Long-range data transmission in a fault-tolerant quantum bus architecture

We propose a fault-tolerant scheme for generating long-range entanglement at the ends of a rectangular array of qubits of length $R$ with a square cross-section of $m = O(\log^2 R)$ qubits. It is realized by a constant-depth circuit producing a constant-fidelity Bell-pair (independent of $R$) for local stochastic noise of strength below an experimentally realistic threshold. The scheme can be viewed as a quantum bus in a quantum computing architecture where qubits are arranged on a rectangular 3D grid, and all operations are between neighboring qubits. Alternatively, it can be seen as a quantum repeater protocol along a line, with neighboring repeaters placed at a short distance to allow constant-fidelity nearest-neighbor operations. To show our protocol uses a number of qubits close to optimal, we show that any noise-resilient distance-$R$ entanglement generation scheme realized by a constant-depth circuit needs at least $m = \Omega(\log R)$ qubits per repeater.

The ability of operating on spatially separated qubits has been recognized early on as a key requirement for a working quantum computer\(^1\)\(^2\). Long-range quantum communication is also a fundamental primitive at the core of quantum information-theoretic\(^3\) and cryptographic\(^6\)\(^7\) protocols in quantum networks, as well as e.g., distributed quantum computation\(^8\)\(^9\)\(^10\) or distributed sensing\(^11\)\(^12\). With increasingly powerful capabilities of present-day devices that can act as local nodes in such a network, finding appropriate protocols for entanglement sharing and quantum communication over faulty networks remains an active and central theme of present-day research.

Here we focus on long-range entanglement generation schemes which can be used in quantum computing devices where two-qubit operations are restricted to be geometrically local. This kind of restriction typically arises from technological constraints and is determined by the geometric arrangement of qubits in space as well as limits on the ability to couple distant qubits. Quantum algorithmic and fault-tolerance-related constructions, on the other hand, typically do not refer to notions of geometric locality. For example, while recent quantum low-density parity check (LDPC) codes are seen as excellent candidates for resource-efficient realizations of fault-tolerant quantum computing, their application involves long-range interactions when qubits are embedded in Euclidean space. A recent result implies that the number of such long-range gates needed scales at least polynomially with the size of the code\(^11\). This motivates the need to design protocols for realizing long-range interactions. Using teleportation, this goal can be rephrased as that of generating a long-range Bell pair in a short time with local operations. Any corresponding construction, such as ours, can be used as a quantum (data) bus in scalable quantum devices\(^1\)\(^4\).

Our proposal addresses the specific constraints on near-term devices by optimizing the following aspects:

(i) qubit efficiency versus distance: Our protocol fault-tolerantly generates a Bell pair on two qubits separated by distance $R$, using a grid of qubits of size $d \times d \times R$ with $d = O(\log R)$. That is, it uses only $O(R\log^2 R)$ qubits in total.

(ii) simplicity (geometric locality) of operations: Our scheme only involves one- and two-qubit gates between nearest-neighbor qubits in the architecture. The architecture could be realized, e.g., by stacking planar chip-like devices, and connecting neighboring devices through the third dimension.

(iii) resilience to general errors: we provide a rigorous proof of the robustness of our scheme in the form of a fault-tolerance threshold theorem allowing for fully general (circuit-level) local stochastic noise\(^15\). In this model, a random (multiqubit) Pauli error $E_t$ acts on the system at each time step $t$ of the quantum circuit. The probability that the Pauli error $E_t$ at each time step $t$ acts non-trivially on any subset of $k$ qubits is assumed to be at most $p^k$ for some $p \in [0, 1]$. The parameter $p$ is called the strength of the local stochastic noise, and we write $E_t \sim N(p)$. This notion only depends on the marginal distribution of $E_t$. In particular, our analysis captures situations where errors are correlated both in space and time, and we assume all involved operations (state preparation, gates and measurements) to be non-ideal, i.e., affected by noise. We note that — because of the simplicity of the scheme — we can explicitly specify and analyze all involved operations without resorting to fault-tolerant gadgets (e.g., for encoding and recovery) that are often treated in a black-box manner.

---

1Department of Mathematics, School of Computation, Information and Technology, Technical University of Munich, Garching, Germany. 2Munich Center for Quantum Science and Technology (MCQST), Munich, Germany. e-mail: choe.shinho@gmail.com
short coherence times and minimal latency: The required coherence time of each qubit is independent of the overall communication distance $R$. It only needs to be of the order of the local operation time $t_{\text{local}}$. This is the maximal time needed to perform a single- or two-qubit gate, a single-qubit measurement, or an entangling gate between two neighboring qubits.

Our protocol minimizes the amount of time $T_{\text{eq}}$ it takes to establish long-range entanglement from a product state of all qubits. The long-range entangled state is available after a time of order $T_{\text{eq}} = O(t_{\text{local}})$.

In the area of quantum computing architectures, our protocol could be especially beneficial in settings where gates or measurements need to be performed between spatially distant qubits. Such a need arises for example when using a quantum low-density parity check (LDPC) code whose stabilizer generators are not spatially local. With our scheme, we can design a quasi-2D architecture that allows for the application of joint measurements on any subset of size $\ell = O(1)$ of a total of $n$ qubits, see Fig. 1. Each such $\ell$-qubit measurement is executed with constant fidelity in a constant amount of time (up to Pauli correction whose classical postprocessing can typically be deferred to a later stage). The architecture requires $O(n \cdot \text{polylog}(n))$ qubits in total. When combined with the recently discovered good quantum LDPC codes that have parameters close to optimal\(^{16}\), this logarithmic overhead only leads to an encoding rate that decays as an inverse poly-logarithmic function.

Our scheme provides a candidate architecture for transferring quantum information within a quantum device, but can alternatively also be used to transfer information between devices. In this setting, one typically considers a line of $R$ equidistantly placed repeaters, where each repeater controls $m$ qubits, neighboring repeaters are connected by (imperfect) quantum communication as well as classical communication links, and the goal is to establish entanglement between the first and the last repeater (separated by a distance proportional to $R - 1$). Assuming that every repeater has full control over its qubits, i.e., can apply any quantum circuit, this setup allows (by teleportation between repeaters) to imperfectly realize an adaptive quantum circuit on $mR$ qubits which only uses gates between pairs of qubits associated with neighboring repeaters, and gates acting only on its own qubits. Any repeater protocol for distance-$R$ entanglement generation amounts to a circuit of this form, and, conversely, circuits with these restrictions can be seen as repeater protocols.

In our architecture, qubits are located on the sites of a slab-like 3D rectangular lattice of dimensions $d \times d \times R$. We can define, for each $j \in \{1, \ldots, R\}$, a repeater $R_j$, which is composed of the qubits in the “slice” $[1, \ldots, d] \times [j]$. Then our circuit construction immediately translates into a repeater protocol with the following desirable features for each repeater:

(a) long communication distance: As a function of the total communication distance $R$, every repeater only needs to manipulate $m = O(\log R)$ number of qubits. This is in line with the best currently known proposals\(^{17,18}\), which have poly-logarithmic scaling for long-range quantum communication. We also provide a converse bound $\Omega(\log R)$ on the number of qubits per repeater for low-latency protocols as considered in this work.

(b) simplicity of involved operations: Assuming the repeaters are connected appropriately (by $m$ links along the direction of communication), all operations at each repeater $R_j$, $j \in \{1, \ldots, R\}$ are geometrically local when the $m = O(\log R)$ qubits are arranged on a 2D array. In contrast, in existing works on repeater-based quantum communication, repeaters are often assumed to provide universal quantum computing power locally, and no geometric locality considerations are taken into account. For example, in ref. 19, the authors discuss entanglement distribution protocols which use quantum error-correcting codes\(^{22}\). This family of protocols proceeds by communicating encoded quantum information: Repeaters apply a recovery operation to each codeblock before retransmitting the entire block to the next repeater. This operation is arguably more involved that the operations required in our scheme.

Purely error-correction based schemes have a loss-tolerance limited by 50% because of the no-cloning theorem. To sidestep this limitation, the more recent work\(^{25}\) combines error correction with hashing-based entanglement purification. In this scheme, intermediate error correction operations are realized based on stabilizer (resource) states prepared at each repeater. The proposed scheme achieves a constant rate of qubit transmission. Here error correction is based on measurement-based gate teleportation using special (graph) states encoding the recovery map at each repeater location. While such schemes are ultimately desirable, the generation of these resource states is arguably more challenging to realize with near-term devices than the protocol discussed here. In particular, protocols preparing these resource states require long-range gates and non-constant circuit depth in each repeater. These operations are more involved that those in our scheme, and, in particular, do not give a low-latency scheme because of the extensive circuit depth.

(c) resilience to full circuit-level local stochastic noise: It is sufficient if each repeater realizes circuits up to local stochastic noise. In particular, no fault-tolerant quantum computer is needed at any one of the repeaters. In contrast to our circuit-level analysis, existing proposals often resort to the black-box use of fault-tolerant gadgets (e.g., for encoding and recovery), essentially amounting to the assumption that each repeater has full quantum fault-tolerance capabilities. Furthermore, existing works on quantum repeaters argue that faulty operations can be modeled by erroneous communication links (or corresponding Bell states) by commuting errors forward respectively backward in time. Errors arising in this way are generally not going to identically and independently distributed. Nevertheless, the provided analysis in existing works typically assumes this i.i.d. form of the noise. Our work overcomes this mismatch between the actually occuring noise and the error analysis by avoiding such simplistic assumptions altogether.

(d) constant coherence times and minimal latency: Every repeater only needs to maintain coherence for a time of order $t_{\text{local}}$. Furthermore, the long-range entangled state is available after a time of order $T_{\text{eq}}$. Then our circuit construction immediately translates into a repeater protocol with the following desirable features for each repeater:

\[ T_{\text{eq}} = O(t_{\text{local}}). \]
Communication applications, e.g., transmission of qubits over an optical fiber, are challenging. In the optical setup, it is typically assumed that the loss (or error) rate then depends on the distance between repeaters by means of extra processes such as entanglement distillation. The complexity of these distillation steps precludes the possibility of a generation scheme having the low-latency property (d). Thus, there is a non-zero constant $T_{\text{ent}}$ such that each measurement result is independent. Then, the probability that the distance between repeaters is limited, i.e., they cannot be arbitrarily far apart. In this case, low-latency long-distance entanglement is possible.

Our setup (as described by (c)) differs in another important aspect from commonly made assumptions in quantum communications. Our analysis also applies to correlated errors. This makes our protocol particularly attractive for using it as a quantum bus (rather than for point-to-point communication), as quantum computations can propagate errors leading to correlations between different timesteps. Analysing only independent errors will generally not be sufficient to ensure the protocol is suitable for use in noisy quantum computational devices.

To motivate the problem of fault-tolerant long-distance entanglement generation with low latency, let us recall entanglement swapping. We consider $R$ sites (or “repeaters”) lying on a line, where each pair of neighboring sites shares a Bell state $|\Phi\rangle := (|00\rangle + |11\rangle)/\sqrt{2}$, see Fig. 2(a). Suppose we measure two qubits at each site in the Bell basis. Then, the resulting two-qubit state on two opposite sites on the line is one of the Bell states.

Entanglement swapping circuit. It generates a Bell state of two qubits indexed by $q_1$ and $q_2$ with certainty, but it is not fault-tolerant under noise, such as i.i.d. measurement error. (b) Our protocol. Each repeater has $m$ qubits. The gates $H$ and $CZ$ are tensor products of Hadamard gates and controlled-$Z$ gates, respectively, on certain subsets of qubits clarified in Result.
of the corresponding architecture is that it encodes logical information into a surface code at each repeater. We emphasize, however, that this only motivates the scheme in a similar way, e.g., cluster-state measurement-based quantum computation is related to braiding surface code anyons. For example, the scheme does not involve intermediate encoding/recovery operations associated with these surface codes.

**Results**

Our work builds on the fault-tolerance properties of the 3D cluster state. In seminal work ref. 36, Rausendorf, Bravyi and Harrington proved that for temperatures below some threshold, the thermal state of the cluster state Hamiltonian on a simple 3D cubic lattice has localizable entanglement: applying single-qubit Pauli measurements to each “bulk” qubit results – up to Pauli corrections determined by the measurement outcomes – in an encoded Bell state of 2D surface codes lying on two opposing faces of the cube. Since the initial thermal state is the cluster state with independent and identically distributed Pauli-Z-errors on each qubit, and the cluster state can be created in constant depth with local operations, this provides a scheme for long-range (encoded) entanglement generation by a constant-depth circuit with local gates. While the assumed (physically motivated) noise model is somewhat restricted, the authors of ref. 36 also point out that more general errors on the initial cluster state such as independent single-qubit depolarizing noise can be tolerated when using ideal (non-faulty) operations at the two surfaces.

The scheme of ref. 36 produces a surface-code encoded Bell pair, i.e., the resulting entanglement lives in a logical 2-qubit subspace of a many-qubit system. In applications such as the one considered in ref. 37, this is a desirable feature: the encoded information can be operated on without decoding, thus preserving fault-tolerance properties. In a communication scenario, however, it may be more desirable to establish entanglement between two physical qubits instead. A scheme for achieving this kind of localization of the entanglement was proposed in ref. 38 by modifying the decoding, thus preserving fault-tolerance properties. In a communication setting, the protocol describes a procedure of entanglement sharing between the two end nodes in a 1D network. Considering the slice of the lattice C containing the qubit q1 as the first slice (associated with repeater R1), we call the j-th slice the repeater Rj.

Let n = Θ(dR) be the number of vertices of the lattice C. The protocol starts by preparing the 3D cluster state W[0]⊗n of the lattice C, where

\[ W = H^Ω_{[4]} \prod_{(i,j)} CZ_{u,v} H^Ω_{[4]} \]  

(1)

is a product of Hadamard gates and controlled-Z gates

\[ H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \text{and} \quad CZ_{u,v} = |0\rangle\langle 0|_u \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}_v + |1\rangle\langle 1|_u \otimes \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}_v, \]

and the product of controlled-Z gates CZ_{u,v} is taken over all edges \{u, v\} of the lattice C.

We first observe that the product Π_{i=1}^{n} CZ_{u,v} of controlled-Z gates in Eq. (1) can be written as a depth-6 quantum circuit as described in Figs. 4 and 5. Here the operations in Fig. 4 are "local" for each repeater, and are applied within each repeater Rj in parallel for j = 1, ..., R. The operations illustrated in Fig. 5 couple neighboring odd-even pairs \{(2k−1)R, 2kR\}, k = 1, ..., \lfloor R/2 \rfloor of repeaters, and even-odd pairs \{(2k−1)R, 2kR+1\}, k = 1, ..., \lfloor R/2 \rfloor of repeaters in two rounds.

In more detail, the first four layers consist of CZ gates on the edges connecting two qubits in the same node. The fifth layer consists of CZ gates acting on \{(2j−1)R and 2jR\} repeaters for j = 1, ..., \lfloor R/2 \rfloor (see solid lines in Fig. 5), and the sixth layer consists of CZ gates acting on 2jR and (2j+1)R repeaters for j = 1, ..., \lfloor R/2 \rfloor (see dashed lines in Fig. 5). In summary, we can write W as

\[ W = H^Ω_{[4]} U'_6 U'_5 U'_4 U'_3 U'_2 U'_1 H^Ω_{[4]}, \]

where all U'j for j = 2, ..., 5 together with U'_6 and U'_6 consist of CZ gates. Note that this depth-6 circuit is also illustrated in Fig. 2(b).
Fig. 3 | A cluster state on the lattice $C[d \times d \times R]$ with $R = 29 \gg d = 4$. Each sphere represents a qubit. Also indicated is the measurement pattern used in our protocol, consisting of single-qubit Pauli-$X$ and Pauli-$Z$ measurements. Entanglement is established between the qubits $q_1$ and $q_2$. The set of qubits is partitioned into three subsets, namely $\mathcal{X}$, $\mathcal{Z}$ and $\{q_1, q_2\}$. In the course of the protocol, the qubits in the subsets $\mathcal{X}$ and $\mathcal{Z}$ are measured in the Hadamard- and computational basis, respectively.

Fig. 4 | Sequence of quantum operations applied within each individual repeater (slice). Each CZ is a product of CZ gates on the black edges connecting the qubits (white balls). The gray edges of the graph at each layer represent the CZ gates applied in the previous steps. Each repeater runs this procedure in parallel. In the end, each repeater has a planar graph state.

Fig. 5 | Quantum operations applied to nearest-neighboring repeaters consisting of two steps after the process illustrated in Fig. 4. The first step consists of two-qubit CZ operations between repeater $\mathcal{R}_{2j-1}$ and repeater $\mathcal{R}_{2j}$ (solid lines) for all $j = 1, \ldots, \lfloor R/2 \rfloor$. The second step is of two-qubit CZ operations between repeater $\mathcal{R}_{2j}$ and repeater $\mathcal{R}_{2j+1}$ (dashed lines) for all $j = 1, \ldots, \lfloor R/2 \rfloor$. 

https://doi.org/10.1038/s41534-024-00928-4
We can further merge the layers $H^n$ and $U_i$ into one layer $U_1 = U_1'H^n$, which is a product of two-qubit gates $CZ_{a,v}H_a$ acting on different qubits. Analogously, we define the unitary $U_6 = H^nU_6'$ as the layer obtained by merging two layers $U_6^T$ and $H^n$. Finally, we can write $W$ as a depth-6 Clifford circuit

$$W = U_6 \cdots U_1,$$

where each $U_j$ for $j = 1, \ldots, 6$ consists of $CZ_{a,v}H_aCZ_{a,v}$ and $CZ_{a,v}H_a$ gates acting on pairwise disjoint qubits $\{u, v\}$.

To describe the remaining steps of the protocol, we consider the partition $X \cup Z \cup \{q_1, q_2\}$ of these $n$ qubits which is represented in Fig. 3. With the cluster state $W|0\rangle^n$, we perform single-qubit measurements on the qubits except for two qubits $q_1$ and $q_2$. Here, each qubit in $X$ and $Z$ is measured in the Hadamard- and computational basis, respectively.

We can construct functions $\alpha, \beta : \{0, 1\}^X \times Z \rightarrow \{0, 1\}$ such that given the measurement results $s \in \{0, 1\}^X \times Z$, the post-measurement state on $q_1$ and $q_2$ is $Z^s_{q_1}X^0_{q_2}|\Phi\rangle_{q_1,q_2}$, if all operations are perfect. Therefore, given the two bits $\alpha(s)$ and $\beta(s)$, one can apply the Pauli “correction” operator $Z^\alpha X^\beta$ to qubit $q_2$ to obtain the desired state $|\Phi\rangle_{q_1,q_2}$. Note that for many applications which can be written as Clifford circuits, it is sufficient to compute the two bits $\alpha(s)$ and $\beta(s)$ to correct the Pauli operator in “software”, i.e., one can commute the Pauli correction $Z^\alpha X^\beta$ through a stabilizer circuit without the need for physically applying a correction.

Before we move on to the precise description of the components of the protocol, we observe that the protocol satisfies the low-latency and geometric locality requirements discussed in the Introduction: The whole process can be written in a constant-depth quantum circuit with Pauli gates controlled by classical bits obtained by computing the bits $\alpha(s)$ and $\beta(s)$ (see Fig. 2(b)). Moreover, the involved qubits are arranged in a slab-like 3D lattice, as shown in Fig. 3, and operations are between neighboring qubits only. A precise description of the lattice $[d \times d \times R]$ as well as the validity of the protocol is deferred to Sections 4 and 5 of the Supplementary Information.

Here we state the definition of $\alpha$ and $\beta$ and discuss when the protocol fails to output the correct Bell basis state in the scenario where the protocol runs with the noisy cluster state $EW|0\rangle^n$ corrupted by some $n$-qubit random Pauli $E$ and all measurements are perfect. (We also discuss more general circuit-level noise below.)

First, note that since all qubits in $X$ and $Z$ are measured in Hadamard basis and computational basis, respectively, we may assume that the random Pauli $E$ is of the form

$$E = E_{q_1,q_2}X_{q_1}Z_{q_2},$$

where $E_{q_1,q_2}$ is a random Pauli supported on the qubits $q_1$ and $q_2$, $E_{q_1}$ is an $X$-type random Pauli supported on $X$, and $E_{Z}$ is a $Z$-type random Pauli supported on $Z$. In particular, from now on, we think of $E_{q_1}$ and $E_{Z}$ respectively as a (random) subset of $X$ and $Z$.

The main subroutine contained in the (algorithmic) definition of $\alpha$ and $\beta$ is solving minimum matching problems on the (primal) and dual decoding graphs

$$T_{d,\text{dec}} = (V_{d,\text{dec}}, E_{d,\text{dec}}) \quad \text{and} \quad T_{d,\text{dec}^*} = (V_{d,\text{dec}^*}, E_{d,\text{dec}^*})$$

illustrated in Fig. 6. The full definitions of the decoding graphs are deferred to Section 4 of the Supplementary Information, but here we state all properties of these graphs to give the definition of $\alpha$ and $\beta$.

1. Each qubit in $X \cup Z$ corresponds to one of the edges of $T_{d,\text{dec}}$ or $T_{d,\text{dec}^*}$. With this correspondence, the subset $E_{X} \cup E_{Z} \subseteq X \cup Z$ can be represented as a subset of the edges of the decoding graphs. Let us denote these subsets of the edges by

$$E_{\text{dec}} = (E_{X} \cup E_{Z}) \cap E_{d,\text{dec}}, \quad E_{\text{dec}^*} = (E_{X} \cup E_{Z}) \cap E_{d,\text{dec}^*}.$$

2. Each vertex of the decoding graphs is either internal or external. The set of external vertices of a decoding graph $T_{d,\text{dec}}$ is partitioned into two “membranes” $M_1$ and $M_2$. We collect all edges connecting the vertices in $M_1$ with internal vertices and let $L_{d,\text{X}}$ denote the set of these edges. We analogously define the membranes $M_{1^*}$ and $M_{2^*}$ of the other decoding graph $T_{d,\text{dec}^*}$ and the set of edges $L_{d,Z}$. 

Fig. 6 | The decoding graph $T_{d,\text{dec}}$ as well as the dual decoding graph $T_{d,\text{dec}^*}$. External vertices (black spheres) and internal vertices (white spheres) are illustrated. The external vertices incident to the edges in $L_{d,\text{X}}$ and $L_{d,\text{Z}}$ form the membranes $M_1$ and $M_2$. Labels on the external vertices are used in Supplementary Information for the analysis on the failure probability of the protocol.
Given the measurement result \( s \in \{ 0, 1 \}^{X \cup Z} \), the bits \( \alpha(s) \) and \( \beta(s) \) are computed as follows:

1. Considering \( s \) as a subset of the set of edges \( E_{\text{dec}} \cup E_{\text{dec}'} \), partition \( s \) into two subsets
   \[
   s = s_{\text{dec}} \cup s_{\text{dec}'} \quad \text{where} \quad s_{\text{dec}} = s \cap E_{\text{dec}}, \quad s_{\text{dec}'} = s \cap E_{\text{dec}'}.
   \]

2. Compute the boundaries \( \partial_{T_{\text{dec}} \cup T_{\text{dec}'}}(s_{\text{dec}}) \) and \( \partial_{T_{\text{dec}} \cup T_{\text{dec}'}}(s_{\text{dec}'}) \) of \( s_{\text{dec}} \) and \( s_{\text{dec}'} \) in the decoding graphs \( T_{\text{dec}} \) and \( T_{\text{dec}'} \). Here, the boundary \( \partial_T(E) \) of a subset \( E \subset E \) of edges in a decoding graph \( T \) is defined as the set of internal vertices of \( T \) incident to odd number of edges in \( E \).

3. Compute minimal matchings
   \[
   m = \text{MinMatch}_{\partial_{T_{\text{dec} \cup T_{\text{dec}'}}}(s_{\text{dec}})}(\partial_{T_{\text{dec} \cup T_{\text{dec}'}}}(s_{\text{dec}})) \quad m^* = \text{MinMatch}_{\partial_{T_{\text{dec} \cup T_{\text{dec}'}}}(s_{\text{dec}'})}(\partial_{T_{\text{dec} \cup T_{\text{dec}'}}}(s_{\text{dec}'}))
   \]

   of the internal vertices \( \partial_{T_{\text{dec}}} (s_{\text{dec}}) \) and \( \partial_{T_{\text{dec}'} \cup T_{\text{dec}'}}(s_{\text{dec}'}) \). Here, a minimal matching \( \text{MinMatch}_{\partial_T}(V) \) of a subset \( V \) of internal vertices in a decoding graph \( T \) is a minimal subset of edges of \( T \) whose boundary is \( V \).

4. Finally, compute the bits
   \[
   \alpha(s) = \left\langle s_{\text{dec}} \otimes m, L_{\text{d}, X} \right\rangle, \quad \beta(s) = \left\langle s_{\text{dec}' \cap m^*}, L_{\text{d}, X}^* \right\rangle.
   \]

Here, for two sets \( A \) and \( B \), we denote \( A \oplus B \) the symmetric difference of \( A \) and \( B \), or equivalently the bit-wise summation over modulo 2 if all sets are considered as bitstrings. The notation \( \langle A, B \rangle \) denotes the parity of the number of elements in the intersection \( A \cap B \), or equivalently the usual inner product on the bitstring defined as the parity of the sum of the bits obtained by bit-wise multiplication of \( A \) and \( B \) which are considered as bitstrings.

A key fact here is that the two subsets \( s_{\text{dec}} \) and \( E_{\text{dec}} \) have the same boundaries in the decoding graph \( T_{\text{dec} \cup T_{\text{dec}'}} \), i.e.,
\[
\partial_{T_{\text{dec} \cup T_{\text{dec}'}}}(s_{\text{dec}}) = \partial_{T_{\text{dec} \cup T_{\text{dec}'}}}(E_{\text{dec}}).
\]

Analogously, we have
\[
\partial_{T_{\text{dec} \cup T_{\text{dec}'}}}(s_{\text{dec}'}) = \partial_{T_{\text{dec} \cup T_{\text{dec}'}}}(E_{\text{dec}'})
\]

for the subsets \( s_{\text{dec}} \) and \( E_{\text{dec}} \) of the edges of the decoding graph \( T_{\text{dec} \cup T_{\text{dec}'}} \). In particular, the boundaries of the errors \( E_{\text{dec}} \) and \( E_{\text{dec}'} \) can be computed from the measurement outcome \( s \). The main role of the function \( \alpha \) is finding the estimate
\[
\hat{E}_{\text{dec}}(s) := m = \text{MinMatch}_{\partial_{T_{\text{dec} \cup T_{\text{dec}'}}}(s_{\text{dec}})}(\partial_{T_{\text{dec} \cup T_{\text{dec}'}}}(s_{\text{dec}}))
\]

of \( E_{\text{dec}} \) with the information about its boundary in the decoding graph by solving a minimum matching problem on the decoding graph \( T_{\text{dec} \cup T_{\text{dec}'}} \). Analogously, a subroutine of \( \beta \) computes a minimal matching
\[
\hat{E}_{\text{dec}'}(s) := m^* = \text{MinMatch}_{\partial_{T_{\text{dec} \cup T_{\text{dec}'}}}(s_{\text{dec}'})}(\partial_{T_{\text{dec} \cup T_{\text{dec}'}}}(s_{\text{dec}'}))
\]

of the boundary of \( E_{\text{dec}'} \) associated with the dual decoding graph \( T_{\text{dec} \cup T_{\text{dec}'}} \). With Edmond’s algorithm\(^{1,11,12}\), if the repeater \( k \), which contains \( q_2 \), receives all measurement results from all other \( k - 1 \) repeaters, then it can evaluate the estimates \( \hat{E}_{\text{dec}}(s) \) and \( \hat{E}_{\text{dec}'}(s) \) in a time \( O(\hat{S} \cdot R^4) \). The “flow” of the classical information, which is the measurement results, is illustrated as double line segments in Fig. 2).

A necessary condition for the failure of the protocol, i.e., the case where the output state is not the desired Bell state \( |\Psi \rangle \), is that the Pauli \( E_{\text{dec}} \) on the qubits \( q_1 \) and \( q_2 \) is non-trivial. Geometrically, this implies that at least one of the subsets \( E_{\text{dec}} \oplus E_{\text{dec}'}(s) \subset E_{\text{dec} \cup T_{\text{dec}'}} \) or \( E_{\text{dec}'} \oplus E_{\text{dec}}(s) \subset E_{\text{dec} \cup T_{\text{dec}'}} \) of the edges connects two membranes \( M_1 \) and \( M_2 \) or \( M_1 \) and \( M_2' \), of the decoding graphs at the opposite sides, consisting of external vertices, see Fig. 7. The proof of the necessity of this condition for failure of the protocol is deferred to Section 4 of the Supplementary Information.

A threshold theorem for low-latency long-range entanglement generation
We prove that the entanglement generation protocol using a 3D cluster state is robust under local stochastic noise of strength below a threshold, even with a communication distance \( R \gg d \). The formal definition of local stochastic errors is as follows.

**Definition 1.** (Local stochastic errors) Let \( |n \rangle := \{ n \}, \ldots, \delta \rangle \) denote the set of \( n \) qubits. A random \( n \)-qubit Pauli \( E \) is local stochastic with strength \( p \in [0, 1] \) if
\[
\Pr[S \subseteq \text{supp}(E)] \leq p^{|S|} \quad \text{for all} \quad S \subseteq |n\rangle.
\]

where \( \text{supp}(E) \) is the set of all qubits where \( E \) acts non-trivially. We write \( E \sim \mathcal{N}(p) \) if \( E \) is local stochastic with strength \( p \).
We first state our result on fault-tolerance of the protocol in the case where the protocol runs with the noisy cluster state $E W | 0 \rangle^{\otimes n}$ where $E \sim \mathcal{N}(p)$ is a random Pauli which is local stochastic with strength $p$, and there is no other source of error.

**Theorem 2.** (Long-range entanglement generation). Consider a cluster state $W | 0 \rangle^{\otimes n}$ associated with the lattice $\mathcal{C}[d \times d \times R]$ where $d \geq 3$ is arbitrary. Suppose the measurement pattern in Fig. 3 is applied to a corrupted cluster state $E W | 0 \rangle^{\otimes n}$, where $E \sim \mathcal{N}(p)$ is a local stochastic error. Then there are two efficiently computable functions $\alpha, \beta$ taking the collection $s$ of measurement outcomes to bits such that the post-measurement state on qubits $q_1, q_2$ is the state $Z_{q_1}^{\otimes s} W_{q_2}^{\otimes} \Phi$ with probability at least $1 - 5000p$, for any $0 < p \leq p_0 := \frac{1}{5000} \approx 2 \times 10^{-4}$ and any $R \leq \frac{1}{\sqrt{d}} \left( \frac{1}{10 \sqrt{p}} \right)^{d-2}$.

Here, we only give a sketch of the proof of Theorem 2 presented in Section 5 of the Supplementary Information. In the proof of Theorem 2, we show that for a random Pauli $E \sim \mathcal{N}(p)$, the subset $E_{d, \text{dec}} \oplus E_{\text{dec}}(s)$ of the edges contains a path connecting the two membranes of the decoding graph $T_{d, \text{dec}}$ at the opposite sides with probability at most $O(p)$. We also show a similar result for the subset $E_{\text{dec}} \oplus E_{\text{dec}}(s)$ of $E_{d, \text{dec}}$. The upper bound on the failure probability is obtained by applying the union bound to these events.

We also show that fault-tolerance of the protocol with this simplified noise model (with noise only on the cluster state $W | 0 \rangle^{\otimes n}$, i.e., Theorem 2 implies that the protocol is robust under circuit-level local stochastic noise of strength below some constant threshold where Pauli error occurs before and after every layer of the state preparation, unitary operation, and single-qubit measurements, see the Methods section.

**Optimality: A converse to the threshold theorem.** Considering each $(2d - 1) \times (2d - 1)$-slice of the lattice $\mathcal{C}[d \times d \times R]$ as a repeater, we have $m = O(d^2)$ qubits per repeater. Theorem 2 thus implies that below the threshold error strength $p_0$,

$$m = O(\log^2 R)$$

qubits per repeater are sufficient to establish constant-fidelity over distance $R$. Thus, the cluster-state-based protocol satisfies the properties discussed in the introduction.

In this subsection, we study to which extent the protocol provided by Theorem 2 is optimal among protocols with low latency. Specifically, we ask if the relationship between the number $m$ of qubits at each site and the communication distance $R$ can be better than polylogarithmic (cf. (2)). We find that this is not the case, i.e., the scaling of our protocol is essentially optimal.

**Theorem 3.** (Converse for low latency schemes). Let $n$ be an entanglement generation of low-latency over communication distance $R$ which uses $m$ qubits per repeater. If its output state achieves fidelity at least $\frac{1}{2} (1 + e^{-c}) \approx 0.684$ under arbitrary local stochastic noise of strength $p$ below a constant threshold $p_0 > 0$, then we must have $m = \Omega(\log R)$.

The proof of Theorem 3 replaces the analysis of a low-latency protocol $\pi$ by that of an entanglement-assisted protocol $\pi'$, and follows from the fact that for any $p \in (0, 1)$, the qubit depolarizing channel $E_{\text{dec}}(p) = (1 - p) I + p \rho(\pi')$ is a convex combination of an entanglement-breaking channel and the identity channel (see Supplementary Information Section 6 for details). It implies that local stochastic noise with any constant noise strength prevents entanglement generation over distances greater than exponential in the number of qubits at each site. Comparing the lower bound from the converse theorem with the number of qubits per site from our protocol (2), we conclude that the polylogarithmic scaling on the number $m$ of Theorem 2 can at best be improved to a logarithmic scaling in $R$.

Closing the log-factor gap between the converse Theorem 3 and achievability Theorem 2 appears to require a new approach, however: We can show that the analysis of Theorem 2 leading to the scaling (2) is tight. In fact, the scaling is necessary for any procedure which is based on the cluster state and uses the same syndrome information using functions $\alpha$ and $\beta$ as in our protocol: We show the lower bound

$$m = \Omega(\log^2 R)$$

for any such protocol. This result is obtained by a detailed analysis of the associated decoding problem, see Supplementary Information Section 6. It applies, in particular, to natural variants of our scheme that are obtained, e.g., by replacing Edmonds' minimal matching algorithm by other procedures, including heuristic algorithms. This indicates that improving over our protocol will require a different resource state or measurement pattern.

**Understanding the protocol with surface code single-shot decoding.** We give a modular way of understanding the protocol, giving a high-level overview focused on the two main building blocks:

(I) **Generating a surface-code-encoded Bell pair:** The generation of a cluster state on an elongated lattice $\mathcal{C}[d \times d \times R]$ and subsequent single-qubit measurement of “bulk” qubits not belonging to the two boundaries with third coordinate equal to $u_1 = 1$ and $u_1 = R$, respectively. Up to local unitaries, this part of the procedure is identical to the original scheme $^{[3]}$. This generates a surface-code encoded Bell pair up to a (computable) Pauli correction and residual local stochastic noise $^{[3]}$, Result 3).

(II) **Single-shot decoding of the surface code:** A single-shot decoding procedure for transferring surface-code encoded quantum information onto a single physical qubit. We note that the same idea was used in $^{[4]}$ for logical $T$-basis measurement. We show that this procedure is resilient to local stochastic noise: There is a threshold $p_0^\text{dec}$ such that if the initial encoded state $\Psi$ is corrupted by a local stochastic noise with strength $p \leq p_0^\text{dec}$, then the fidelity of the resulting single-qubit state with the logical state $\Psi$ is linearly close to 1 as a function of the noise strength $p$. We refer to Supplementary Information Section 1 and 3 for details.

From a conceptual point of view, it is clear that combining (I) with (II) results in a procedure for generating long-range entanglement. However, our analysis of the resulting fidelity is not simply a combination of the analytical results stated above for each building block (I) and (II). This is because a threshold result derived in this way does not yield parameters of practical relevance. Instead, we establish a significantly tighter threshold estimate by a direct analysis of the achieved fidelity for the entire (combined) process.

**Discussion**

In this work, we give a complete analysis of the robustness of the distance-$R$ entanglement generating protocol based on the 3D cluster state of size $(2d - 1) \times (2d - 1) \times R$ with $d = O(\log R)$ under circuit-level local stochastic noise. Our main result is a fault-tolerance threshold theorem giving a bound on the tolerable noise strength.

We also show that our protocol based on the 3D cluster state gives rise to a low-latency protocol for long-range entanglement generation by a set of $R$ equidistant repeaters on a line. We prove that fault-tolerant generation of one Bell pair between two end nodes by any low-latency protocol requires at least $m = \Omega(\log R)$ qubits per node. Our construction uses $m = O(\log^2 R)$ qubits at each repeater, which we show is optimal for the 3D cluster-state approach. We note that it is still an open problem whether there exists a low-latency protocol which saturates the lower bound $m = \Omega(\log R)$.

**Methods**

In this section, we explain how to reduce the problem of showing fault-tolerance of low-latency protocols with circuit-level noise to the setting of a
noisy cluster state $EW|0\rangle^\otimes$ (and noise-free operations otherwise). We note that this reduction is same as that of ref. 37, and it is ubiquitous in the error correction literature, see e.g., refs. 15,45.

Recall that the protocol we proposed (without the correction operator) is essentially preparing the state $W|0\rangle^\otimes$ followed by the single-qubit measurements, where $W$ is a Clifford unitary

$$W = U_\Delta \cdots U_1$$

of depth $\Delta = 6$, where each $U_i$ is a tensor product of single- or two-qubit Clifford unitaries whose supports are disjoint.

A noisy implementation of the whole protocol with noise strength $p \in [0,1]$ can be described as the protocol performing the perfect single-qubit measurements on a noisy input state

$$E_{\Delta+1} E_A U_{\Delta} E_{\Delta-1} \cdots E_1 U_0 E_0|0\rangle^\otimes$$

with $\Delta = 6$. Here, all $E_j \sim N(p)$ are local stochastic errors of strength $p$ on all qubits. In particular, $E_0$ denotes the Pauli error caused by noisy state preparation, and $E_{\Delta+1}$ denotes the Pauli error which is introduced to describe the measurement error.

Proving the robustness of the protocol with a noisy state of the form (3) seems to be complicated. However, we can sidestep this problem with the following observation: Since all $U_i$ are Clifford unitaries, we can “commute” all Pauli errors $E_j$ to the left. For example, there exists a Pauli $E_0$ such that $E_0 U_1 = U_1 E_0$, i.e., $E_0 = U_1 E_0 U_1$ is obtained by conjugating $E_0$ with the Clifford unitary $U_1$. Then, consider the Pauli $E_0 E_0$ and apply the similar technique with the Clifford unitary $U_2$. By repeating this procedure, we end up with some Pauli $E$ such that the noisy state (3) is equal to

$$EW|0\rangle^\otimes.$$

More importantly, $E \sim N(p')$ is also a local stochastic error of strength $p' = 4p^{11}$. This fact follows from [37, Lemma 12]: The main idea here is using the commuting technique described above with the following update rules of the noise strength: If $E \sim N(p)$ and $E' \sim N(q)$ are Paulis and $U$ is a depth-1 Clifford unitary consisting of one- and two-qubit gates, then $E \cdot E' \sim N(2\max\{\sqrt{p}, \sqrt{q}\})$ and $UEU^\dagger \sim N(\sqrt{2}p')$, see also [10, Lemma 11] for the proof of the update rules. Therefore, if we show that the the protocol is robust for a noisy initial cluster state $EW|0\rangle^\otimes$ (see Eq. (4)), where the only error during the whole process is the random Pauli $E \sim N(p')$ with $p' \leq p_0$, then we can say that the protocol is robust with the circuit-level noise model with random Paulis $E_j \sim N(p)$ in Eq. (3) for all $p \leq p_0$ with some constant threshold $p_0 > 0$.

Data availability

Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

Received: 21 December 2023; Accepted: 2 December 2024; Published online: 26 December 2024

References

1. DiVincenzo, D. P. The physical implementation of quantum computation. Quantum. 48, 771–783 (2000).
2. Preskill, J. Quantum computing: pro and con. Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences 454, 469–486 (1998).
3. Gisin, N. & Thew, R. Quantum communication. Nature Photonics 1, 1749–4893 (2007).
4. Kimble, H. The quantum internet. Nature 453, 1476–4687 (2008).
5. Wehner, S., Elkouss, D. & Hanson, R. Quantum internet: A vision for the road ahead. Science 326, 890988 (2018).
6. Bennett, C. & Brassard, G. Quantum cryptography: Public key distribution and coin tossing. Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing, Bangalore 175–179 (1984).
7. Gisin, N., Ribordy, G., Tittel, W. & Zbinden, H. Quantum cryptography. Rev. Mod. Phys. 74, 145–195 (2002).
8. Cirac, J. I., Ekert, A. K., Huelga, S. F. & Macchiavello, C. Distributed quantum computation over noisy channels. Phys. Rev. A 59, 4249–4254 (1999).
9. Meter, R. V., Munro, W. J., Nemoto, K. & Itoh, K. M. Arithmetic on a distributed-memory quantum multicomputer. J. Emerg. Technol. Comput. Syst. 3 (2008).
10. Robert, B. et al. Efficient distributed quantum computing. Proc. R. Soc. A 469 (2013).
11. Kómár, P. et al. A quantum network of clocks. Nature Physics 10, 582–587 (2014).
12. Eldredge, Z., Foss-Feig, M., Gross, J. A., Rolston, S. L. & Gorshkov, A. V. Optimal and secure measurement protocols for quantum sensor networks. Phys. Rev. A 97, 042337 (2018).
13. Baspin, N. & Krishna, A. Connectivity constrains quantum codes. Quantum 6, 711 (2022).
14. Brennen, G. K., Song, D. & Williams, C. J. Quantum-computer architecture using nonlocal interactions. Phys. Rev. A 67, 050302 (2003).
15. Gottesman, D. Fault-tolerant quantum computation with constant overhead (2013). https://arxiv.org/abs/1310.2984.
16. Panteleev, P. & Kalachev, G. Asymptotically good quantum and locally testable classical ldpc codes. In Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2022, 375–388 (Association for Computing Machinery, New York, NY, USA, 2022). https://doi.org/10.1145/3519935.3520017.
17. Fowler, A. G. et al. Surface code quantum communication. Phys. Rev. Lett. 104, 180503 (2010).
18. Muralidharan, S., Kim, J., Lütkenhaus, N., Lukin, M. D. & Jiang, L. Ultrafast and fault-tolerant quantum communication across long distances. Phys. Rev. Lett. 112, 250501 (2014).
19. Muralidharan, S. et al. Optimal architectures for long distance quantum communication. Scientific Reports 6 (2016).
20. Muralidharan, S., Zou, C.-L., Li, L., Wen, J. & Jiang, L. Overcoming erasure errors with multilevel systems. New Journal of Physics 19, 013026 (2017).
21. Munro, W., Stephens, A., Devitt, S., Harrison, K. & Nemoto, K. Quantum communication without the necessity of quantum memories. Nature Photonics 6 (2013).
22. Zwerver, M., Pirker, A., Dunjko, V., Briegel, H. J. & Dür, W. Long-range big quantum-data transmission. Physical Review Letters 120, 030503 (2018).
23. Dai, W., Peng, T. & Win, M. Z. Optimal Remote Entanglement Distribution. IEEE Journal on Selected Areas in Communications 38, 540–556 (2020).
24. Sinclair, N. et al. Spectral multiplexing for scalable quantum photonics using an atomic frequency comb quantum memory and feed-forward control. Phys. Rev. Lett. 113, 053603 (2014).
25. Guha, S. et al. Rate-loss analysis of an efficient quantum repeater architecture. Phys. Rev. A 92, 022337 (2015).
26. Liu, X., Zhou, Z.-Q., Hua, Y.-L., Li, C.-F. & Guo, G.-C. Semihierarchical quantum repeaters based on moderate lifetime quantum memories. Phys. Rev. A 95, 012319 (2017).
27. Briegel, H.-J., Dür, W., Cirac, J. I. & Zoller, P. Quantum repeaters: The role of imperfect local operations in quantum communication. Phys. Rev. Lett. 81, 5932–5935 (1998).
28. Duan, L.-M., Lukin, M. D., Cirac, J. I. & Zoller, P. Long-distance quantum communication with atomic ensembles and linear optics. Nature 414, 413–418 (2001).
29. Collins, O. A., Jenkins, S. D., Kuzmich, A. & Kennedy, T. A. B. Multiplexed Memory-Insensitive Quantum Repeaters. Physical Review Letters 98, 060502 (2007).
Acknowledgements

SC and RK gratefully acknowledge support by the European Research Council under grant agreement no. 101001976 (project EQUIPTNT). Figures were produced using VESTA [46]. RK thanks Isaac Kim for discussions on surface codes.

Author contributions

SC and RK both significantly contributed to this work.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41534-024-00928-4.

Correspondence and requests for materials should be addressed to Shin Ho Choe.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2024