adjacent to $v$, the marginal state $Tr_{\{a,b\}} |G\rangle |G\rangle_V := \rho_{ab}$ is maximally mixed.

Let $S$ denote the stabilizer group of the zero-dimensional Hilbert space $\{|G\rangle_V\}$. First, any stabilizer $S \in S$ can be written as the product of stabilizer generators: $S = \prod_{\forall v \in V} S_v$, where $V_0 \subseteq V$ and $S_v := X_v \prod_{\forall v' \in N(v)} Z_{v'}$. If $V_0$ contains a vertex $c \neq a, b$, $S$ must contain $X_c$ or $Y_c$, since no stabilizer generators besides $S_c$ contain $X_c$. If otherwise, $V_0$ is one of $\emptyset$, $\{a\}$, $\{b\}$, and $\{a, b\}$. Except when $V_0$ is empty (namely, $S$ is identity), there exists a vertex $c \neq a, b$ such that $S$ contains $Z_c$, since $N(a)$ and $N(b)$ are not empty, $b \notin N(a)$, $a \notin N(b)$, and $N(a) \neq N(b)$. Therefore, every single- or two-qubit Pauli operator that is not identity on $a$ and $b$ cannot be a stabilizer, thus it anticommutes with at least one stabilizer. (If such an operator $P_aP_b$ commutes with all stabilizers, $P_aP_b |G\rangle_V$ is also stabilized by $S$, which means that $P_aP_b |G\rangle_V = |G\rangle_V$.) Consequently, $Tr(P_aP_b\rho_{ab}) = \langle G | P_aP_b | G \rangle = 0$ for every single- or two-qubit Pauli operator $P_aP_b$ that is not identity. The state $\rho_{ab}$ satisfying this condition is unique and maximally mixed.

Appendix B: Theoretical estimations of the loss thresholds in the scheme without logical encoding

We here present the method to estimate the loss thresholds theoretically for the scheme without logical encoding. Note that a fusion fails ($q_{\text{sign}} = 1/2$ and $q_{\text{let}} = 0$) with probability $p_l(1-\eta)^2$ and it detects a loss ($q_{\text{sign}} = q_{\text{let}} = 1/2$) with probability $1 - (1-\eta)^2$. We first consider the case that star clusters are not post-selected. For a central qubit $q$ to be not deficient, the following conditions should be satisfied simultaneously:

1. Two step-1 fusions in the star cluster containing $q$ succeed.
2. Four step-1 fusions in the four adjacent star clusters (one for each) do not detect losses.
3. Four step-2 fusions involved in the star cluster containing $q$ do not detect losses. Two among them (that make $q$ deficient if they fail) succeed.
4. $q$ itself does not suffer a loss.

From above, we obtain the probability that a central qubit in the final lattice is intact: $p_{\text{int}}(\eta, p_l) = (1-p_l)^4(1-\eta)^2$. If star clusters are post-selected, the first and second conditions are no longer needed, thus we get $p_{\text{int}}(\eta, p_l) = (1-p_l)^2(1-\eta)^9$. Regarding a 50% chance of a Z-error as erasing the qubit by measuring it in the Z-basis and ignoring correlated errors, a photon loss threshold $p_{\text{th}}$ can be estimated by solving $1 - p_{\text{prc}} = p_{\text{int}}(p_{\text{th}}, p_l)$, where $p_{\text{prc}} = 0.249$ is the known cubic-lattice bond percolation threshold 35 36.

Appendix C: Concatenated BSM scheme of the parity state codes

The $(n, m)$ parity state encoding defines a basis as $|0_L\rangle := |+^{(m)}\rangle_{\otimes n}$ and $|1_L\rangle := |-^{(m)}\rangle_{\otimes n}$, where $|\pm^{(m)}\rangle := (|H\rangle + |V\rangle)^{\otimes m} \pm (|H\rangle - |V\rangle)^{\otimes m}$. The bases of lattice-, block-, and physical-level Hilbert spaces are respectively given as $\{|0_L\rangle, |1_L\rangle\}$, $\{|\pm^{(m)}\rangle\}$, and $\{|H\rangle, |V\rangle\}$. With the above bases, the Bell states are defined as $\{|\Phi^+\rangle, |\Psi^\pm\rangle\}$,
shown in the main text. For example, \(|\psi^\pm\rangle\) if detectors (A, C) or (B, D) detect one photon respectively and \(|\psi^-\rangle\) if detectors (A, D) or (B, C) detect one photon respectively. If otherwise, it fails or detects a loss, which can be discriminated by the total number of detected photons if PNRDs are used. Two distinguishable Bell states can be to perform a BSM in a concatenated manner: A lattice-level BSM (BSM \(P\)) where
\[
\Phi^\pm := |0_L\rangle |0_L\rangle \pm |1_L\rangle |1_L\rangle
\]
and
\[
\Psi^\pm := |0_L\rangle |1_L\rangle \pm |1_L\rangle |0_L\rangle.
\]
The Bell states of each level can be decomposed into those of the lower level as follows:
\[
|\Phi^\pm\rangle = 2^{-\frac{m_1}{2}} \sum_{l: \text{even}(\text{odd})} \mathcal{P} \left[ |\phi^{-\pm}\rangle^{\otimes l} |\phi^{\pm}\rangle^{\otimes n-l} \right],
\]
\[
|\Psi^\pm\rangle = 2^{-\frac{m_2}{2}} \sum_{l: \text{even}(\text{odd})} \mathcal{P} \left[ |\psi^{-\pm}\rangle^{\otimes l} |\phi^{\pm}\rangle^{\otimes n-l} \right],
\]
\[
|\phi^{\pm}\rangle_{(m)} = 2^{-\frac{m}{2}} \sum_{k: \text{even} \leq m} \mathcal{P} \left[ |\phi^{\pm}\rangle^{\otimes k} |\phi^{\pm}\rangle^{\otimes m-k} \right],
\]
\[
|\psi^{\pm}\rangle_{(m)} = 2^{-\frac{m}{2}} \sum_{k: \text{odd} \leq m} \mathcal{P} \left[ |\psi^{\pm}\rangle^{\otimes k} |\phi^{\pm}\rangle^{\otimes m-k} \right],
\]
where \(\mathcal{P}[\cdot]\) means the summation of all the permutations of the tensor products inside the bracket. It makes it possible to perform a BSM in a concatenated manner: A lattice-level BSM (BSM_{lat}) is done by \(n\) block-level BSIs (BSM_{blk}'s), each of which is again done by \(m\) physical-level BSIs (BSM_{phys}'s). We refer to the sign (letter) result obtained from a lattice-, block-, or physical-level BSM as a lattice-, block-, or physical-level sign (letter), respectively.

1. Original scheme

We first review the original concatenated BSM scheme of the parity state codes in Ref. 32. A BSM_{phys} can discriminate only two among the four Bell states. Three types of BSM_{phys}'s (\(B_\psi, B_+, \) and \(B_-\)) are considered, which discriminate \(\{|\psi^+\rangle, |\psi^-\rangle\}, \{\phi^+\rangle, |\psi^+\rangle\}\), and \(\{|\phi^-\rangle, |\psi^-\rangle\}\), respectively. A BSM_{phys} has four possible outcomes: two successful cases (e.g., for \(B_\psi, |\psi^+\rangle\) and \(|\psi^-\rangle\)), “failure,” and “detecting a photon loss.” The BSM scheme for single-photon polarization qubits 40 is presented in Fig. 6. Failure and loss are discriminated by the number of total photons detected by the photon detectors. Since two photons may enter a single detector, it is assumed that PNRDs resolving up to two photons are used. Note that, even in the failure cases, either sign or letter still can be determined definitely. (For example, if a \(B_\psi\) fails, we can learn that the letter is \(\phi\).) On the other hand, if it detects a loss, we cannot get neither sign nor letter. We also note that the failure case contains many combinations of detector outcomes and each of them may individually give some information on the input state. However, we suppose that we forget the exact combination and only know that it fails.

A BSM_{blk} is done by \(m\)-times of BSM_{phys}'s. Each block is composed of \(m\) photons, thus we consider \(m\) pairs of photons selected respectively in the two blocks. First, \(B_\psi\) is performed on each of such pairs in order until it Either succeeds, detects a loss, or consecutively fails \(j\) times, where \(j \leq m - 1\) is a predetermined number. Then a sign \(s = \pm\) is selected by the sign of the last \(B_\psi\) outcome if it succeeds and randomly if it fails or detects a loss. After that, \(B_s\)'s are performed for all the left pairs of photons.
The block-level sign (letter) is determined by the physical-level signs (letters) of the $m$ BSM_{phy}'s. In details, the block-level sign is chosen (a) to be the same as $s$ if the last $B_s$ succeeds or any $B_s$ succeeds, and (b) to be the opposite of $s$ if the last $B_s$ does not succeed and any $B_s$ fails. (c) Otherwise (namely, if the last $B_s$ does not succeed and all the $B_s$'s detect losses), the block-level sign is not determined. The block-level letter is determined only when all the physical-level letters are determined, namely, when no losses are detected and all $B_s$'s succeed. For such cases, the block-level letter is $\phi$ ($\psi$) if the number of $\psi$ in the BSM_{phy} results is even (odd).

Next, a BSM_{lat} is done by $n$-times of BSM_{blc}'s. The lattice-level sign is determined only when all the block-level signs are determined; it is (+) if the number of (−) in the BSM_{blc} results is even and it is (−) if the number is odd. The lattice-level letter is equal to any determined block-level letter. Thus, if all BSM_{blc}'s cannot determine letters, the lattice-level letter is not determined as well.

2. Modified scheme for our model

In our MBQC model with the parity state codes, we also consider the case where only on-off detectors are used. The concatenated BSM scheme should be slightly modified for this case.

Since failure and loss cannot be discriminated, a BSM_{phy} now has three possible outcomes: two successful cases and failure. Consequently, in a BSM_{blc}, $B_s$’s are performed until it either succeeds or consecutively fails $j$ times. The way to determine the block-level sign and letter is the same as the original scheme, except that the case (c) when determining the sign no longer occurs. The biggest difference from the original scheme is that the determined sign and letter may be wrong. We explicitly calculate these error probabilities in the next appendix.

In a BSM_{lat}, the lattice-level sign is determined by the block-level signs by the same method as the original scheme, although it may be wrong with a nonzero probability as well. On the other hand, the lattice-level letter is not determined by a single block-level letter as the original scheme; instead, we use a weighted majority vote of block-level letters. The weight of each block-level letter is given as

$$w(i) = \prod_{i \in t_s} \frac{1 - q_{lett}^{(i)}}{q_{lett}^{(i)}} / \prod_{i \in t_s} \frac{1 - q_{lett}^{(i)}}{q_{lett}^{(i)}} = \exp\left(\sum_{i=1}^n w(i)\right),$$

where $q_{lett}^{(i)}$ and $w(i)$ are respectively the letter error probability and the weight of the $i$th block. Note that the third equality comes from the fact that a lattice-level Bell state is decomposed into block-level Bell states of the same letter, as shown in Eq. (C1a) and (C1b).

Appendix D: Error probabilities of a BSM of parity state qubits under a lossy environment

We here derive the error probabilities of a BSM on two qubits (say, qubits 1 and 2) encoded with a parity state code. We suppose the loss model in the main text with the photon loss probability $\eta$ and denote $x := (1 - \eta)^2$, which is the probability that a BSM_{phy} does not detect photon losses. We also assume that the initial marginal state on qubits 1 and 2 before suffering losses is the equal mixture of four lattice-level Bell states. In other words, the four Bell states have the same prior probability, we get

$$\frac{\text{Pr}(\Phi|I_\phi, I_\phi)}{\text{Pr}(\Psi|I_\phi, I_\phi)} = \frac{\text{Pr}(I_\phi, I_\phi, \Phi) \text{Pr}(\Phi)}{\text{Pr}(I_\phi, I_\phi, \Psi) \text{Pr}(\Psi)} = \frac{\text{Pr}(I_\phi, I_\phi, \Phi)}{\text{Pr}(I_\phi, I_\phi, \Psi)} = \prod_{i \in t_s} \left(1 - q_{lett}^{(i)}\right) \prod_{i \in t_s} q_{lett}^{(i)} = \prod_{i \in t_s} \left(1 - q_{lett}^{(i)}\right) \prod_{i \in t_s} q_{lett}^{(i)}$$

We explicitly calculate these error probabilities in the next appendix.
1. With PNRDs resolving up to two photons

The case where PNRDs resolving up to two photons are used is already well explained in Ref. [32]. We here review the contents to be self-contained.

a. Block-level BSM (BSM\(_{blc}\))

Considering the determination of a block-level sign and letter in Appendix C, the outcome of a BSM\(_{blc}\) is included in one of the following three cases: (Success) Both the sign and letter are identified if no losses are detected and all the \(B_{\pm}\)'s succeed. (Failure) Neither sign nor letter is identified if no \(B_{\psi}\)'s succeed and all \(B_{\pm}\)'s detect losses. (Sign discrimination) Only the sign is identified if otherwise. The block-level sign (or letter) is selected randomly if it is not identified. The probabilities of these cases are respectively

\[
p_s = \left[1 - 2^{-(j+1)}\right]x^n, \quad p_t = \sum_{l=0}^{j} \left(\frac{x}{2}\right)^l (1-x)^{m-l}, \quad p_{sd} = 1 - p_s - p_t. \tag{D1}
\]

b. Lattice-level BSM (BSM\(_{lat}\))

For a BSM\(_{lat}\), let \(N_s\) (\(N_f\)) denote the number of successful (failed) BSM\(_{blc}\)'s. The lattice-level letter is identified if \(N_s \geq 1\) (namely, if at least one block-level letter is identified) and the sign is identified if \(N_f = 0\) (namely, if all block-level signs are identified). Hence, the outcome of a BSM\(_{lat}\) is included in one of the following four events:

\[
\begin{align*}
S & (\text{Success}) : \quad N_s \geq 1 \land N_f = 0, \\
D_L & (\text{Letter discrimination}) : \quad N_s, N_f \geq 1, \\
D_S & (\text{Sign discrimination}) : \quad N_s = N_f = 0, \\
F & (\text{Failure}) : \quad N_s = 0 \land N_f \geq 1.
\end{align*} \tag{D2}
\]

The sign and letter error probabilities \((q_{\text{sign}}, q_{\text{lett}})\) of the BSM\(_{lat}\) for each event are \((0, 0)\) for \(S\), \((1/2, 0)\) for \(D_L\), \((0, 1/2)\) for \(D_S\), and \((1/2, 1/2)\) for \(F\). The probabilities of the events are respectively given as

\[
\begin{align*}
P_S &= (1-p_t)^n - p_{sd}^n, \\
P_{D_L} &= 1 - (1-p_s)^n + (1-p_t)^n - p_{sd}^n, \\
P_{D_S} &= p_{sd}^n, \\
P_F &= (1-p_s)^n - p_{sd}^n. \tag{D3}
\end{align*}
\]

2. With on-off detectors

Next, we consider the case where only on-off detectors are used; namely, a photon loss in a BSM\(_{phy}\) is not detected and just leads to its failure.

a. Block-level BSM (BSM\(_1\))

Each outcome of a BSM\(_{blc}\) is uniquely identified by a triple \(O = (r, s, U)\), where \(r \in \mathbb{Z}_{j+1}\) is the number of failed \(B_{\psi}\)'s, \(s = \pm\) is the sign chosen by the successful \((r+1)\)th \(B_{\psi}\) (if \(r < j\)) or randomly (if \(r = j\)), and \(U\) is a \((m-r)\)-element tuple composed of \(\phi\), \(\psi\), and \(f\) (failure) indicating the outcomes of the BSM\(_{phy}\)'s from the \((r+1)\)th to the the last. (If \(r < j\), the first component of \(U\) is always \(\psi\) and the other components are determined by the \(B_s\)'s. If \(r = j\), all the components are determined by the \(B_s\)'s.) Let \(N_e(U)\) for \(e \in \{\phi, \psi, f\}\) denote the number of \(e\) in \(U\).
TABLE I. The error probabilities and the total probabilities of the \( j + 3 \) events [defined in Eq. (D4)] on the outcome of a BSM\(_{\text{blc}}\) when only on-off detectors are used. We define \( x := (1 - \eta)^2 \), where \( \eta \) is the photon loss rate.

| Event (\( \mathcal{E} \)) | Sign error probability \( q_{\text{sign}}^{\text{blc}} \) | Letter error probability \( q_{\text{lett}}^{\text{blc}} \) | Total probability \( p_x \) |
|---|---|---|---|
| \( \mathcal{S}_r \) \((0 \leq r \leq j)\) | 0 | \( 1/(2 - |x(2 - x)|)^{m-r}/2 \) | \( 1 - \sum_r p_{\mathcal{S}_r} - p_{\mathcal{F}} \) |
| \( \mathcal{F} \) \((1 - x)^{m-j}/[1 + (1 - x)^{m-j}] \) | 1/2 | \( (1 - x)/2 \) | \( 1 - (1 - x)/2 \) |
| \( \mathcal{D} \) | 0 | \( 1/[1 + (1 - x)^{m-j}] \) |

Then a BSM\(_{\text{blc}}\) outcome \( O \) is included in one of the following \( j + 3 \) events:

\[
\begin{align*}
\mathcal{S}_r &:= \{(r,s,U)|N_f(U) = 0\} \quad (0 \leq r \leq j), \\
\mathcal{F} &:= \{(j,s,U)|N_f(U) = m - j\}, \\
\mathcal{D} &:= \mathcal{O} \setminus \left[\mathcal{F} \cup \bigcup_{r=0}^{j} \mathcal{S}_r\right],
\end{align*}
\]

where \( \mathcal{O} \) is the set of all possible outcomes.

The probability that each event occurs and the error probabilities \( (q_{\text{sign}}^{\text{blc}}, q_{\text{lett}}^{\text{blc}}) \) conditioning to the event are explicitly presented in Table I. Note that, if \( \eta = 0 \), the events \( \mathcal{S}_r, \mathcal{F}, \), and \( \mathcal{D} \) correspond to success, failure, and sign discrimination. We show these results one by one from now on.

Before that, we note that every positive operator-valued measure (POVM) element of a lossy BSM\(_{\text{phy}}\) has vanishing off-diagonal entries in the Bell basis; see Appendix E for the proof. Also, each POVM element of a lossy BSM\(_{\text{blc}}\) [denoted by \( M^0_{\text{blc}} \) for each outcome \( O = (r,s,U) \)] is the tensor product of particular POVM elements of the lossy BSM\(_{\text{phy}}\)’s. Thus, the conditional probability of getting \( O \) from a block-level Bell state |\( B \rangle \) is

\[
\Pr(O | B) = \langle B| M^0_{\text{blc}} |B \rangle = \frac{1}{2^{m-1}} \sum_i \langle B_i| M^0_{\text{D}} |B_i \rangle = \frac{1}{2^{m-1}} \sum_i \Pr(O | B_i),
\]

where \( |B_i \rangle \)'s are the terms constituting the summation in Eq. (C1c) or (C1d) such that \( |B \rangle = \frac{1}{\sqrt{2^{m-1}}} \sum_i |B_i \rangle \). In other words, when calculating \( \Pr(O | B) \), it is enough to find \( \Pr(O | B_i) \)'s and then take their average.

The posterior probability of a block-level Bell state |\( B \rangle \) under a given outcome \( O \) is

\[
\Pr(B | O) = \frac{\Pr(O | B)}{\sum_{|B'| \in \mathcal{B}_{\text{blc}}} \Pr(O | B')},
\]

where \( \mathcal{B}_{\text{blc}} \) is the set of the four block-level Bell states. Thus, the result of the BSM\(_{\text{blc}}\) is selected randomly in the set \( R(O) := \arg \max_B \Pr(O | B) \). The sign (letter) error probability as a function of \( O \) is

\[
q_{\text{sign(lett)}}(O) = \frac{1}{|R(O)|} \sum_{|B| \in R(O)} \left[ \Pr(F_{\text{sign(lett)}}(B)|O) + \Pr(F_{\text{sign}} \circ F_{\text{lett}}(B)|O) \right],
\]

where \( |F_{\text{sign(lett)}}(B) \rangle \) is the Bell state obtained by flipping the sign (letter) from \( |B \rangle \) (e.g., \( F_{\text{sign}}(\phi^\pm) = \phi^\mp \)).

We now calculate the probabilities in Table I one by one. Let us first consider an outcome \( O = (r,s,U) \in \mathcal{S}_r \), where \( N_f(U) = 0 \). Regarding a single term in the decomposition of \( |\phi^\pm_{(m)} \rangle \) [see Eq. (C1c)], if there are total \( k \) of \( |\psi^\pm \rangle \)'s, the first \( r \) physical levels contain \( k - N_\psi(U) \) of |\( \psi^\pm \rangle \)'s, which should suffer photon losses by the definition of \( r \). If \( r < j \), \( s \) selected by the successful \( (r + 1) \)th \( B_\psi \) is certainly \pm, the sign of \( |\phi^\pm_{(m)} \rangle \). If \( r = j \), the randomly selected \( s \) may or may not be corrected; however, the latter case is out of \( \mathcal{S}_r \) since all the following \( B_\psi \)'s must fail. The remaining \( m - r \) BSM\(_{\text{phy}}\)'s should not suffer photon losses since \( N_f(U) = 0 \). Hence, for all \( U \) satisfying \( N_f(U) = 0 \), we get

\[
\Pr(r, \pm, U | \phi^\pm_{(m)}) = \frac{1}{2^{m-1}} \sum_{k: \text{even} \leq r + N_\psi} \binom{r}{k - N_\psi} (1 - x)^{k - N_\psi} \frac{1}{2^{m-r}} x^{m-r} \]

\[
= \frac{1}{2^{m-r}} \left[ (1 - \frac{x}{2})^r \left( \frac{x}{2} \right)^{m-r} + (-1)^{N_\psi} \left( \frac{x}{2} \right)^m \right],
\]

\[
\Pr(r, \mp, U | \phi^\pm_{(m)}) = 0,
\]
where $N_\psi = N_\psi(U)$. Similarly, we get

$$\Pr(r, \pm, U|\psi_{(m)}^\pm) = \frac{1}{2^m} \left[ (1 - \frac{x}{2})^r (\frac{x}{2})^{m-r} - (-1)^N_\psi (\frac{x}{2})^m \right],$$

$$\Pr(r, \mp, U|\psi_{(m)}^\pm) = 0.$$  \hspace{1cm} (D8)

From Eqs. (D5)–(D8), we obtain

$$q_{\text{blc}}^{\text{bloc}}(O) = 0 =: q_{\text{sign}}^{\text{bloc}}(S_r),$$

$$q_{\text{lett}}^{\text{bloc}}(O) = \frac{(1 - \frac{x}{2})^r (\frac{x}{2})^{m-r} - (\frac{x}{2})^m}{2 (1 - \frac{x}{2})^r (\frac{x}{2})^{m-r}} = \frac{1}{2} \left[ 1 - (\frac{x}{2})^r \right] =: q_{\text{lett}}^{\text{bloc}}(S_r).$$  \hspace{1cm} (D9)

Note that the error probabilities are the same for all $O \in S_r$. The total probability that the event $S_r$ occurs is

$$p_{\text{bloc}} := \frac{1}{4} \sum_{O \in S_r} \sum_{|B| \in \text{Bloc}} \Pr(O|B) = \frac{1}{2} \frac{1}{2^m} \left( 1 - \frac{x}{2} \right)^r (\frac{x}{2})^{m-r} 2^{m-r-1} = \frac{1}{2} (1 - \frac{x}{2})^r x^{m-r},$$

where the factor $2^{m-r-1}$ is the number of possible $U$’s for a given value of $r$.

Next, we consider $O = (j, s, U) \in F$, where $N_f(U) = m - j$, namely, all the $B_s$’s fail. Regarding a single term in the decomposition of $\langle \phi_{(m)}^\pm \rangle$, all the $|\psi^\pm\rangle$’s in the first $j$ physical levels should suffer photon losses. If $s = \pm$, all the following $B_\pm$’s should suffer photon losses as well. If $s = \mp$, all the following $B_\mp$’s fail regardless of photon losses. We thus get

$$\Pr(j, \pm, U = (f, \cdots, f)|\phi_{(m)}^\pm) = \frac{1}{2} \left( 1 - \frac{x}{2} \right)^j,$$

$$\Pr(j, \mp, U = (f, \cdots, f)|\phi_{(m)}^\pm) = \frac{1}{2} \left( 1 - \frac{x}{2} \right)^j,$$

where $k_1$ ($k_2$) in the summation indicates the number of $|\psi^\pm\rangle$’s in the first $j$ (last $m - j$) physical levels. Similarly, the same results are obtained for $|\psi_{(m)}^\pm\rangle$:

$$\Pr(j, \pm, U = (f, \cdots, f)|\psi_{(m)}^\pm) = \frac{1}{2} \left( 1 - \frac{x}{2} \right)^j (1 - x)^{m-j},$$

$$\Pr(j, \mp, U = (f, \cdots, f)|\psi_{(m)}^\pm) = \frac{1}{2} \left( 1 - \frac{x}{2} \right)^j.$$

The corresponding error probabilities are

$$q_{\text{blc}}^{\text{bloc}}(O) = \frac{(1 - x)^{m-j}}{1 + (1 - x)^{m-j}} =: q_{\text{sign}}^{\text{bloc}}(F),$$

$$q_{\text{lett}}^{\text{bloc}}(O) = \frac{1}{2} =: q_{\text{lett}}^{\text{bloc}}(F)$$

and the total probability of the event $F$ is

$$p_F = \frac{1}{2} \sum_{s = \pm} \sum_{|B| \in \text{Bloc}} \Pr(j, s, U = (f, \cdots, f)|B) = \frac{1}{2} \left( 1 - \frac{x}{2} \right)^j [1 + (1 - x)^{m-j}].$$

Lastly, we consider $O = (r, s, U) \in D$. If $r < j$, $N_f(U) > 0$ by the definition of $D$ and $N_f(U) < m - r$ since the first component of $U$ is always $\psi$. If $r = j$, $0 < N_f(U) < m - j$ by the definition of $D$. Therefore, regardless of $r$, $U$ contains at least one failure and one success ($\psi$ or $\phi$). Thanks to the successful BSMphy’s, the sign of the result is identified without an error. On the other hand, the letter is not identified because of the failures. We can see intuitively without calculation that the letter error probability is 1/2: Even if there is only one failure in $U$, the letter information of the corresponding physical-level Bell state is completely lost, considering that the marginal state of a block-level Bell state on a single physical level is $|\phi^\pm\rangle|\phi^\pm\rangle + |\psi^\pm\rangle|\psi^\pm\rangle$. Thus, the block-level letter information (determined by the parity of the number of BSMphy outcomes with $\psi$) is completely lost as well. To rewrite the results, we get

$$q_{\text{sign}}^{\text{bloc}}(D) = 0, \quad q_{\text{lett}}^{\text{bloc}}(D) = \frac{1}{2}, \quad p_D = 1 - \sum_{r=0}^{j} p_{S_r} - p_F.$$
b. Lattice-level BSM (BSM\textsubscript{lat})

Each \(n\)-tuple of events composed of \(S_r (0 \leq r \leq j)\), \(\mathcal{F}\), and \(\mathcal{D}\) corresponds to a possible outcomes of a BSM\textsubscript{lat}, whose probability can be directly obtained from Table \[\text{III}\]. Let us consider such an \(n\)-tuple \(E = (\mathcal{E}_1, \ldots, \mathcal{E}_n)\). A lattice-level sign error occurs when there is an odd number of block-level sign errors and \(\mathcal{F}\) is the only event where a block-level sign error may occur; thus, the sign error probability is

\[
q_{\text{sign}} = \sum_{i: \text{odd} \leq N_{\mathcal{F}}} \left( \frac{N_{\mathcal{F}}}{i} \right) q_{\text{sign}}^{\text{blc}}(\mathcal{F})^i \left[ 1 - q_{\text{sign}}^{\text{blc}}(\mathcal{F}) \right]^{N_{\mathcal{F}}-i} = \frac{1}{2} - \frac{1}{2} |1 - 2q_{\text{sign}}^{\text{blc}}(\mathcal{F})|^{N_{\mathcal{F}}},
\]

where \(N_{\mathcal{F}}\) is the number of \(\mathcal{F}\)‘s in \(E\).

A lattice-level letter error occurs when the weighted majority vote of the block-level letters gives a wrong answer. We consider i.i.d. random variables \(\Lambda_1, \ldots, \Lambda_n\) such that \(\Lambda_i \sim \text{Bernoulli}(q_i)\) for each \(i\) where \(q_i := q_{\text{lett}}^{\text{blc}}(\mathcal{E}_i)\), which indicates whether a letter error occurs in the \(i\)th block. A lattice-level letter error occurs if

\[
\sum_i (2\Lambda_i - 1) \log \frac{1 - q_i}{q_i} =: V(\Lambda_1, \ldots, \Lambda_n)
\]

is larger than zero or if it is equal to zero and the randomly selected letter is wrong. Therefore, we get

\[
q_{\text{lett}} = \Pr(V(\Lambda_1, \ldots, \Lambda_n) > 0) + \frac{1}{2} \Pr(V(\Lambda_1, \ldots, \Lambda_n) = 0)
\]

\[
= \sum_{(\lambda_1, \ldots, \lambda_n) \in \mathbb{Z}_2^n} \prod_{i=1}^n \Pr(\Lambda_i = \lambda_i) \left\{ \Theta[V(\lambda_1, \ldots, \lambda_n) > 0] + \frac{1}{2} \Theta[V(\lambda_1, \ldots, \lambda_n) = 0] \right\}
\]

\[
= \sum_{(\lambda_1, \ldots, \lambda_n) \in \mathbb{Z}_2^n} \prod_{i=1}^n \left[ q_i^{\lambda_i}(1 - q_i)^{1-\lambda_i} \right] \left[ \frac{1}{2} \text{sgn}(V(\lambda_1, \ldots, \lambda_n)) + \frac{1}{2} \right],
\]

\[
= \frac{1}{2} + \frac{1}{2} \sum_{(\lambda_1, \ldots, \lambda_n) \in \mathbb{Z}_2^n} \prod_{i=1}^n \left[ q_i^{\lambda_i}(1 - q_i)^{1-\lambda_i} \right] \text{sgn} \left( \sum_{i=1}^n (2\lambda_i - 1) \log \frac{1 - q_i}{q_i} \right),
\]

where \(\Theta[C]\) for a condition \(C\) is equal to 1 if \(C\) is true and 0 if it is false, and \(\text{sgn}(a) = a/|a|\) if \(a \neq 0\) and 0 if \(a = 0\).

Appendix E: Proof of vanishing off-diagonal entries of the POVM elements of a lossy BSM\textsubscript{phy}

Here, we prove that every POVM element of a lossy BSM\textsubscript{phy} has vanishing off-diagonal entries in the Bell basis when on-off detectors are used. Let \(\Lambda_q\) denote the photon loss channel of a loss rate \(q\) defined as \(\Lambda_q(\sigma) := (1 - q)\sigma + q|0\rangle\langle 0|\) for a single-qubit state \(\sigma\) and the vacuum state \(|0\rangle\). By substituting \(\sigma = |\psi\rangle\langle\psi|\) for an arbitrary pure state \(|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle\), we get \(\Lambda_q(|H\rangle|V\rangle) = (1 - q)|H\rangle|V\rangle\). Thus,

\[
\begin{align*}
(\Lambda_q \otimes \Lambda_q) (|\psi^+\rangle\langle\psi^+|) &= (1 - q)^2 |\psi^+\rangle\langle\psi^+| + q(1 - q)(|H\rangle|0\rangle + |V\rangle|0\rangle + |0\rangle|H\rangle + |0\rangle|V\rangle), \\
(\Lambda_q \otimes \Lambda_q) (|\phi^±\rangle\langle\psi^±|) &= (1 - q)^2 |\phi^±\rangle\langle\psi^±| + q(1 - q)(|H\rangle|0\rangle - |V\rangle|0\rangle - |0\rangle|H\rangle + |0\rangle|V\rangle),
\end{align*}
\]

(E1)

\[
(\Lambda_q \otimes \Lambda_q) (|\phi^+\rangle\langle\phi^+|) = (1 - q)^2 |\phi^+\rangle\langle\phi^+| + q(1 - q)(|H\rangle|0\rangle + |V\rangle|0\rangle + |0\rangle|H\rangle - |0\rangle|V\rangle),
\]

\[
(\Lambda_q \otimes \Lambda_q) (|\psi^+\rangle\langle\psi^+|) = (1 - q)^2 |\psi^±\rangle\langle\psi^±| + q(1 - q)(|H\rangle|0\rangle - |V\rangle|0\rangle - |0\rangle|H\rangle + |0\rangle|V\rangle).
\]

Now, let \(M_+\) and \(M_f\) denote the POVM elements of a lossy \(B_\psi\) corresponding to the outcomes \(|\psi^±\rangle\) and failure, respectively. By modelling a lossy \(B_\psi\) as a photon loss channel followed by an ideal \(B_\psi\), \(M_i\) for \(i \in \{+, -, f\}\) satisfies

\[
\text{Tr}[M_i \rho] = \text{Tr}[\Pi_i (\Lambda_q \otimes \Lambda_q) (\rho)],
\]

(E2)

for any two-qubit state \(\rho\), where

\[
\Pi_+ := |\psi^±\rangle\langle\psi^±|, \quad \Pi_- := |\psi^\mp\rangle\langle\psi^\mp|,
\]

\[
M_i := |\phi^±\rangle\langle\phi^±| + |\phi^\mp\rangle\langle\phi^\mp| + |H\rangle\langle H| + |V\rangle\langle V| + |0\rangle\langle 0| + |1\rangle\langle 1|,
\]

are the projectors of an ideal \(B_\psi\) with a lossy input. From Eqs. (E1) and (E2), we obtain \(\langle\psi^±| M_i |\phi^±\rangle = \langle\psi^±| M_i |\phi^\mp\rangle = \langle\psi^±| M_i |\psi^±\rangle = \langle\psi^±| M_i |\psi^\mp\rangle = 0\) for every \(i \in \{+, -, f\}\). Similar arguments can be done for a lossy \(B_+\) and \(B_-\) as well.
FIG. 7. Physical-level graph structures of post-H microclusters for the HIC and HIS configurations when the \((n,m)\) parity state encoding is used. The squares (circles) indicate lattice-level (physical-level) qubits and black squares (circles) mean that the lattice-level (physical-level) \(H\) gates are applied on the qubits after the cluster states are created. Repetitive subgraph structures are abbreviated as blue dashed squares or circles with numbers as defined in the top of the figure.

Appendix F: Microclusters for the scheme with the parity state encoding

In this appendix, we first derive the physical-level graph structures of post-\(H\) microclusters required for the scheme with the parity state encoding and then present the method to generate them.

1. Physical-level graph structures of post-\(H\) microclusters

The physical-level graph structures of the central and side post-\(H\) microclusters are shown in Fig. 7 for the two \(H\)-configurations. Here, the squares (circles) indicate lattice-level (physical-level) qubits. If a square (circle) is filled with black, it means that the lattice-level (physical-level) Hadamard gate is applied on the qubit after the cluster state is created. Repetitive subgraph structures (namely, sets of vertices sharing the same neighborhood) are abbreviated as blue dashed squares or circles with numbers.

The first step to derive the graph structures is to investigate how a cluster state is transformed if a Hadamard gate \((H_1)\) is applied on one of the qubits (say, qubit 1) and then a \(cz\) gate \((CZ_{12})\) is applied on qubit 1 and another qubit (say, qubit 2) that is not adjacent to qubit 1. Note that, in the Heisenberg picture, the \(cz\) gate transforms the Pauli-\(X\) operators of the qubits as \(X_1 \rightarrow X_1Z_2\) and \(X_2 \rightarrow Z_1X_2\), while it leaves the Pauli-\(Z\) operators the same. For a qubit \(i\), \(S_i := X_i \prod_{j \in N(i)} Z_j\), where \(N(i)\) is the set of qubits adjacent to qubit \(i\), is a stabilizer of the initial cluster.
FIG. 8. (a) Transformation of a cluster state by applying a Hadamard gate followed by applying a CZ gate. (b) Physical-level graph structure of the state $|+_{L}\rangle = |0_{L}\rangle + |1_{L}\rangle$. (c) Physical-level graph structure of a lattice-level three-qubit linear cluster state.

FIG. 9. Encoding circuit of the state $|+_{L}\rangle := |0_{L}\rangle + |1_{L}\rangle$ in the $(3,3)$ parity state code, which employs multiple copies of the state $|+\rangle := |h\rangle + |v\rangle$, CZ gates, and Hadamard gates. The label $[i,j]$ for each physical-level qubit indicates the index $i$ of the block and the index $j$ of the photon in the block.

The stabilizers $S_{1}$ and $S_{1}S_{2}$ are transformed by $CZ_{12}^2H_{1}$ as

$$S_{1} = X_{1} \prod_{j \in N(1)} Z_{j} \rightarrow Z_{1} \prod_{j \in N(1)} Z_{j} = H_{1} \left( X_{1} \prod_{j \in N(1)} Z_{j} \right) H_{1},$$

$$S_{1}S_{2} = X_{1}X_{2} \prod_{j \in N(1)\triangle N(2)} Z_{j} \rightarrow X_{2} \prod_{j \in N(1)\triangle N(2)} Z_{j} = H_{1} \left( X_{2} \prod_{j \in N(1)\triangle N(2)} Z_{j} \right) H_{1},$$

where $A\Delta B := A \cup B \setminus (A \cap B)$ for two sets $A$ and $B$. Also, for each qubit $i \in N(1)$,

$$S_{i} = X_{i} \prod_{j \in N(i)} Z_{j} \rightarrow X_{i}X_{1}Z_{2} \prod_{j \in N(i)\setminus \{1\}} Z_{j} = H_{1} \left( X_{i} \prod_{j \in N(i)\setminus \{2\}} Z_{j} \right) H_{1}.$$

Therefore, the overall effect of the process is, for each qubit $i$ adjacent to qubit 1, to flip the connectivity of the qubits 2 and $i$ (namely, connect them if they are disconnected and disconnect them if they are already connected) and then apply $H_{1}$. An example of this transformation is presented in Fig. 8(a).

Next, we obtain the graph structure of the state $|+_{L}\rangle := |0_{L}\rangle + |1_{L}\rangle$. Figure 9 shows the encoding circuit of the state for the $(3,3)$ parity state code, which employs multiple copies of the state $|+\rangle := |h\rangle + |v\rangle$, CZ gates, and Hadamard
FIG. 10. Circuit to implement the lattice-level Hadamard gate of the (3, 3) parity state code.

gates. Here, we label the $j$th physical qubit of the $i$th block by $[i,j]$. It is straightforward to generalize it for any parity state code. The graph structure of $|+_L\rangle$ shown in Fig. 8(b) is obtained by preparing $nm$ isolated vertices and tracking the transformation of the graph via the $cz$ and Hadamard gates in the circuit.

A lattice-level $cz$ gate $C^L$ is done by $m^2$ physical $cz$ gates:

$$C^L = \prod_{i,j \leq m} C_{i,j}^{Z}, \quad (F1)$$

where $C_{i,j}^{Z}$ is the $cz$ gate between the $[i,j]$ qubit of the first lattice-level qubit and the $[k,l]$ qubit of the second lattice-level qubit. It can be verified as follows: The stabilizer generators of the $(n,m)$ parity state code are

$$\begin{cases} X_{ij}X_{i(j+1)} \quad (\forall i \leq n, \forall j \leq m - 1), \\ Z_{ij}Z_{j+1}j \quad (\forall i \leq n - 1) \end{cases}$$

and the lattice-level Pauli operators are

$$X_L = X_{11} \cdots X_{n,1}, \quad Z_L = Z_{11} \cdots Z_{1,m},$$

where $X_{ij}$ ($Z_{ij}$) is the Pauli-$X$ ($-Z$) operator on the $[i,j]$ qubit. It is straightforward to see that the RHS of Eq. (F1) commutes with all the stabilizers and transforms the lattice-level Pauli operators correctly.

Combining the above results on the $|+_L\rangle$ state and the lattice-level $cz$ gate, we attain the graph structure of a lattice-level three-qubit linear cluster state shown in Fig. 8(c). The only left ingredient is the lattice-level Hadamard gate ($H^L$).

We now describe the generation of post-$H$ microclusters. The physical-level graphs of post-$H$ microclusters (see Fig. 7) can be decomposed into multiple components through the process shown in Fig. 11. Each repetitive subgraph structure connected with multiple vertices is separated and connected with only one vertex. The results are shown in Fig. 12; the post-$H$ microclusters can be generated by preparing individual components first and then merging them through fusions. For microclusters that are not shown in the figure (including the central microcluster of the HIC configuration and other microclusters that do not meet the presented conditions), the decomposition is not necessary; that is, the entire physical-level graph itself is a single component. This process may greatly reduce the number of required fusions due to the reduction of the number of edges as shown in Fig. 11.

Each component can be generated by the following method systematically. First, we consider a merging graph of a cluster state where vertices correspond to 3-GHZ states and edges indicate required merging operations (BSMs or fusions). A merging graph $G_{\text{mrg}}$ is constructed by the following process [see Fig. 13(a)–(c) for an example]:

2. Generation of post-$H$ microclusters
FIG. 11. Decomposition of a cluster state by separating repetitive subgraph structures that are connected with multiple vertices.

FIG. 12. Decomposition of post-\(H\) microclusters into multiple components done by separating connected pairs of repetitive subgraph structures. Individual components are prepared first and then merged through fusions to construct post-\(H\) microclusters. The central post-\(H\) microcluster of the HIC configuration does not have connected pairs of repetitive subgraph structures, thus it is not presented here.

1. The graph \(G_{mrg} = (V, E)\) is initialized by the physical-level graph \(G_{phy}\) of the cluster state \(|G_{phy}\rangle\). Let us define \(v_{\text{root}}(v) := v\) for each vertex \(v\) with degree 2 in \(V\).

2. Let us define \(V_{\text{deg} \geq 3} := \{v \in V | \deg(v) \geq 3\}\). This set is fixed and not updated during the entire process. For
each vertex \( v \in V_{\text{deg} \geq 3} \), perform the following:

(a) Remove \( v \) from \( G_{\text{mrg}} \) and add \( d_v - 1 \) new vertices, where \( d_v := \text{deg}(v) \). Let \( V_{\text{new}} = \left( v_{\text{new}}^{(1)}, \ldots, v_{\text{new}}^{(d_v-1)} \right) \) denote the series of the new vertices and \( V_{\text{nhg}} = \left( v_{\text{nhg}}^{(1)}, \ldots, v_{\text{nhg}}^{(d_v)} \right) \) denote the series of the vertices that were adjacent to \( v \) before removing it. The order of the vertices in these series can be freely chosen.

(b) Connect the vertices in \( V_{\text{new}} \) linearly; namely, connect \( \left( v_{\text{new}}^{(1)}, v_{\text{new}}^{(2)} \right), \left( v_{\text{new}}^{(2)}, v_{\text{new}}^{(3)} \right), \ldots \).

(c) Choose one of the vertices in \( V_{\text{new}} \) randomly. Let us term this vertex the root vertex \( v_{\text{root}}(v) \).

(d) Let us define \( V'_{\text{new}} \) by omitting \( v_{\text{root}}(v) \) from \( V_{\text{new}} \) and keeping the order of the vertices. For each \( i = 1, \cdots, d_v - 2 \), connect the \( i \)th element of \( V'_{\text{new}} \) with \( v_{\text{nhg}}^{(i)} \).

(e) Connect \( v_{\text{new}}^{(1)} \) with \( v_{\text{nhg}}^{(3)} \) and \( v_{\text{new}}^{(2)} \) with \( v_{\text{nhg}}^{(2)} \).

(f) Every edge connecting two vertices in \( V_{\text{new}} \) is labelled as “internal.” Every edge connecting a vertex in \( V_{\text{new}} \) and a vertex in \( V_{\text{nhg}} \) is labelled as “external.”

3. Remove all vertices with degree 1 from the graph.

Note that there are two types of randomness in the above process for each vertex \( v \in V_{\text{deg} \geq 3} \): the order of the vertices in \( V_{\text{nhg}} \) and the choice of \( v_{\text{root}}(v) \). The merging graph is uniquely identified by \( G_{\text{phy}} \) and these two factors, which is important when optimizing the resource overheads (see Appendix H).

Let us define the 1-GHZ state for an integer \( l \geq 3 \) by the state \( |\text{GHZ}_l\rangle := |H\rangle^\otimes l + |\rangle^\otimes l \). Note that it is a post-\( H \) cluster state with a star graph \( S_{l-1} \) where the Hadamard gates are applied on all the leaves of the graph; namely, \( |\text{GHZ}_l\rangle = H_2 \cdots H_l |C_{12} Z \cdots C_{l-1} Z| + \rangle^\otimes l \). We refer to the first qubit of the above expression as the root qubit of the state and the other qubits as its leaf qubits. The root qubit can be chosen arbitrarily since the state is symmetric about the permutations of its support qubits.

For a given merging graph \( G_{\text{mrg}} \), the original cluster state can be generated up to known local Pauli operators by the following process [see Fig. 13(d) for an example]:

1. Place a 3-GHZ state at each vertex in \( G_{\text{mrg}} \). The root qubits of the 3-GHZ states are chosen arbitrarily and fixed during the entire process.

2. For each internal edge, a pair of qubits is selected respectively from the two involved 3-GHZ states such that one qubit is a root qubit, the other is a leaf qubit, and both of them are not selected previously. The root qubits of the 3-GHZ states on the vertices \( v_{\text{root}}(v) \)'s chosen while generating \( G_{\text{mrg}} \) should not be selected here.
3. For each external edge, a pair of qubits is selected respectively from the two involved 3-GHZ states such that both of the qubits are leaf qubits and not selected previously.

4. For each pair of qubits selected above, perform a BSM if it is from an internal edge and perform a fusion if it is from an external edge.

5. For each remained qubit that was originally a leg qubit of a 3-GHZ state, apply the Hadamard gate on the qubit.

The above process is equivalent to preparing a \([\deg(v) + 1]\)-GHZ state for each vertex \(v\) in \(G_{\text{phy}}\) with degree larger than 1 and then merging them through fusions along the edges in \(G_{\text{phy}}\). In detail, performing a BSM on a pair of qubits selected in step 2 corresponds to merging two GHZ states (say, \(|\text{GHZ}_{l_1}\rangle\) and \(|\text{GHZ}_{l_2}\rangle\)) to form an \((l_1 + l_2 - 2)\)-GHZ state. For each vertex \(v\) in \(G_{\text{phy}}\), \(G_{\text{mrg}}\) contains \(\deg(v) - 1\) vertices connected by internal edges, thus merging this number of 3-GHZ states forms a \([\deg(v) + 1]\)-GHZ state. Then, by performing a fusion on each pair of qubits selected in step 3, these GHZ states are merged with each other to form the desired cluster state.

It is worth noticing that the order of all the BSMs and fusions throughout the generation of a post-\(H\) microcluster can be chosen arbitrarily, although the generation of individual components and the process to merge them are separately depicted here for convenience. It is because different BSMs or fusions always act on different pairs of qubits. This degree of freedom severely affects the resource overhead of generating the cluster state, as will be discussed in Appendix H.

Appendix G: Details on the error simulations

In this appendix, we describe the error simulation method in details. We first introduce the parameters that determine the details of the MBQC model:

- **encoding**: If True, the scheme using the parity state encoding is used. If False, the scheme without logical encoding is used.
- **pssl**: If True, star clusters in which all the step-1 fusions are successful are post-selected to be used to construct the RTCS lattice. If False, all generated star clusters are used regardless of the fusion results.
- **hic**: If True, the configuration of the Hadamard gates is HIC. If False, it is HIS. It is always True if encoding is False.
- **d**: Code distance.
- **\(\eta\)**: Photon loss rate.

There are some additional parameters depending on the value of encoding. If encoding is True,

- **pnrd**: If True, PNRDs resolving up to two photons are used. If False, on-off detectors are used.
- **n, m**: The \((n, m)\) parity state code is used to encode side qubits.
- **j**: The maximal number of \(B_{\psi}\)'s in a BSM_{blc}. (See the CBSM scheme in Appendix C)

If encoding is False,

- **\(p_f\)**: Failure rate of a fusion when \(\eta = 0\). If \(\eta > 0\), the failure rate is \(p_f(1 - \eta)^2\).

Fixing the values of the above parameters, we consider an RTCS lattice whose boundaries are in the form of a cuboid. Let us denote the three axes of the cuboid by the \(x\)-, \(y\)-, and \(t\)-axis and the corresponding boundaries as the \(x\)-, \(y\)-, and \(t\)-boundaries. The \(t\)-axis is also referred to as the simulated time axis. The cuboid has the widths of \(d - 1\) unit cells along the \(x\)- and \(y\)-axis to make the code distance as \(d\). On the other hand, its width along the \(t\)-axis is \(4d + 1\) unit cells, which is arbitrarily set to be larger enough than \(d\) for reducing the effects of errors near the \(t\)-boundaries. The \(x\)- and \(t\)-boundaries are set to be primal, while the \(y\)-boundaries are set to be dual. In other words, the \(x\)- and \(t\)-boundaries adjoin normally on primal unit cells, while the \(y\)-boundaries cross the middle of primal unit cells. Since we are only interested in error chains connecting the opposite \(x\)-boundaries, it is assumed that the qubits on the \(t\)-boundaries do not have errors.

We use a Monte-Carlo method for the simulations. Each trial is proceeded as follows:
1. Sample the outcomes of all fusions in steps 1 and 2 (only step 2 if \(\text{pssl} = \text{True}\)). If \(\text{encoding} = \text{True}\), the probabilities presented in Appendix [D] which depend on the values of \(\text{prr}, n, m, j, \) and \(\eta\), are used. If it is \(\text{False}\), a fusion fails \((q_{\text{lett}} = 0 \text{ and } q_{\text{sign}} = 1/2)\) with probability \(p_{f}(1-\eta)^{2}\) and it detects a loss \((q_{\text{lett}} = q_{\text{sign}} = 1/2)\) with probability \((1 - (1 - \eta)^{2})\).

2. For each fusion outcome, the corresponding error probabilities \((q_{\text{sign}}, q_{\text{lett}})\) are obtained and whether the fusion has a sign or letter error is randomly determined by the probabilities. These error probabilities and errors are then propagated to appropriate central qubits determined by the value of \(\text{nic}\). For each central qubit \(i\), the presence or absence of an error and its probability are assigned to a boolean variable \(\text{error}_{i}\) and a floating-point variable \(q_{\text{err},i}\), respectively.

3. For each central qubit \(i\), a photon loss is sampled with probability \(\eta\). If it has a loss, \(q_{\text{err},i}\) is updated to 0.5 and \(\text{error}_{i}\) is flipped with probability 50%.

4. The syndrome of each parity-check operator (which corresponds to a primal unit cell) is determined by the values of \(\text{error}_{i}\)'s of the qubits in the support of the operator.

5. The syndromes are decoded to infer the locations of the errors. We use the weighted minimum-weight perfect matching (MWPM) decoder via PyMatching package \[34\] where the weight for each qubit \(i\) is \(\log((1 - q_{\text{err},i})/q_{\text{err},i})\). (If \(q_{\text{err},i} = 0\), the weight is infinity, which is handled by ignoring the qubit in the input for the decoder.) Exceptionally, if every value of \(q_{\text{err},i}\) is either 0 or 1/2, the qubits with \(q_{\text{err},i} = 1/2\) are given the weight of one, not zero, for a technical reason.

6. The remaining errors are obtained by comparing the original and estimated errors. If the number of the remaining errors on one side of the \(x\)-boundaries is odd, we regard that this trial has a logical error.

The logical error rate \(p_{L}\) for a given parameter setting is obtained by repeating the above process a sufficient number of times. In detail, we repeat the process until \(\Delta p_{L}/p_{L} \leq 0.1\) is reached where \(\Delta p_{L}\) is half the width of the 99% confidence interval. The logical error rates \(p_{L}^{(9)}(\eta)\), \(p_{L}^{(11)}(\eta)\) are calculated while varying \(\eta\) for two code distances \(d = 9, 11\) and the loss threshold \(\eta_{\text{th}}\) is obtained by finding the largest \(\eta\) satisfying \(p_{L}^{(11)}(\eta) + \Delta p_{L}^{(11)}(\eta) < p_{L}^{(9)}(\eta) - \Delta p_{L}^{(9)}(\eta)\).

### Appendix H: Details on the resource analysis

#### 1. Methods

We first describe the details on the resource analysis, particularly focusing on the problem of calculating \(N_{\text{GHZ}}^{\text{central}}(N_{\text{GHZ}}^{\text{side}})\); the expected number of required 3-GHZ states to generate one star cluster. Let \(N_{\text{GHZ}}^{\text{central}}(N_{\text{GHZ}}^{\text{side}})\) denote the expected number of required 3-GHZ states to generate one central (side) microcluster. Then we get

\[
N_{\text{GHZ}}^{*} = \left\{ \begin{array}{ll}
(N_{\text{GHZ}}^{\text{central}} + N_{\text{GHZ}}^{\text{side}})/p_{\text{succ,step1}} + N_{\text{GHZ}}^{\text{side}})/p_{\text{succ,step1}} & \text{if } \text{pssl} = \text{True}, \\
N_{\text{GHZ}}^{\text{central}} + 2N_{\text{GHZ}}^{\text{side}} & \text{if } \text{pssl} = \text{False},
\end{array} \right.
\]

where \(p_{\text{succ,step1}}\) is the average success probability of step-1 fusions and \(\text{pssl}\) is defined in Appendix [G]. Each microcluster is generated by merging multiple 3-GHZ states via BSMs or fusions according to a merging graph, as presented in Appendix [F2]. We consider using only microclusters generated without failed merging operations.

Note that, if two cluster states composed of physical-level qubits respectively require \(N_{1}\) and \(N_{2}\) of 3-GHZ states to generate, it averagely requires

\[
N_{1} + fN_{2} := \frac{2}{(1-\eta)^{2}}(N_{1} + N_{2}),
\]

3-GHZ states to successfully generate their merged cluster state. The factor \(2/(1-\eta)^{2}\) is the inverse of the success probability of the fusion. Since \(+f\) is not associative, the merging order is important when merging three or more cluster states.

For the generation of a post-\(H\) microcluster \(|\text{MC}\rangle\), \(N_{\text{MC}}^{\text{GHZ}}(N_{\text{GHZ}}^{\text{central}}\) or \(N_{\text{GHZ}}^{\text{side}}\)\) is affected by the choice of a merging graph \(G_{\text{merg}}\) for each of its components and the order of BSMs and fusions (see Appendix [F2]). To compute \(N_{\text{MC}}^{\text{GHZ}}\), we take an approach that merging graphs are chosen randomly and the order of BSMs and fusions is selected by an algorithm found heuristically:
1. For each component (see Fig. 12 of |MC|), one of its merging graph is randomly selected through the process in Appendix F.2.

2. Let \( G_{\text{comb}} \) be the disjoint union of the merging graphs. For each fusion required to generate |MC| from the components, connect the corresponding two vertices in \( G_{\text{comb}} \). More precisely, if two vertices \( v_1, v_2 \) (respectively from the physical-level graphs of two components) are fused, \( v_{\text{root}}(v_1) \) and \( v_{\text{root}}(v_2) \) in \( G_{\text{comb}} \) are connected.

3. Allocate a weight of 1 to every vertex in \( G_{\text{comb}} \). Let \( N_v \) denote the weight of vertex \( v \). For an edge \( e \) connecting \( v_1 \) and \( v_2 \), define \( N_e := N_{v_1} + f N_{v_2} \).

4. Repeat the follows until \( G_{\text{comb}} \) has only one vertex.
   (a) Find the set \( E_{\text{min}, \text{wgt}} \) of edges with the smallest value of \( N_e \).
   (b) Using an edge coloring algorithm, allocate “colors” to all edges so that different edges sharing a vertex have different colors and as few colors as possible are used.
   (c) Partition \( E_{\text{min}, \text{wgt}} \) into disjoint subsets by the colors of the edges. Find the largest subset \( E_{\text{mrg}} \) among them. If such a subset is not unique, choose one randomly.
   (d) For each edge \( e \) in \( E_{\text{mrg}} \), perform the follows.
      i. Contract the two vertices \((v_1, v_2)\) connected by \( e \) into a single vertex \( v_e \). Namely, \( v_1 \) and \( v_2 \) are replaced with \( v_e \) and every edge connected to \( v_1 \) or \( v_2 \) is connected to \( v_e \) instead. This process may make multiple edges connecting the same pair of vertices, which is allowed. If there are other edges connecting \( v_1 \) and \( v_2 \) besides \( e \), they are transformed into loops connected to \( v_e \).
      ii. Allocate \( N_e \) to \( v_e \).

5. \( N_{\text{GHZ}}^{\text{MC}} \) is equal to the weight of the only vertex left in \( G_{\text{comb}} \).

In the middle of the above process, each vertex of \( G_{\text{comb}} \) represents a connected subgraph of the intermediate cluster state and its weight is equal to the expected number of 3-GHZ states to construct the subgraph state. Thus, each contraction in step 2(d) corresponds to merging two independent cluster states to form a larger cluster state.

Steps 2(a)–(c) shows the strategy to determine the order of the merging operations. Although it may be not an optimal strategy, it is based on the following two intuitions: First, it is better to merge cluster states with small weights first, since \( (N_1 + f N_2) + f N_3 < N_1 + f (N_2 + f N_3) \) if \( N_1 < N_2 < N_3 \). Secondly, it is better to perform merging operations in parallel as much as possible and the edges with the same color (assigned by an edge coloring algorithm) can be merged parallelly. For coloring edges, we use the greedy graph coloring function \texttt{coloring.greedy.color} in NetworkX package \cite{NetworkX} with the strategy \texttt{largest.first}. Since the function performs vertex coloring, we input the line graph of \( G_{\text{mrg}} \) into the function.

To obtain the results in Fig. 1(b), we sample 1200 values of \( N_{\text{GHZ}}^{\text{MC}} \) through the aforementioned process. Let \( N_1 \) \( (N_2) \) be the minimal values of \( N_{\text{GHZ}}^{\text{MC}} \) for the first 600 (total 1200) samples. If \( N_1 = N_2 \), the value is returned. If otherwise, we sample 1200 values of \( N_{\text{GHZ}}^{\text{MC}} \) again and denote the minimal \( N_{\text{GHZ}}^{\text{MC}} \) for the total 2400 samples by \( N_3 \). If \( N_2 = N_3 \), the value is returned. If otherwise, we sample 2400 merging graphs again and so on. By varying the total number of samples in this way, it is possible to increase the odds that we reach close to the real optimal value.

2. Detailed results

Here, we show detailed simulation results on resource overheads in addition to the results shown in the main text. Table 11 shows the optimal parameter settings and the values of \( N_{\text{GHZ}}^{\text{*}} \) for various target loss thresholds. These results correspond to the data points along the upper envelopes in Fig. 5 of the main text. Figure 14 presents the values of \( N_{\text{GHZ}}^{\text{*}} \) as a function of the loss rate \( \eta \leq \eta_{th} \) for several optimal parameter settings. It is observed that \( N_{\text{GHZ}}^{\text{*}} \) varies much less than one order of magnitude when \( \eta \leq \eta_{th} \).

---

[1] T. C. Ralph and G. J. Pryde, Chapter 4 - Optical Quantum Computation (Elsevier, 2010) pp. 209–269.

[2] R. Raussendorf and H. J. Briegel, A one-way quantum computer, Phys. Rev. Lett. 86, 5188 (2001).

[3] R. Raussendorf, D. E. Browne, and H. J. Briegel, Measurement-based quantum computation on cluster
TABLE II. Optimal parameter settings and the corresponding values of $N^\ast_{\text{GHZ}}$ for various target loss thresholds and settings of \textit{pnrd} (whether PNRDs resolving up to two photons or on-off detectors are used) and \textit{pssl} (whether star clusters are post-selected or not).

| Threshold | $N^\ast_{\text{GHZ}}$ | $n$ | $m$ | $j$ | $H$-configuration |
|-----------|-------------------|----|----|----|----------------|
| When \textit{pnrd,pssl} = (True,True) | | | | | |
| 0.002 | $1.6 \times 10^2$ | 1 | 3 | 2 | HIC |
| 0.009 | $3.3 \times 10^2$ | 1 | 4 | 3 | HIS |
| 0.02 | $3.9 \times 10^2$ | 2 | 2 | 1 | HIC |
| 0.03 | $8.8 \times 10^2$ | 3 | 2 | 1 | HIC |
| 0.035 | $1.0 \times 10^3$ | 2 | 3 | 2 | HIC |
| 0.036 | $1.8 \times 10^3$ | 4 | 2 | 1 | HIC |
| 0.04 | $1.9 \times 10^3$ | 2 | 4 | 3 | HIC |
| 0.052 | $2.7 \times 10^3$ | 3 | 3 | 1 | HIC |
| 0.067 | $5.3 \times 10^3$ | 4 | 3 | 1 | HIC |
| 0.074 | $8.6 \times 10^3$ | 5 | 3 | 1 | HIC |
| 0.085 | $2.3 \times 10^4$ | 5 | 4 | 2 | HIC |
| When \textit{pnrd,pssl} = (False,True) | | | | | |
| 0.001 | $8.8 \times 10^2$ | 1 | 4 | 3 | HIC |
| 0.003 | $1.5 \times 10^3$ | 1 | 5 | 4 | HIS |
| 0.009 | $1.0 \times 10^4$ | 2 | 3 | 2 | HIC |
| 0.012 | $2.1 \times 10^4$ | 2 | 4 | 3 | HIC |
| 0.013 | $3.5 \times 10^4$ | 2 | 5 | 4 | HIC |
| 0.022 | $1.2 \times 10^5$ | 3 | 3 | 1 | HIC |
| 0.024 | $3.2 \times 10^5$ | 3 | 4 | 2 | HIC |
| 0.032 | $1.3 \times 10^6$ | 4 | 3 | 1 | HIC |
| 0.035 | $3.9 \times 10^6$ | 4 | 4 | 2 | HIC |
| 0.037 | $4.3 \times 10^7$ | 5 | 4 | 1 | HIC |
| When \textit{pnrd,pssl} = (False,False) | | | | | |
| 0.001 | $9.9 \times 10^2$ | 2 | 3 | 2 | HIS |
| 0.003 | $1.5 \times 10^3$ | 2 | 4 | 3 | HIC |
| 0.004 | $1.9 \times 10^3$ | 2 | 4 | 2 | HIC |
| 0.005 | $1.9 \times 10^3$ | 3 | 3 | 2 | HIC |
| 0.008 | $2.3 \times 10^3$ | 3 | 3 | 1 | HIC |
| 0.013 | $3.7 \times 10^3$ | 4 | 3 | 1 | HIC |
| 0.014 | $4.4 \times 10^3$ | 3 | 4 | 2 | HIC |
| 0.016 | $4.7 \times 10^3$ | 4 | 3 | 1 | HIC |
| 0.02 | $6.0 \times 10^3$ | 5 | 3 | 1 | HIC |
| 0.023 | $9.6 \times 10^3$ | 4 | 4 | 2 | HIC |
| 0.025 | $1.4 \times 10^4$ | 4 | 5 | 3 | HIC |
| 0.028 | $1.5 \times 10^4$ | 5 | 4 | 2 | HIC |
| 0.03 | $1.6 \times 10^4$ | 5 | 4 | 1 | HIC |
| 0.032 | $2.3 \times 10^4$ | 5 | 5 | 2 | HIC |

![FIG. 14. Average numbers $N^\ast_{\text{GHZ}}$ of required 3-GHZ states per center qubit as a function of the photon loss rate $\eta$ for several parameter settings in Table II.](image)

[4] R. Raussendorf, J. Harrington, and K. Goyal, A fault-tolerant one-way quantum computer, \textit{Ann. Phys.} \textbf{321}, 2212 (2006).

[5] R. Raussendorf, J. Harrington, and K. Goyal, Topological fault-tolerance in cluster state quantum computation, \textit{New J. Phys.} \textbf{9}, 199 (2007).

[6] A. G. Fowler and K. Goyal, Topological cluster state quantum computing, \textit{Quantum Info. Comput.} \textbf{9}, 721–738 (2009).

[7] D. Herr, A. Paler, S. J. Devitt, and F. Nori, Lattice surgery on the Raussendorf lattice, \textit{Quantum Sci. Technol.} \textbf{3}, 035011 (2018).

[8] B. J. Brown and S. Roberts, Universal fault-tolerant
[54x249][22] D. Herr, A. Paler, S. J. Devitt, and F. Nori, A local
[54x134][25] H. A. Zaidi and P. van Loock, Beating the one-half limit
[54x207][23] Y. Li, P. C. Humphreys, G. J. Mendoza, and S. C.
[54x280][21] F. Ewert and P. van Loock, 3
[54x280] W. P. Grice, Arbitrarily complete Bell-state measure-
[54x176][24] S. Takeda, T. Mizuta, M. Fuwa, P. Van Loock, and A. Fu-
[54x353][19] S. Omkar, S.-H. Lee, Y. S. Teo, S.-W. Lee, and H. Jeong,
[54x636][11] D. E. Browne and T. Rudolph, Resource-efficient lin-
[54x667][10] S. Bravyi and A. Kitaev, Universal quantum computa-
[54x427][17] S. Omkar, Y. S. Teo, S.-W. Lee, and H. Jeong, Highly
[54x489][15] H. Jeong and M. S. Kim, Efficient quantum computation
[54x604][12] J. M. Auger, H. Anwar, M. Gimeno-Segovia, T. M. Stace,
[54x678][30] M. Pant, D. Towsley, D. Englund, and S. Guha, Perco-
[54x720], 033305 (2020).
[140x364], 052324 (2015).
[230x479], 042305 (2002).
[235x186], 041007 (2015).
[266x448], 060501 (2019).
[278x155], 315 (2021).
[286x552], 015005 (2017).
[305x730], 260501 (2013).
[317x542]\[30\] M. Pant, D. Towsley, D. Englund, and S. Guha, Perco-
[317x720], 033305 (2020).
[317x144], edited by G. Varoquaux, T. Vaught, and J. Mill-
[335x521]Commun. 10.1145/3505637 (2021).
[335x531]architectural requirements for generating universal pho-
[335x584]with nondeterministic entangling gates, Phys. Rev. A 97.
[335x646]Rev. Lett. 140403 (2018).
[341x217]et al. Fusion-based quantum
computing, arXiv preprint arXiv:2101.09310 (2021).
[345x374], 200502 (2010).
[363x144], Fusion-based quantum
computing, arXiv preprint arXiv:2101.09310 (2021).
[377x699]near determinism of universal cluster states with probabilistic Bell measurements and three-qubit re-
source states, Phys. Rev. A 91, 042301 (2015)
[388x521], 1 (2019).
[395x730], 230 (1998)
[407x437], 39 (2005).
[408x458], 052303 (2019).
[421x374], 200502 (2010).
[442x646], 020502 (2015).
[460x604], 042301 (2015).
[462x228], Fusion-based quantum
computing, arXiv preprint arXiv:2101.09310 (2021).
[467x228], Fusion-based quantum
computing, arXiv preprint arXiv:2101.09310 (2021).
[488x196], 595 (2021)
[493x312], 595 (2021)
[500x756]22
[526x500], 100501 (2005)
[537x196], 39 (2005).
[543x312], 595 (2021)
[552x756]22
[560x563], Quantum Science and Technology 3
[560x343]57
[573x489], 140403 (2018).
[582x709]measurement success probability with predetection squeezing and imperfect photon-number-resolution detectors, Phys. Rev. A 99, 032302 (2019)
[590x720], 030301 (2018).
[602x730], 260501 (2013)
[610x756]22
[618x500], 100501 (2005)
[627x521], 1 (2019).
[635x531]near deterministic creation of universal cluster states with probabilistic Bell measurements and three-qubit re-source states, Phys. Rev. A 91, 042301 (2015)
[640x584]with nondeterministic entangling gates, Phys. Rev. A 97.
[643x646], 020502 (2015).
[645x657]near-deterministic creation of universal cluster states with probabilistic Bell measurements and three-qubit re-source states, Phys. Rev. A 91, 042301 (2015)
[648x594]and D. E. Browne, Fault-tolerant quantum computation
with passive linear optics and unentangled ancillae, Phys. Rev. Lett. 113 (2014).
[660x678], 032302 (2022).
[667x552]operator and quantum teleportation, Phys. Rev. A 64, 052308
(2001).
[680x604], 042301 (2015).
[692x625]Near-deterministic creation of universal cluster states with probabilistic Bell measurements and three-qubit re-source states, Phys. Rev. A 91, 042301 (2015)
[695x636]and Greenberger-Horne-Zeilinger states to ballistic universal quantum computation, Phys. Rev. Lett. 115, 020502 (2015)
[706x667]T. Rudolph, From three-photon Greenberger-Horne-
Zeilinger states to ballistic universal quantum computation, Phys. Rev. Lett. 115, 020502 (2015)
[714x688], 031505 (2020).
[722x720], 033305 (2020).
[116x720], 033305 (2020).
[140x364], 052324 (2015).
[158x259], 140403 (2014).
[230x479], 052303 (2019).
[235x186], 041007 (2015).
[266x448], 060501 (2020).
[278x155], 315 (2021).
[286x552], 015005 (2017).
[305x730], 260501 (2013).
[317x542]\[30\] M. Pant, D. Towsley, D. Englund, and S. Guha, Perco-
[317x720], 033305 (2020).
[317x144], edited by G. Varoquaux, T. Vaught, and J. Mill-
[335x521]Commun. 10.1145/3505637 (2021).
[335x531]architectural requirements for generating universal pho-
[335x584]with nondeterministic entangling gates, Phys. Rev. A 97.
[335x646]Rev. Lett. 140403 (2018).
[341x217]et al. Fusion-based quantum
computing, arXiv preprint arXiv:2101.09310 (2021).
[345x374], 200502 (2010).
[363x144], Fusion-based quantum
computing, arXiv preprint arXiv:2101.09310 (2021).
[377x699]near determinism of universal cluster states with probabilistic Bell measurements and three-qubit re-source states, Phys. Rev. A 91, 042301 (2015)
[388x521], 1 (2019).
[395x730], 230 (1998)
[407x437], 39 (2005).
[421x374], 200502 (2010).
[436x604], 042301 (2015).
[457x625]Near-deterministic creation of universal cluster states with probabilistic Bell measurements and three-qubit re-source states, Phys. Rev. A 91, 042301 (2015)