Distributed quantum computation based on small quantum registers

Share Your Story

Distributed Quantum Computation Based-on Small Quantum Registers

Liang Jiang[1], Jacob M. Taylor[1,2], Anders S. Sørensen[3], Mikhail D. Lukin[1]

[1] *Department of Physics, Harvard University, Cambridge, Massachusetts 02138*
[2] *Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 and*
[3] *Quantop and The Niels Bohr Institute, University of Copenhagen, DK-2100 Copenhagen Ø, Denmark*

We describe and analyze an efficient register-based hybrid quantum computation scheme. Our scheme is based on probabilistic, heralded optical connection among local five-qubit quantum registers. We assume high fidelity local unitary operations within each register, but the error probability for initialization, measurement, and entanglement generation can be very high ($\sim 5\%$). We demonstrate that with a reasonable time overhead our scheme can achieve deterministic non-local coupling gates between arbitrary two registers with very high fidelity, limited only by the imperfections from the local unitary operation. We estimate the clock cycle and the effective error probability for implementation of quantum registers with ion-traps or nitrogen-vacancy (NV) centers. Our new scheme capitalizes on a new efficient two-level pumping scheme that in principle can create Bell pairs with arbitrarily high fidelity. We introduce a Markov chain model to study the stochastic process of entanglement pumping and map it to a deterministic process. Finally we discuss requirements for achieving fault-tolerant operation with our register-based hybrid scheme, and also present an alternative approach to fault-tolerant preparation of GHZ states.

I. INTRODUCTION

The key challenge in experimental quantum information science is to identify isolated quantum mechanical systems with good coherence properties that can be manipulated and coupled together in a scalable fashion. Recently, considerable advances have been made towards interfacing of individual qubits in the optical and microwave regimes. These include advances in cavity QED [1, 2] as well as in probabilistic techniques for entangling remote qubits [3, 4, 5]. At the same time, substantial progress has been made towards the physical implementation of few-qubit quantum registers using systems of coupled trapped ions [6, 7, 8], neutral atoms [9], or solid-state qubits based on either electronic and nuclear spins in semiconductors [10, 11, 12] or superconducting islands [13, 14].

While the precise manipulation of large, multi-qubit systems still remains an outstanding challenge, various approaches for connecting such few qubit sub-systems into large scale circuits have been investigated [5, 15, 16, 17, 18]. These studies suggest that hybrid schemes, which benefits from short range interactions for local coupling and (optical) long range interactions for non-local coupling, might be an effective way toward large scale quantum computation: small local few-qubit quantum systems may be controlled with very high precession using optimal control techniques [19, 20], and in practice it may be more feasible to operate several such small subsystems compared to the daunting task of high-precession control of a single large quantum system with thousands of qubits. Optical techniques for quantum communication can then be used to connect any two sub-systems. For example, we may directly transfer a quantum state from one sub-system to another via an optical channel [21] (see Fig. 1a), which immediately provides an efficient way to scale-up the total number of the physical qubits we can manipulate coherently. In particular the use of optical means for connecting different subsystems has the advantage that it allows for fast non-local operations over large distances. This is advantageous for quantum error correction since the existence of such non-local coupling operations alleviates the threshold requirement for fault-tolerant quantum computation [22].

In practice, however, it is very difficult to have a perfect optical connection. In particular, there is excitation loss associated with the optical channel, due to scattering or absorption. For a lossy channel, it is therefore more desirable to use it to generate entanglement between different sub-systems (see Fig. 1b), rather than for direct state transfer. The entanglement generation is then heralded by the click patterns from the photon detectors. Such detection-based scheme is intrinsically robust against excitation loss in the channel, since it only reduces the success probability but does not affect the entanglement fidelity. This entanglement can then be used as a resource to teleport quantum state from one sub-system to another [23]. More generally, entanglement provides a physical resource to implement non-local unitary coupling gates (such as CNOT gate) [24, 25, 26, 27]. If there is only one physical qubit for each sub-system, so-called cluster states [28] can be created based on probabilistic, heralded entanglement generation (see Ref. [5] and references therein). Such cluster states can be used for universal quantum computation [29]. If there are two physical qubits available for each sub-system, cluster states can be created deterministically [30]; meanwhile one can also use these two-qubit sub-systems to implement any quantum circuit directly [4].

Furthermore, realistic optical channel connecting subsystems has other imperfections beside excitation loss, such as the distortion of the polarization or shape of the wave-packet. These imperfections will reduce the fidelity of the heralded entanglement generated between the sub-
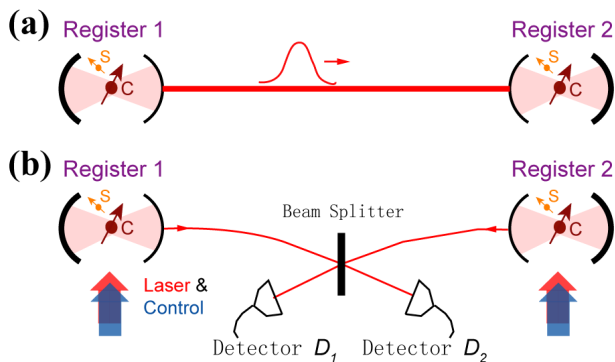
FIG. 1: (Color online) Two schemes to couple different registers. (a) Deterministic state transfer from one register to the other [21]. (b) Probabilistic entanglement generation, heralded by distinct detector click patterns [3, 31, 32]. In the text, we argue that probabilistic, heralded entanglement generation is sufficient for deterministic distributed quantum computation.

systems. To overcome these imperfections in the channel, entanglement purification schemes have been proposed, which may create some high-fidelity Bell pairs from many low-fidelity ones [33, 34]. In particular, the entanglement pumping scheme originally presented in the context of quantum communication over long distances [31, 35, 36] provides a very efficient purification scheme in terms of local physical resources, and Dür and Briegel [15] first proposed to use such entanglement pumping scheme for quantum computation. In principle, the infidelity of the purified Bell pair shared by the sub-systems can be very low, and is only limited by the error probability from local operations. In Ref. [15] it was found that *three* auxiliary qubits (requiring a total of five qubits including the storage and communication qubits) for each sub-system provide enough physical resources to obtain high fidelity entangled pairs via entanglement pumping.

In order to implement the idea of distributed quantum computation using realistic optical channel and imperfect operations, it is necessary to consider the following questions: What are the minimal local physical resources needed for robust entanglement generation? What is the time overhead associated with entanglement generation? Can we extend the robustness to other imperfections, such as errors from initialization and measurement?

Motivated by these considerations, we study the practical implementation of robust quantum registers for scalable applications. In Ref. [37] we have proposed an entanglement purification scheme that only requires *two* auxiliary qubits for robust entanglement generation. We have found that the time overhead associated with entanglement generation ranges from a factor of 10 to a few 100, depending on the initial and targeting infidelities. We have also suggested to use *one more* auxiliary qubit to suppress errors from initialization and measurement. Thus, our hybrid scheme also requires only 5 (or fewer)

qubits with local deterministic coupling, while providing additional improvements over the protocol of Ref. [15]: reduced measurement errors, higher fidelity, and more efficient entanglement purification. In this paper, we will provide a detailed discussion on the register-based, hybrid quantum computation scheme presented in Ref. [37].

The paper is organized as follows. In Sec. II, we will introduce the concept of quantum register and discuss two experimental implementations. In Sec. III, we will review the idea of universal quantum computation based on two-qubit quantum registers. In Sec. IV, we will specify the error models for imperfections and provide the basic ideas underlying our robust operations. In Sec. V, we will describe the robust measurement/initialization scheme. In Sec. VI, we will present our bit-phase two-level entanglement pumping scheme. In Sec. VII, we will introduce the Markov chain model to quantitatively analyze the time overhead and residual infidelity associated with the stochastic process of entanglement pumping, and discuss further improvement upon our two-level entanglement pumping scheme. In Sec. VIII, we will map our stochastic, hybrid, and distributed quantum computation scheme to a deterministic computation model that is characterized by two quantities (the clock cycle and effective error probability), and estimate the practical values for these quantities. We will also consider the constraint set by the finite memory lifetime and determine the achievable performance of hybrid distributed quantum computation. Finally, in Sec. IX, we will discuss using our hybrid scheme for fault-tolerant quantum computation with quantum error correcting codes, and provide a resource-efficient approach for fault-tolerant preparation of the GHZ states.

II. QUANTUM REGISTER AND EXPERIMENTAL IMPLEMENTATIONS

We define a *quantum register* as a few-qubit device (see Fig. 2a) that contains one *communication* qubit (c), with a photonic interface; one *storage* qubit (s), with very good coherence times; and several *auxiliary* qubits ($a_1, a_2, ...$), used for purification and error correction. A critical requirement for a quantum register is high-fidelity unitary operations between the qubits within a register.

The quantum registers considered here can be implemented in several physical systems, but in this paper we shall focus on two specific systems where these considerations are particularly relevant. First, ion traps have been used to demonstrate all essential elements of quantum registers. (1) The ion qubits may play the role of communication qubits: they can be initialized and measured efficiently using optical pumping and cycling transitions respectively, and they can also be prepared in highly entangled states with the polarization of single photons [39, 40]. Very recently, entanglement generation between ion qubits from two remote traps has been demonstrated
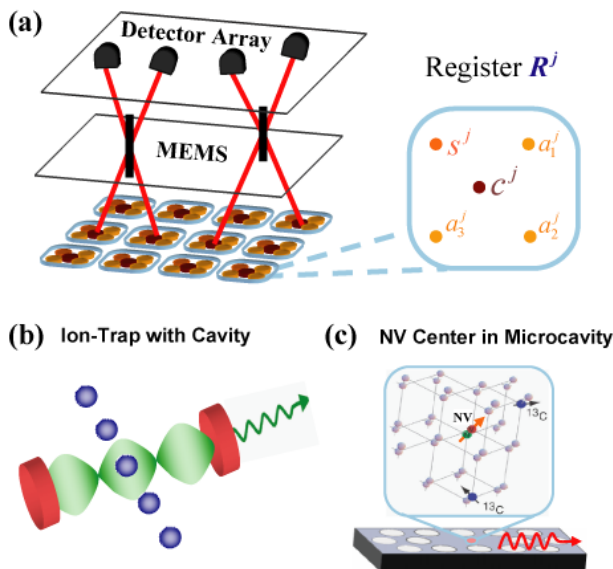
FIG. 2: (Color online) Distributed quantum computer. (a) Illustration of distributed quantum computer based on many quantum registers. Each register has five physical qubits, including one communication qubit (c), one storage qubit (s), and three auxiliary qubits ($a_{1,2,3}$). Local operations for qubits within the same register have high fidelity. Entanglement between non-local registers can be generated probabilistically [3, 31, 32]. Devices of optical micro-eletro-mechanical systems (MEMS) [38] can efficiently route photons and couple arbitrary pair of registers. Detector array can simultaneously generate entanglement for many pairs of registers. (b) An ion trap coupled to a cavity also provides a promising candidate for distributed quantum computation. A single ion is resonantly coupled to the cavity and serves as the communication qubit; while the others can be storage or auxiliary qubits. (c) Nitrogen-vacancy (NV) defect center in photonic crystal micro-cavity. The inset shows the atomic structure of the NV center [10], which forms a quantum register. The electronic spin localized at the vacancy is optically accessible (measurement/initialization) and can play the role of the communication qubit. The nuclear spins from proximal ^{13}C atoms constitute the storage and auxiliary qubits, which are coherently controlled via hyperfine interaction and rf pulses.

[41, 42]. This experiment directly demonstrates the non-local connection required for our hybrid approach. Since the photon collection and detection efficiency is not perfect, the entanglement generation is a probabilistic process. However, the entanglement generation is also a heralded process, because different click patterns from the detectors can be used to identify each successful entanglement generation. As we will discuss Sec. III, such probabilistic, heralded entanglement generation process is already sufficient to implement deterministic non-local coupling gates. (2) The ion qubits can be good storage qubits as well. Coherence time of approximately 10 seconds has been demonstrated in ion traps [43, 44], which is $10^{6\sim7}$ times longer than the typical gate operation time which is at the order of μs. Since fault-tolerant quantum computation only requires the coherence time to be approximately 10^4 times longer than the gate time, the very long coherence of the ions provides new opportunities, such as performing non-local coupling gates with some extra time overhead. (3) Coherent manipulation of few ions in the same trap has been demonstrated [45, 46], allowing gates to be implemented among the qubits in the register. (4) High fidelity operations between the ion qubits within the ion trap has also been demonstrated [19].

A second promising candidate for implementing quantum registers is the nitrogen vacancy (NV) centers in diamond. Each NV center can be regarded as an ion trap confined by the diamond crystal, which can be treated as a single register. The qubits for each NV register consists of one electronic spin associated with the defect and several nuclear spins associated with the proximal C-13 nuclei. The electronic spin is optically active, so that it can be measured and initialized optically. With optical cavities or diamond-based photonic crystal micro-cavities [47] one could enhance photon collection efficiency towards unity. Furthermore, the electron spin can be coherently manipulated by microwave pulses [10]. The electronic spin is thus suitable as the communication qubit. The nuclear spins are coupled to the electronic spin via hyperfine interactions. One can either use these hyperfine interaction to directly rotate the nuclear spins [48, 49], or apply radio-frequency pulses to address individual nuclear spins spectroscopically [50]. These nuclear spins have very long coherence times approaching seconds [12], and can be good storage and auxiliary qubits. Furthermore, the optical manipulation of the electronic spin can be well decoupled from the nuclear spins [51]. It can be inferred from the recent experiment [12] that the fidelity of local operations between electronic and nuclear spins is higher than 90%. While the fidelity is still low for the procedures considered here, we believe that it can be significantly improved (to higher than 0.999) by optimal control techniques [20, 50], such as composite pulses [52] and numerically optimized GRAPE pulses [53].

III. UNIVERSAL QUANTUM COMPUTATION WITH TWO-QUBIT REGISTERS: FUNDAMENTALS

We now consider universal quantum computation via the simplest possible two-qubit registers [4, 30]. Each register has one qubit for communication and the other qubit for storage. We can use probabilistic approaches from quantum communication ([31] and references therein) to generate entanglement between communication qubits from two *arbitrary* non-local registers. The probabilistic entanglement generation creates a Bell pair conditioned on certain measurement outcomes, which are distinct from outcomes of unsuccessful entanglement generation. If the entanglement generation fails, it can be re-attempted until success, with an exponen-
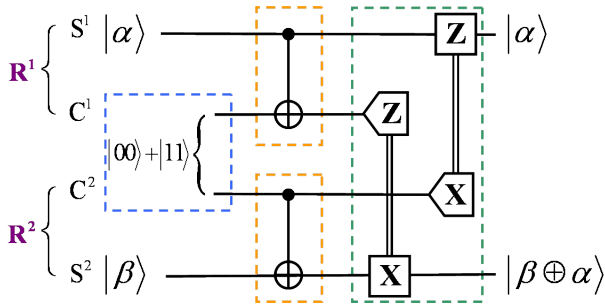
FIG. 3: (Color online) Quantum circuit for non-local CNOT gate between two registers R^1 and R^2. The circuit starts with the Bell state $|\Phi^+\rangle_{c^1,c^2} = (|00\rangle + |11\rangle)/\sqrt{2}$ for the communication qubits c^1 and c^2 (the left blue box). After local unitary operations within each register (the middle orange boxes), qubits c^1 and c^2 are projectively measured in Z and X basis, respectively (the right green box). According to Eq. (1), up to some local unitary gates (the right green box), this circuit implements the non-local CNOT gate to qubits s^1 and s^2 from two quantum registers.

FIG. 4: (Color online) General circuit for non-local Controlled-U gate [27], with the storage qubit S^1 as control and the storage qubit S^2 as target. One Bell pair is consumed for this operation.

IV. ERRORS AND IMPERFECTIONS

In practice, the qubit measurement, initialization, and entanglement generation can be noisy with error probabilities as high as $\sim 5\%$, due to practical limitations such as imperfect cycling transitions, finite collection efficiency, and poor interferometric stability. As a result, there will be a large error probability in non-local gate circuits. In contrast, local unitary operations may fail infrequently ($p_L \lesssim 10^{-4}$) when quantum control techniques for small quantum system are utilized [19, 20]. We now show that the most important sources of imperfections, such as imperfect initialization and measurement errors for individual qubits in each quantum register, and entanglement generation errors between registers, can be corrected with a modest increase in register size. We determine that with just *three* additional auxiliary qubits and high-fidelity local unitary operations, all these errors can be efficiently suppressed by repeated quantum non-demolition (QND) measurement [54] and entanglement purification [35, 36]. This provides an extension of Ref. [15] that mostly focused on suppressing errors from entanglement generation.

We will use the following error model for the entire paper: (1) The imperfect local two-qubit operation U_{ij} is

$$U_{ij}\rho U_{ij}^\dagger \rightarrow (1 - p_L)\, U_{ij}\rho U_{ij}^\dagger + \frac{p_L}{4}\, \mathrm{Tr}_{ij}\left[\rho\right] \otimes \mathbf{I}_{ij} \qquad (2)$$

where $\mathrm{Tr}_{ij}\left[\rho\right]$ is the partial trace over the qubits i and j, and $\mathbf{I}_{ij}$ is the identity operator for qubits i and j. This error model describes that with a probability $1 - p_L$ the gates perform the correct operation and with a probability p_L the gates produce a complete random output for the two involved qubits [71]. (2) The imperfect initialization of state $|0\rangle$ will prepare a mixed state

$$\rho_0 = (1 - p_I)\, |0\rangle\langle 0| + p_I\, |1\rangle\langle 1|, \qquad (3)$$

which has error probability p_I, i.e., it prepares the wrong state with a probability p_I. (3) The imperfect measurement of state $|0\rangle$ will correspond to the projection oper-

tially decreasing probability of continued failure.

When the communication qubits (c^1 and c^2) are prepared in the Bell state, we can immediately perform the non-local CNOT gate on the storage qubits (s^1 and s^2) using gate-teleportation between registers R^1 and R^2. The gate-teleportation circuit in Fig. 3 implements (before the conditional Pauli operations) the following map

$$|\phi\rangle_{s^1,s^2} |\Phi^+\rangle_{c^1,c^2} \rightarrow (\sigma_{s^1}^z)^{m_{c^2}} (\sigma_{s^2}^x)^{m_{c^1}} \mathrm{CNOT}_{s^1,s^2} |\phi\rangle_{s^1,s^2}, \qquad (1)$$

where $|\Phi^+\rangle_{c^1,c^2} = (|00\rangle + |11\rangle)/\sqrt{2}$, $\mathrm{CNOT}_{i,j}$ is the controlled-NOT (CNOT) gate with the ith qubit as control and the jth qubit as target, and $m_i = 0, 1$ is the measurement result for qubit i from the circuit in Fig. 3.

By consuming one Bell pair, one can implement any non-local controlled-U gate between two storage qubits [27], as shown in Fig. 4. Since operations on a single qubit can be performed within a register, the CNOT operation between different quantum registers is in principle sufficient for universal quantum computation [4]. Similar approaches are also known for deterministic generation of graph states [30] —an essential resource for one-way quantum computation [29].

We emphasize that deterministic entanglement generation is *not* required, which opens up a wide range of possibilities of entanglement generation. It is experimentally challenging to achieve deterministic quantum state transfer directly [21], but we are able to achieve the same task by probabilistic entanglement scheme and two-qubit quantum registers [4].
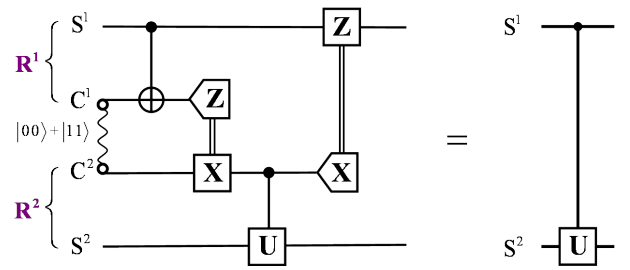
ator

$$P_0 = (1 - p_M) |0\rangle \langle 0| + p_M |1\rangle \langle 1|, \qquad (4)$$

This operator describes that a qubit prepared in state $|0\rangle$ or $|1\rangle$ will give rise to the opposite measurement output with the measurement error probability p_M. (4) Finally, the entanglement fidelity for a non-ideal preparation of state $|\Phi^+\rangle$ is defined as

$$F = \langle \Phi^+| \rho |\Phi^+\rangle, \qquad (5)$$

and the infidelity is just $1 - F$. The fidelity does, however, not completely characterize the produced entangled state. Depending on the exact method used to generate the entangled state, one can in some situations argue that the error will predominantly be, e.g., only a phase error [3, 32, 55], whereas in other situations it will be a combination of phase and bit flip errors (see [31] and references therein). Below we shall therefore both consider the situation where we only have a dephasing error as well the situation, where we have a more complicated depolarizing error (exact definition given later). As we shall see, the knowledge that the error is of a particular type (e.g. only dephasing error) provides a significant advantage for purification.

We will also assume a separation of error probabilities: any internal, unitary operation within the register fails with extremely low probability, p_L, while all operations connecting the communication qubit to the outside world (initialization, measurement, and entanglement generation) fail with error probabilities that can be several orders of magnitude higher.

$$p_L \ll p_I, p_M, 1 - F. \qquad (6)$$

In terms of these quantities the error probability in the non-local CNOT gate in Fig. 3 is

$$p_{CNOT} \sim (1 - F) + 2p_L + 2p_M., \qquad (7)$$

because we use one entangled state, two local operations, and two measurements. In the next two sections, we will show how to use robust operations to dramatically improve the fidelity for these non-local coupling gates.

Robust measurement can be implemented by repeated QND measurement, i.e., a majority vote among the measurement outcomes (Fig. 5), following a sequence of CNOT operations between the auxiliary/storage qubit and the communication qubit. This also allows *robust initialization* by measurement. High-fidelity, *robust entanglement generation* is achieved via entanglement pumping [15, 35, 36] (Figs. 6 and 7), in which lower fidelity entanglement between the communication qubits is used to purify entanglement between the auxiliary qubits, which can then be used for non-local CNOT operations. To make the most efficient use of physical qubits, we introduce a new entanglement pumping scheme. In our *bit-phase two-level entanglement pumping* scheme, we first use unpurified Bell pairs to repeatedly pump (purify) against bit-errors (Fig. 7a), and then
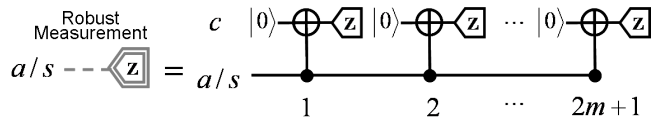

FIG. 5: The robust measurement scheme based on repeated quantum non-demolition (QND) measurements and majority vote. Each QND measurement consists of initializing, coupling, and measuring the communication qubit. The QND measurements are repeated $2m+1$ times using the same communication qubit.

use the bit-purified Bell pairs to repeatedly pump against phase-errors (Fig. 7b).

Entanglement pumping, like entanglement generation, is probabilistic; however, failures are detected. In computation, where each logical gate should be completed within the allocated time (clock cycle), failed entanglement pumping can lead to gate failure. To demonstrate the feasibility of our approach for quantum computation, we will analyze the time required for robust initialization, measurement and entanglement generation, and show that the failure probability for these procedures can be made sufficiently small with reasonable time overhead.

V. ROBUST MEASUREMENT & INITIALIZATION

In this section, we will analyze the robust measurement scheme based on repeated QND measurement, discuss the recent experimental demonstration of robust measurement in the ion-trap system, and present two approaches to robust initialization.

The measurement circuit shown in Fig. 5 yields the correct result based on a majority vote from $2m+1$ consecutive readouts (bit-verification). Since the evolution of the system (CNOT gate) commutes with the measured observable (Z operator) of the auxiliary/storage qubit, it is a quantum non-demolition (QND) measurement [54], which can be repeated many times. The error probability for such a majority vote measurement scheme is:

$$\varepsilon_M \approx \sum_{j=m+1}^{2m+1} \binom{2m+1}{j} (p_I + p_M)^j + \frac{2m+1}{2} p_L \qquad (8)$$

where the last term conservatively estimates the probability for bit-flip error of the auxiliary/storage qubit during the repeated QND measurement. For simplicity, we will use Eq. (8) for our conservative estimate of error probability for repeated QND measurement. Suppose $p_I = p_M = 5\%$, we can achieve $\varepsilon_M \approx 8 \times 10^{-4}$ by choosing $m^* = 6$ for $p_L = 10^{-4}$, or even $\varepsilon_M \approx 12 \times 10^{-6}$ by choosing $m^* = 10$ for $p_L = 10^{-6}$. For convenience of discussion, we shall add ε_M to the set of imperfection parameters: $(1 - F, p_I, p_M, p_L, \varepsilon_M)$. The time for robust

measurement is

$$\tilde{t}_M = (2m+1)(t_I + t_L + t_M). \quad (9)$$

where t_I, t_L, and t_M are times for initialization, local unitary gate, and measurement, respectively.

Measurements with very high fidelity (ε_M as low as 6×10^{-4}) have recently been demonstrated in the ion-trap system [56], using similar ideas as above. There are several possibilities to further improve the performance of the repeated QND measurement. (1) We may use maximum likelihood estimate (MLE) to replace the majority vote for repeated measurements with multi-value outcomes (e.g., fluorescent intensity) [56]. (2) We may keep updating the error probability using MLE after each measurement. Once the estimated error probability is below some fixed error rate, we stop the repetition of the QND measurement to avoid errors from redundant operations [56]. (3) We may use the implementation of CNOT gate that has small/vanishing bit-flip errors to the control qubit, which will reduce/eliminate the last term in Eq. (8).

The robust measurement scheme also allows to achieve robust initialization by measurement, i.e., by measuring the state of a qubit with the robust measurement scheme, we initialize the qubit into the result of the measurement outcome with an effective initialization error

$$\varepsilon_I \approx \varepsilon_M. \quad (10)$$

Besides the above measurement-based scheme, we may achieve robust initialization using verification-based scheme [72]. For clarity, we shall assume the measurement-based initialization [Eq. (10)] for the rest of the paper.

VI. ROBUST NON-LOCAL TWO-QUBIT GATE

With high-fidelity local unitary gate and repeated QND measurement, the error probability for non-local coupling gates (e.g., Fig. 3 and 4) is

$$p_{CNOT} \sim (1-F) + 2p_L + 2\varepsilon_M, \quad (11)$$

which is dominated by the infidelity of the Bell pair $1-F$, since we assume $p_L \sim \varepsilon_M \ll 1-F$. In this section, we will show how to create high-fidelity Bell pairs between two registers with a reasonable time overhead. We will first briefly review entanglement pumping [31, 35, 36]. Then we will quantitatively analyze the fidelity of the purified Bell pairs for our efficient two-level pumping scheme, and introduce the Markov chain model to calculate the failure probability for entanglement pumping within a given number of attempts. Next we will quantify the performance of the high-fidelity Bell pair generation in terms of the total error probability (or average infidelity) and the time overhead, and discuss the trade-off between these two criteria. Finally, we will mention a non-post-selective pumping scheme which may further reduce the time overhead.
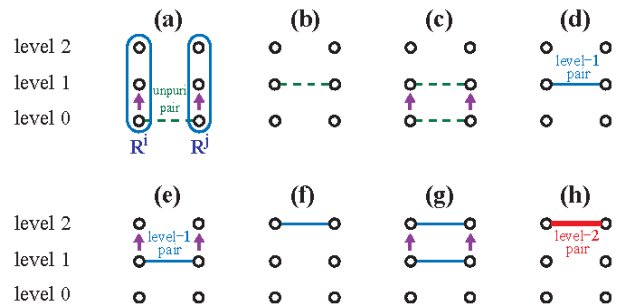

FIG. 6: (Color online) Two-level entanglement pumping between registers R^i and R^j [circled by rounded rectangles in (a)]. (a,b) Generate and store one unpurified Bell pair. (c) Generate another unpurified Bell pair to pump (purify) the previously stored pair. (d) If the purification is successful, we obtain a purified Bell pair (level-1 pair) with higher fidelity; otherwise we discard the stored Bell pair and start the entire pumping process from the beginning. (e-f) The second level of entanglement pumping uses previously purified pairs to purify a stored Bell pair, to obtain a Bell pair with higher fidelity (level-2 pair).

A. Entanglement pumping

We now consider entanglement pumping [31, 35, 36] with high-fidelity local unitary gate and robust measurement. During the entanglement pumping process, we first store one unpurified Bell pair [Fig. 6 (a,b)], and then generate another unpurified Bell pair to purify the previously stored pair [Fig. 6 (c,d)]. If the purification is successful, we will obtain a purified Bell pair with higher fidelity, which can be further purified by repeating the process in Fig. 6 (c,d) with new unpurified Bell pairs; otherwise we discard the stored Bell pair, and start the entire pumping process from the beginning. Sometimes, we may want to introduce a second level of entanglement pumping; that is to use previously purified pairs to purify a stored Bell pair [Fig. 6 (e-h)].

B. Fidelity of entanglement pumping

We now analyze the performance of entanglement pumping for different errors of the unpurified Bell pairs. If the unpurified Bell pair is dominated by one type of error (e.g., dephasing error with density matrix $\rho_{\text{dephasing}} = diag[F, 1-F, 0, 0]$ in the Bell basis $\{|\Phi^+\rangle, |\Phi^-\rangle, |\Psi^+\rangle, |\Psi^-\rangle\}$, defined as $|\Phi^\pm\rangle = (|00\rangle \pm |11\rangle)/\sqrt{2}$, $|\Psi^\pm\rangle = (|01\rangle \pm |10\rangle)/\sqrt{2}$), we can skip the first level pumping. The unpurified pair then immediately becomes a level-1 pair and is purified with the circuit in Fig. 7b. In Fig. 8a we plot the fidelity curve (purified fidelity v.s. number of successful pumping steps) for the one-level pumping process (i.e. $n_b = 0$), where a very high fidelity pair can be created after $n_p = 3$ suc-
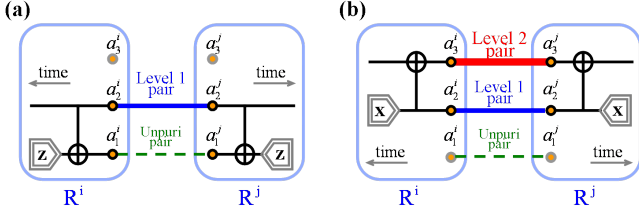
FIG. 7: (Color online) Bit-phase two-level entanglement pumping scheme to create high fidelity entangled pairs between two registers R^i and R^j. (a) Circuit for the first level pumping to purify bit-errors, corresponding to Fig. 6c. (b) Circuit for the second level pumping to purify phase-errors, corresponding to Fig. 6g. The arrows indicate the time direction for each register. Robust measurements are used here. If the two outcomes are the same, it is a successful attempt of pumping; otherwise we generate new pairs and restart the pumping operation from the beginning.
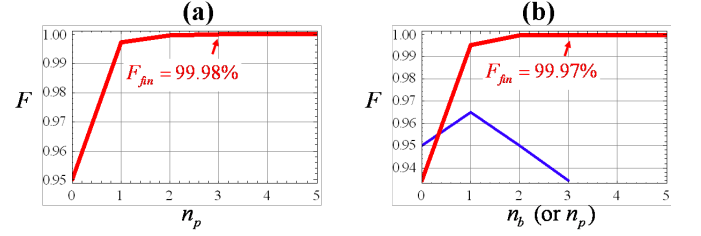

FIG. 8: (Color online) Entanglement fidelity F as a function of the number of successful pumping steps. n_b and n_p are the number of pumping steps used to purify bit-errors and phase-errors, respectively. We assume fixed measurement and local two-qubit gate error rates $\varepsilon_M = p_L = 10^{-4}$. (a) For bit-flip error $\rho_{\text{dephasing}} = diag\,[F, 1-F, 0, 0]$ in the Bell basis $\{|\Phi^+\rangle, |\Phi^-\rangle, |\Psi^+\rangle, |\Psi^-\rangle\}$, with $F = 0.95$, one-level entanglement pumping is sufficient. High fidelity of $F_{fin} = 99.98\%$ can be achieved by $n_p = 3$. (b) For depolarizing error $\rho_{\text{depolarizing}} = diag\,[F, \frac{1-F}{3}, \frac{1-F}{3}, \frac{1-F}{3}]$, with $F = 0.95$, two-level entanglement pumping is needed (see text for more details). The first level pumping only purifies the bit-error, but accumulates the phase-error at the same time, and therefore the (thin blue) fidelity curve for the first level pumping drops for $n_b > 1$. The second level (thick red curve) uses the purified level-1 pair ($n_b = 1$) to pump another stored pair. High fidelity of $F_{fin} = 99.97\%$ can be achieved by $n_p = 3$. Note that we consider the full density matrix for all numerical calculation of entanglement fidelities [57], with the error models given in Eqs. (2,3,4).

cessful pumping steps. Note that we consider the full density matrix for all numerical calculation of entanglement fidelities [57], with the error models given in Eqs. (2,3,4).

If the unpurified Bell pair contains errors from both bit-flip and dephasing processes (e.g., depolarizing error with density matrix $\rho_{\text{depolarizing}} = diag\,[F, \frac{1-F}{3}, \frac{1-F}{3}, \frac{1-F}{3}]$ in the Bell basis), two-level entanglement pumping is needed. We introduce the following *bit-phase two-level pumping scheme* —the first level has n_b steps of bit-error pumping using raw Bell pairs (Fig. 7a) to produce a bit-error-purified entangled pair, and the second level uses these bit-error-purified pairs for n_p steps of phase-error pumping (Fig. 7b). In Fig. 8b we plot the fidelity curves for the first level (thin blue curve) and the second level (thick red curve) of entanglement pumping. One-level pumping is insufficient to achieve high fidelity, but two-level pumping can achieve very high fidelity. With the parameters specified for Fig. 8b, the maximum fidelity is achieved via the optimal choice of control parameters $(n_b^*, n_p^*) = (1, 3)$ for successful pumping steps of the first and second levels, respectively.

For successful purification, the infidelity of the purified pair, $\varepsilon_{E,\text{infid}}^{(n_b, n_p)}$, depends on both the control parameters (n_b, n_p) and the imperfection parameters (F, p_L, ε_M). For depolarizing error, we find

$$\varepsilon_{E,\text{infid}}^{(n_b \geq 1, n_p \geq 1)} \approx \frac{3 + 2n_p}{4} p_L + \frac{4 + 2(n_b + n_p)}{3}(1-F)\varepsilon_M$$
$$+ (n_p + 1)\left(\frac{2(1-F)}{3}\right)^{n_b+1} + \left(\frac{(n_b+1)(1-F)}{3}\right)^{n_p+1} \quad (12)$$

to leading order in p_L and ε_M, for $n_b, n_p \geq 1$. The dependence on the initial infidelity $1 - F$ is exponentially suppressed at a cost of a linear increase of error from local operations p_L and robust measurement ε_M. Measurement-related errors are suppressed by the pref-

actor $1 - F$, since measurement error does not cause infidelity unless combined with other errors. In the limit of ideal operations $(p_L, \varepsilon_M \to 0)$, the infidelity $\varepsilon_{E,\text{infid}}^{(n_b, n_p)}$ can be arbitrarily close to zero, which is rigorously proved in Appendix A. On the other hand, if we use the standard entanglement pumping scheme [35, 36] (that alternates purification of bit and phase errors within each pumping level), the reduced infidelity from two-level pumping is always larger than $(1 - F)^2/9$. Therefore, for very small p_L and ε_M, the new pumping scheme is crucial to minimize the number of qubits per register.

In Fig. 9, we show the contours of the infidelity $\varepsilon_{E,\text{infid}}^{(n_b, n_p)}$ as a function of n_b and n_p, where the contours are labeled by values of $\log_{10} \varepsilon_{E,\text{infid}}^{(n_b, n_p)}$. The parameters for the contour plots are $F = 0.90$ (left) and $F = 0.95$ (right); $\varepsilon_M = p_L = 10^{-4}$ (up) and $\varepsilon_M = p_L = 10^{-6}$ (down). For optimal choice of (n_b^*, n_p^*), the minimal infidelity is limited by ε_M and p_L.

For dephasing error, one level pumping is sufficient (i.e. no bit-error purification, $n_b = 0$). The infidelity is approximately

$$\varepsilon_E^{(0, n_p \geq 1)} \approx (1-F)^{n_p+1} + \frac{2 + n_p}{4} p_L + 2(1-F)\varepsilon_M \quad (13)$$

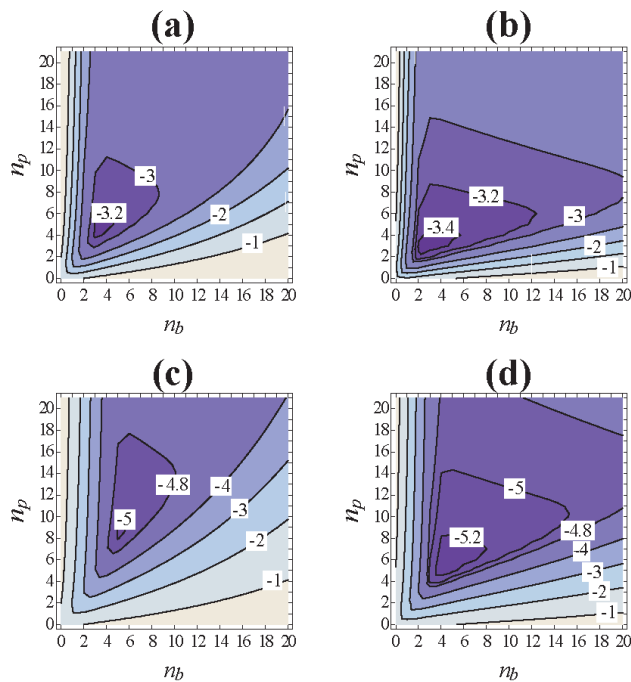by expanding to the leading order in p_L and ε_M.

FIG. 9: (Color online) The contours of the infidelity $\varepsilon_{E,\text{infid}}^{(n_b,n_p)}$ as a function of n_b and n_p for depolarizing error. We use $\log_{10}\varepsilon_{E,\text{infid}}^{(n_b,n_p)}$ to label the contours. The other parameters are $F = 0.90$ (left) and $F = 0.95$ (right); $\tilde{\varepsilon}_M = p_L = 10^{-4}$ (up) and $\tilde{\varepsilon}_M = p_L = 10^{-6}$ (down). With optimal choice of (n_b^*, n_p^*), the minimal infidelity is comparable to the corresponding value of p_L.

VII. MARKOV CHAIN MODEL

The overall success probability can be defined as the joint probability that all successive steps succeed. We use the model of finite-state Markov chain [58] to directly calculate the *failure probability* of (n_b, n_p)-two-level entanglement pumping using N_{tot} raw Bell pairs, denoted as $\varepsilon_{E,fail}^{(n_b,n_p)}(N_{\text{tot}})$.

A. Markov chain model for entanglement pumping

We first use the Markov chain model to study the n-step one-level entanglement pumping. As shown in Fig. 10, we use "0" to denote the *initial state* with no Bell pairs, "1" for the state with one stored unpurified pair, $(j+1)$ for the state with one purified pair surviving j steps of pumping, and "$*$" for the *final state* with the purified pair surviving n steps of pumping. Altogether there are $n+2$ states. The (success) transition probability from state j to state $j+1$ is q_j, while the (failure) transition probability from state j to state 0 is $1-q_j$, for $j = 0, 1, \cdots, n$. Here $q_0 \equiv 1$ [corresponding to deterministic state transfer as shown in Fig. 6 (a,b)] and $q_{j\geq1}$
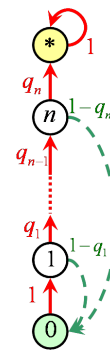

FIG. 10: (Color online) Markov chain model for one-level entanglement pumping. We use "0" to denote the *initial state* with no Bell pairs, "1" for the state with one stored unpurified pair, $(j+1)$ for the state with one purified pair surviving j steps of pumping, and "$*$" for the *final state* with the purified pair surviving n steps of pumping. The (success) transition probability from state j to state $j+1$ is q_j, while the (failure) transition probability from state j to state 0 is $1-q_j$, for $j = 0, 1, \cdots, n$. Here $q_0 \equiv 1$ and $q_{j\geq1}$ can be calculated according to the density matrix of the purified Bell pair surviving j steps of pumping [36]. The final state is self-trapped, and goes back to itself with unit probability. Each transition attempt consumes one unpurified Bell pair.

can be calculated with the density matrix of the purified Bell pair surviving $j-1$ steps of pumping [36]. The final state is self-trapped, and goes back to itself with unit probability, representing that once we have reached the desired final fidelity we no longer make any purification attempts and the system remains in this state with unit probability. Each transition attempt consumes one unpurified Bell pair. We would like to know the probability of reaching the final state "$*$" after N_{tot} attempts. More generally, we might also want to know the probability distribution over all $n+2$ states.

We use a (column) vector $\vec{P}$ with $n+2$ elements to characterize the probability distribution among all $n+2$ states. From the t-th attempt to the $(t+1)$-th attempt, the probability vector evolves from $\vec{P}_t$ to $\vec{P}_{t+1}$ according to the rule

$$\vec{P}_{t+1} = \mathbf{M}\,\vec{P}_t, \tag{14}$$

with the transition matrix

$$\mathbf{M} = \begin{pmatrix} 0 & 1-q_1 & 1-q_2 & \cdots & 1-q_n & 0 \\ 1 & 0 & & & & \\ & q_1 & 0 & & & \\ & & q_2 & 0 & & \\ & & & \cdots & 0 & \\ & & & & q_n & 1 \end{pmatrix}. \tag{15}$$

Since the initial probability vector is $\vec{P}_0 = (1, 0, \cdots, 0)^T$, we can calculate the probability vector after N_{tot} attempts

$$\vec{P}_{N_{\text{tot}}} = \mathbf{M}^{N_{\text{tot}}}\vec{P}_0. \tag{16}$$
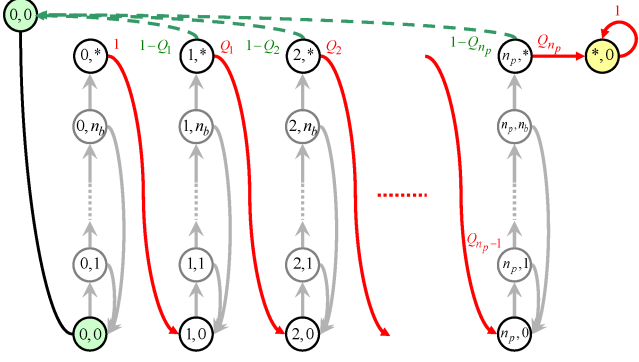
FIG. 11: (Color online) Markov chain model for two-level entanglement pumping. The required pumping steps are n_b and n_p for the two levels, respectively. We use "0,0" to denote the initial state with no Bell pairs, "0, $j+1$" for the state with one purified pair surviving j steps of pumping at the first level, "$k+1, j+1$" for the state with one purified pair surviving k steps of pumping at the second level and one purified pair surviving j steps of pumping at the first level, "$k+1,*$" for the state with one purified pair surviving k steps of pumping at the second level and one purified pair surviving n_b steps of pumping at the first level, and "$*,0$" for the final state with one purified pair surviving n_p steps of pumping at the second level. For the first level pumping, the (success) transition probability from state "k,j" to state "$k,j+1$" is q_j, while the (failure) transition probability from state "k,j" to state "$k,0$" is $1-q_j$, for $j=0,1,\cdots,n_b$. For the second level pumping, the (success) transition probability from state "$k,*$" to state "$k+1,0$" is Q_k, while the (failure) transition probability from state "$k,*$" to state "$0,0$" is $1-Q_k$, for $k=0,1,\cdots,n_p$. The final state is self-trapped, and goes back to itself with unit probability.

The probability vector $\vec{P}_{N_{\text{tot}}}$ describes the entire probability distribution over all states of the Markov chain. The last element of $\vec{P}_{N_{\text{tot}}}$ is the success probability of reaching the final state "$*$" after N_{tot} attempts; the failure probability after N_{tot} attempts is thus

$$\varepsilon_{E,fail}^{(n_b,n_p)}(N_{\text{tot}}) = 1 - (P_{N_{\text{tot}}})_{n+2}. \quad (17)$$

For two-level entanglement pumping, the state transition diagram is shown in Fig. 11. n_b and n_p are the number of pumping steps used to purify bit-errors and phase-errors, respectively. As detailed in Appendix B, we may use a (column) vector $\vec{P}$ with $(n_b+1)(n_p+1)+1$ elements to characterize the probability distribution among all $(n_b+1)(n_p+1)+1$ states. From the t-th attempt to the $(t+1)$-th attempt, the probability vector evolves from $\vec{P}(t)$ to $\vec{P}(t+1)$ according to the same rule as above [Eq. (14)], but with the transition matrix $\mathbf{M}$ given in Eq. (B2).

Similar to one-level pumping, we can calculate the probability vector after N_{tot} attempts using Eq. (16). The probability vector $\vec{P}_{N_{\text{tot}}}$ describes the entire probability distribution over all states of the Markov chain.
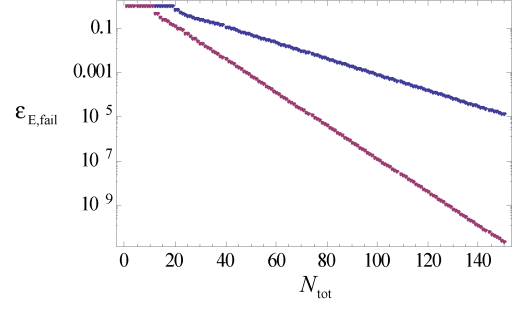

FIG. 12: (Color online) Failure probability $\varepsilon_{E,\text{fail}}$ as a function of N_{tot}. We assume a depolarizing error with $F=0.95$, $p_L = \tilde{\varepsilon}_M = 10^{-4}$. We choose $(n_b,n_p)=(2,3)$ for the lower curve, and $(3,4)$ for the upper curve. For large N_{tot}, the failure probability $\varepsilon_{E,\text{fail}}$ decreases exponentially with N_{tot}.

The last element of $\vec{P}_{N_{\text{tot}}}$ is the success probability of reaching the final state "$*,0$" after N_{tot} attempts; the failure probability after N_{tot} attempts is then

$$\varepsilon_{E,\text{fail}}^{(n_b,n_p)}(N_{\text{tot}}) = 1 - P(N_{\text{tot}})_{(n_b+1)(n_p+1)+1}. \quad (18)$$

In Fig. 12, we plot the failure probability $\varepsilon_{E,\text{fail}}^{(n_b,n_p)}(N_{\text{tot}})$ v.s. N_{tot}, for control parameters $(n_b,n_p)=(2,3)$ and $(3,4)$. For N_{tot} sufficiently large, the failure probability decreases exponentially to zero. For any given parameters, we can efficiently suppress the failure probability with some reasonably large N_{tot}.

B. Total error probability & average infidelity

We now introduce the *total error probability* (TEP) approximated by the sum of the failure probability and the infidelity of the purified Bell pair

$$\varepsilon_E^{(n_b,n_p)}(N_{\text{tot}}) \approx \varepsilon_{E,\text{fail}}^{(n_b,n_p)}(N_{\text{tot}}) + \varepsilon_{E,\text{infid}}^{(n_b,n_p)}. \quad (19)$$

This is a very conservative estimate, since sometimes we do create some partially purified Bell pair though not the targeted purified Bell pair. And here we just say that the state has fidelity zero in these cases.

To consider the possibility of using a partially purified Bell pair for output, we may introduce another useful quantity —the *average infidelity* (AIF) —for the output Bell pair from the robust entanglement generation, where we take into account these partially purified pairs. The average infidelity of the output pair is the weighted average of the infidelity of the Markov chain

$$\delta_E^{(n_b,n_p)}(N_{\text{tot}})$$
$$\equiv 1 - \left\langle F_{N_{\text{tot}}}^{(n_b,n_p)} \right\rangle \quad (20)$$
$$= \sum_{n_b'=0}^{n_b} \sum_{n_p'=0}^{n_p} \varepsilon_{E,\text{infid}}^{(n_b',n_p')} P(N_{\text{tot}})_{(n_b+1)n_p'+n_b'+2} + \frac{1}{2} P(N_{\text{tot}})_1.$$
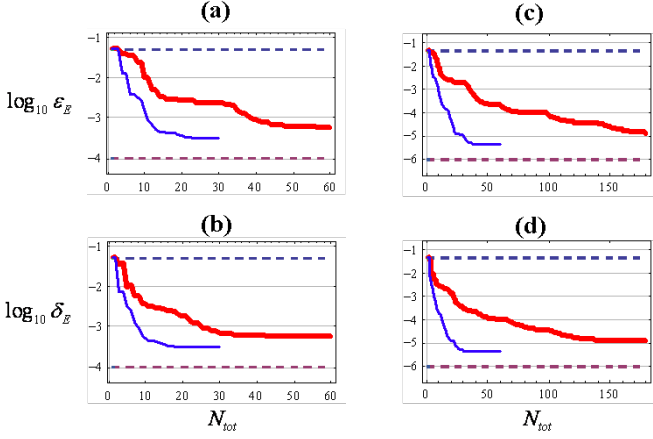
FIG. 13: (Color online) The optimized total error probability ε_E [Eq. (21)] (upper plots) and the optimized average infidelity δ_E [Eq. (22)] (lower plots) as a function N_{tot}. The error probability for the local coupling gates is $p_L = 10^{-4}$ (left plots) and $p_L = 10^{-6}$ (right plots). One-level pumping is used for dephasing error (thin blue curves); two-level pumping is used for depolarizing error (thick red curves). The other parameters $F = 0.95, p_I = p_M = 5\%$ are the same for all plots. Both ε_E and δ_E saturate for large N_{tot}. In each plot, we also show the initial infidelity $1 - F$ (upper blue dashed lines) and the local error probability p_L (lower violet dashed lines).

Here the first term sums over all states of the Markov chain (except for the initial one), each of which has at least one partially purified pair with infidelity $\varepsilon_{E,\text{infid}}^{(n_b', n_p')}$ and probability $P(N_{\text{tot}})_{(n_b+1)n_p' + n_b' + 2}$; the last term comes from the situation that none of the partially purified Bell pairs remain after the last attempt of the entanglement purification and we just use a classically correlated pair with infidelity $1/2$. Generally, the average infidelity is smaller than the total error probability.

We may also optimize the choice of the control parameters (n_b, n_p)

$$\varepsilon_E(N_{\text{tot}}) \equiv \min_{(n_b, n_p)} \varepsilon_E^{(n_b, n_p)}(N_{\text{tot}}) \qquad (21)$$

and

$$\delta_E(N_{\text{tot}}) \equiv \min_{(n_b, n_p)} \delta_E^{(n_b, n_p)}(N_{\text{tot}}). \qquad (22)$$

In Fig. 13, we plot both the optimized total error probability ε_E and the optimized average infidelity δ_E as a function of N_{tot}. Both quantities asymptotically approaches the same minimum value

$$\lim_{N_{\text{tot}} \to \infty} \varepsilon_E(N_{\text{tot}}) = \lim_{N_{\text{tot}} \to \infty} \delta_E(N_{\text{tot}}) = \Delta_{\min}. \qquad (23)$$

Here the minimum value is simply the minimal infidelity of the entanglement purification

$$\Delta_{\min} \equiv \min_{(n_b, n_p)} \varepsilon_{E,\text{infid}}^{(n_b, n_p)} \qquad (24)$$

which is achieved by the control parameters $(n_b, n_p) \equiv (n_b^*, n_p^*)$, for the imperfection parameters $\{p_L, 1 - F, \varepsilon_M\}$.

We remark that a faster and less resource intensive approach may be used if the unpurified Bell pair is dominated by dephasing error. Then, one-level pumping is sufficient (i.e. no bit-error purification, $n_b = 0$). The optimized total error probability and average infidelity (thin blue curves) for this situation are plotted as a function of N_{tot} in Fig. 13.

C. Total time for robust entanglement generation

The total time for robust entanglement generation $\tilde{t}_E$ is proportional to the average number of raw Bell pairs generated $\langle N_{\text{tot}} \rangle$

$$\tilde{t}_E \approx \langle N_{\text{tot}} \rangle \times (t_E + t_L + \tilde{t}_M), \qquad (25)$$

where t_E is the average generation time of the unpurified Bell pair. Note that the entanglement generation itself is a stochastic process. In principle, we may also include the stochastic nature of the entanglement generation by introducing a sub-level of Markov chain to characterize the stochastic entanglement generation. Since each entanglement generation either succeeds or fails, the sub-level Markov chain only involves two states, which can be easily incorporated into the Markov chain models discussed above. After incorporating the sub-level into the Markov chain, each transition corresponds to one attempt of entanglement generation, instead of one attempt of entanglement purification that consumes one unpurified Bell pair previously.

Nevertheless, the number of Bell pairs generated in a given period of time (i.e. N_{tot}) has a distribution. Since the relative deviation of this distribution ($\sim N_{\text{tot}}^{-1/2}$) is fairly small for large N_{tot} (> 20), this only has a minor influence. Thus we replace $\langle N_{\text{tot}} \rangle$ by N_{tot}.

D. Trade-off between gate quality and time overhead

We now consider the balance between the "quality" of the robustly generated entangled pairs and the time overhead N_{tot} associated with the robust generation process. We may use either the optimized total error probability $\varepsilon_E(N_{\text{tot}})$ or the optimized average infidelity $\delta_E(N_{\text{tot}})$ to characterize the quality. Since both quantities approaches the same asymptotic minimum $\Delta_{\min}$ according to Eq. (23), there is only little improvement in the quality of the robust entanglement generation once $\varepsilon_E(N_{\text{tot}})$ or $\delta_E(N_{\text{tot}})$ is comparable to $\Delta_{\min}$ (say $2\Delta_{\min}$). Thus, we find the value for N_{tot} by imposing the relations
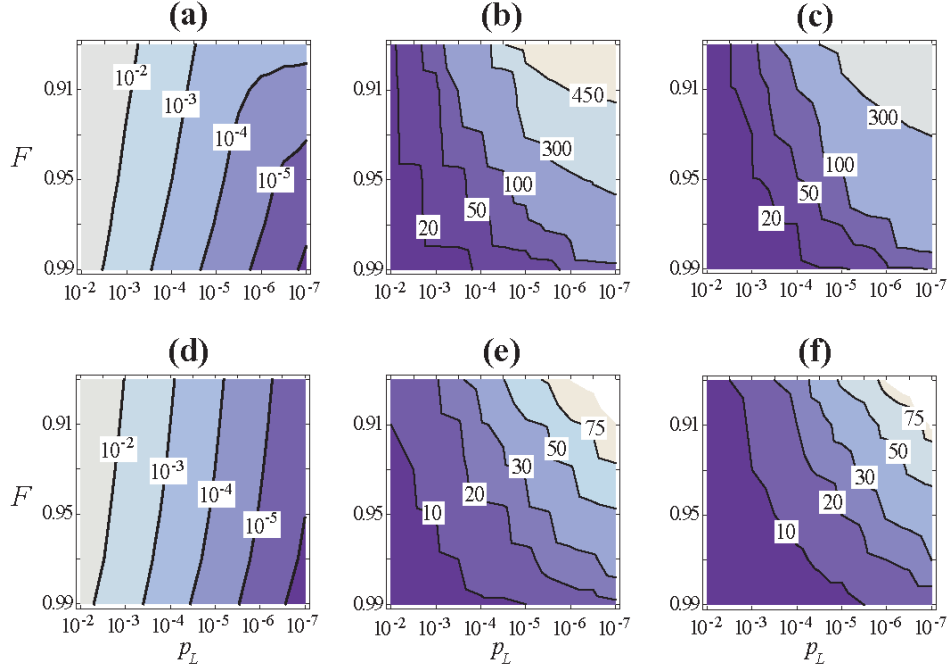
$$\varepsilon_E(N_{\text{tot}}) = 2\Delta_{\min} \qquad (26)$$

FIG. 14: (Color online) Contours of the total error probability ε_E (or average infidelity δ_E) after purification (left), the total number of unpurified Bell pairs N_{tot} associated with ε_E [Eq. (26)] (middle), and N_{tot} associated with δ_E [Eq. (27)] (right). The contours are drawn with respect to the imperfection parameters p_L (horizontal axis) and F (vertical axis). Two-level pumping (up) is used for depolarizing error, and one-level pumping (down) for dephasing error. $p_I = p_M = 5\%$ is assumed.

or

$$\delta_E\left(N_{\text{tot}}\right) = 2\Delta_{\min}. \qquad (27)$$

First, we consider the total error probability $\varepsilon_E\left(N_{\text{tot}}\right)$. The relation in Eq. (26) can be simplified, if we assume fixed control parameters $(n_b, n_p) \equiv \left(n_b^*, n_p^*\right)$ for the left hand side (rather than minimizing over all possible choices of (n_b, n_p)). Combined with Eq. (19), the failure probability should be comparable to the minimal infidelity

$$\varepsilon_{E,\text{fail}}^{\left(n_b^*, n_p^*\right)}\left(N_{\text{tot}}\right) \approx \Delta_{\min}. \qquad (28)$$

Since both the variable $\Delta_{\min}$ and the parameters $\left(n_b^*, n_p^*\right)$ depend on $\{p_L, p_I, p_M, 1 - F\}$, the above relation implicitly determines N_{tot} as a function a function of $\{p_L, p_I, p_M, 1 - F\}$.

In Fig. 14, we plot the contours of ε_E [Eq. (26)] and N_{tot} [Eq. (28)] with respect to the imperfection parameters p_L and $1 - F$, while assuming $p_I = p_M = 5\%$. Actually the choice of p_I and p_M ($< 10\%$) has negligible effect on the contours, since they only modify ε_M marginally. For initial fidelity $F_0 > 0.95$, the contours of ε_E are very close to vertical lines; that is ε_E is mostly limited by p_L with an overhead factor (about 10) very insensitive to F_0. The contours of N_{tot} indicate that the entanglement pumping needs about tens or hundreds of raw Bell pairs to ensure a very high success probability.

Similarly, we may also numerically obtain the value N_{tot} from Eq. (27). The contour plot of N_{tot} with respect to the imperfection parameters p_L and $1 - F$ is also shown in Fig. 14 (c,f). We compare N_{tot}s obtained from two estimates (total error probability [Eq. (26)] and average infidelity [Eq. (27)]). As we expected, the N_{tot} obtained from total error probability is approximately $1.2 \sim 2$ time larger than the N_{tot} obtained from average infidelity, since the former is a more conservative estimate and requires more unpurified Bell pairs. Nevertheless, the difference is small and can be easily accounted by a prefactor of order unity. For clarity, in the rest of the paper we will use the N_{tot} estimated by using total error probability, and sometimes quote the values estimated by using average infidelity.

E. Entanglement pumping with non-post-selective (NPS) scheme

We now consider another entanglement pumping protocol, proposed by Campbell [59]. The entanglement pumping scheme we have considered so far is post-selective (PS); that is we discard the Bell pair if one step of entanglement pumping is not successful. However, the Bell pair may still be highly entangled even if the entanglement pumping failed at some intermediate step. The non-post-selective (NPS) entanglement pumping scheme
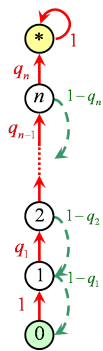
FIG. 15: (Color online) Markov chain model for one-level entanglement pumping with non-post-selective (NPS) scheme. The key difference from the previous Markov chain model with post-selective (PS) pumping scheme (see Fig. 10) is that here the transition for unsuccessful pumping reduces the chain label (score) by 1, while in the previous model the transition for unsuccessful pumping goes back to state "0" (restart of the entire pumping scheme).
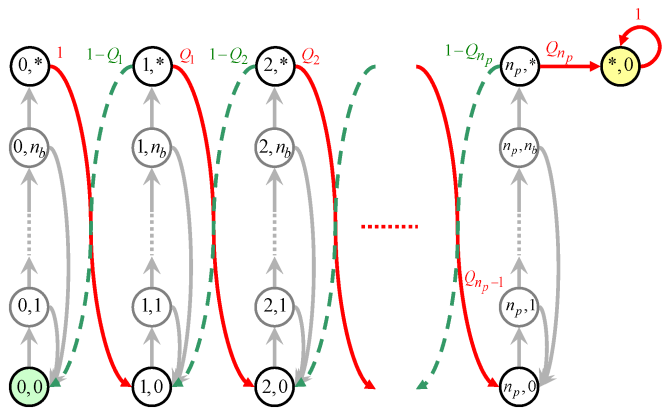

FIG. 16: (Color online) Markov chain model for two-level entanglement pumping with NPS scheme. We still use PS entanglement pumping scheme at the first level to have minimal accumulation of phase errors. Only at the second level, does the NPS scheme work more efficiently than the original PS scheme (see Fig. 11).

[59] keeps track of the evolution of the density matrix of the Bell pair after each step of pumping. The NPS scheme avoids the inefficient restart (i.e., discarding intermediately purified Bell pairs), and it may reduce the time overhead, especially when the unpurified Bell pairs have relatively low fidelity ($F < 0.9$). In Ref. [59], the NPS pumping is discussed in the context of generating a graph state.

We now describe how to use the NPS pumping scheme to generate purified Bell pairs. To simplify the discussion, we first assume that the errors from local measurements and operations are negligible. This assumption enables us to establish a connection between the Markov chain model and the NPS pumping scheme.

Suppose the unpurified Bell pairs have only phase errors, then one level of entanglement pumping is sufficient. For this error model, one can show that a failed attempt produces an EPR pair with a density matrix identical to the one in the previous step [59]. One may introduce an accumulated score associated with entanglement pumping. The score increases by one unit for each attempt of successful pumping, and decrease by one unit for an attempt of unsuccessful pumping. The score for no Bell pair is 0, and for one unpurified Bell pair it is 1. The score exactly corresponds to the state label of the Markov chain (see Fig. 15). After each attempt of pumping, the score changes by ± 1. If the score drops to 0 (i.e. no Bell pair left), it gets back to 1 in the next attempt (i.e., creating a new unpurified Bell pair). The pumping procedure continues, until the score reaches $n + 1$ (i.e., the final state "$*$" in the Markov chain). The key different from the previous Markov chain for post-selective pumping scheme (see Fig. 10) is that here the score decrease by 1 for unsuccessful pumping rather than restart from 0. This modification increases the success probability of the robust entanglement generation.

When the unpurified Bell pairs have both bit-flip and phase errors (e.g., depolarizing error), we may use the bit-phase two-level pumping scheme (see Sec. VI B), which purifies the bit error at the first level and then the phase error at the second level. Since the phase error is not purified at the first level, it accumulates after each attempt of pumping. Therefore, it is better to use PS entanglement pumping scheme at the first level to have minimal accumulation of phase errors. At the second level, the NPS scheme works more efficiently than the PS scheme. The Markov chain circuit for such mixed PS-NPS pumping schemes is shown in Fig. 16.

In practice, the error probability for the local operations is always finite. Then our simple Markov chain model only provides an approximate description for the real process. The approximation comes from the fact that the score is now insufficient to specify the density matrix for intermediate Bell pairs, in the presence of local operational errors. In order to obtain the density matrix for the intermediate state, we need to have the entire list of all previous pumping outcomes. Nevertheless, when the local operational errors are small compared to the infidelity of the intermediate Bell pairs, the Markov chain model still provides an (optimistic) estimate for the total error probability and the average fidelity.

We now compare the N_{tot}s associated with the PS and NPS schemes. The contours of the ratio between the two N_{tot}s is plotted as a function of p_L and F in Fig. 17. As pointed out in Ref. [59], there is a significant improvement by using the NPS scheme (more than a factor of 3), for $F < 0.9$ and $p_L < 10^{-4}$.
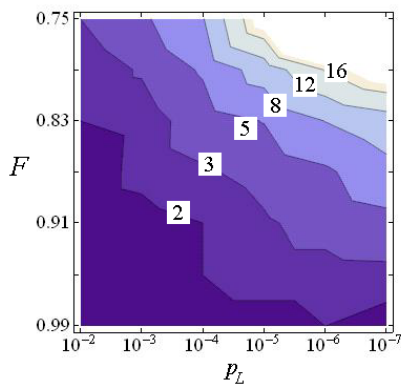
FIG. 17: (Color online) The contours for the ratio between N_{tot} associated with the post-selective (PS) scheme and N_{tot} associated with the non-post-selective (NPS) scheme, as a function of p_L and F. For both schemes, we use the same formula [Eq. (27)], but different Markov chain models (Fig. 10 and 15). The improvement from the NPS scheme becomes significant (more than a factor of 3), for $F < 0.9$ and $p_L < 10^{-4}$.

VIII. MAPPING TO DETERMINISTIC MODEL

In this section, we will map our stochastic, hybrid, and distributed quantum computation scheme to a deterministic computation model, which is characterized by two parameters — the clock cycle and the effective error probability. We will show that even when the underlying operations such as the entanglement generation are non-deterministic, our approach still maintains reasonable fast clock cycle time and sufficiently low effective error probability. We will associate our discussion with achievable experimental parameters, consider the constraint set by the finite memory lifetime, and determine the achievable performance of hybrid distributed quantum computation.

A. Time and error in the theoretical model

All the previous discussions can be summarized in terms of the *clock cycle time*

$$t_C = \tilde{t}_E + 2t_L + \tilde{t}_M \approx \tilde{t}_E , \qquad (29)$$

and the *effective error probability*

$$\gamma = \varepsilon_E + 2p_L + 2\varepsilon_M, \qquad (30)$$

for a general coupling gate between two registers.

We now provide an estimate of the clock cycle time based on realistic parameters. The time for optical initialization/measurement is

$$t_I = t_M \approx \frac{\ln p_M}{\ln(1-\eta)} \frac{\tau}{C}, \qquad (31)$$

with a photon collection/detection efficiency η, vacuum radiative lifetime τ, and the cooperativity (Purcell) factor C for cavity-enhanced radiative decay [60, 61]. Eq. (31) is obtained from the estimate for the measurement error probability $p_M \approx (1-\eta)^{N_{\text{photon}}}$ with $N_{\text{photon}} \approx t_M/(\tau/C)$. We assume that the entanglement is generated based on detection of two photons [3, 32], which takes time

$$t_E \approx (t_I + \tau/C)/\eta^2. \qquad (32)$$

Generally entanglement fidelity is higher for the two-photon schemes than one-photon schemes [31]. In addition, some two-photon schemes have intrinsic purification against bit-flip errors [55]. The time for robust measurement is given in Eq. (9), and the total time for robust entanglement generation is given in Eq. (25).

Combining Eqs. (29), (31), (32), (9) and (25), we obtain the clock cycle time (in units of the local operation time) as a function of other parameters

$$\frac{t_C}{t_L} = f\left[\frac{\tau}{t_L C}, p_M, \eta, m, N_{\text{tot}}\right]. \qquad (33)$$

Meanwhile, we may obtain the relation $m = m[p_L, p_I, p_M]$ by minimizing ε_M with Eq. (8), and find the relation $N_{\text{tot}} = N_{\text{tot}}[p_L, F, \varepsilon_M] = N_{\text{tot}}[p_L, F, 2\Delta_{\min}[p_L, p_I, p_M]]$ using Eqs. (24 and 26). Therefore, we have

$$\frac{t_C}{t_L} = f\left[\frac{\tau}{t_L C}, p_M, \eta, p_L, F\right]. \qquad (34)$$

The dimensionless parameter is the ratio between the times of emitting a single photon and performing a local unitary operation. For systems such as ion-traps and NV centers, this ratio is usually much less than unity (< 0.01).

Similarly, we can obtain the effective error probability in terms of imperfection parameters

$$\gamma = g[p_L, p_M, F], \qquad (35)$$

by combining Eqs. (8), (30) and (26).

In Fig. 18, we plot the clock cycle time t_C and effective error probability γ, for two-level pumping against depolarizing error. Assuming $\eta = 0.2$, we consider the two choices of parameters $1 - F = p_I = p_M = 5\%$ (left) and 1% (right). For each case, we plot the contours of the normalized clock cycle time t_C/t_L as a function of p_L and $\frac{\tau}{t_L C}$, and the effective error probability γ as a function of p_L. The clock cycle time can be reduced by having a fast radiative decay rate τ/C, which can be facilitated by having a large cooperativity factor C. The reduction of the clock cycle time stops once this ratio is below certain value, approximately 0.003 (left) and 0.001 (right), where local gate operation becomes the dominant time consuming step. Similarly, we plot the clock cycle time t_C and effective error probability γ, for one-level pumping against dephasing error in Fig. 19.
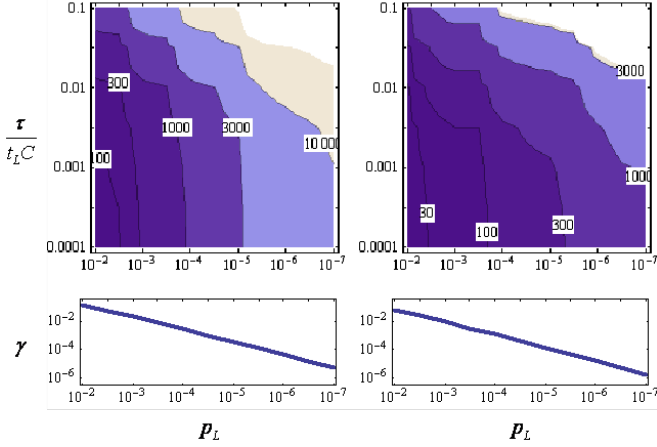
FIG. 18: (Color online) Plots of clock cycle time t_C and effective error probability γ, for two-level pumping against depolarizing errors. Upper plots: contours of the normalized clock cycle time t_C/t_L (upper) as a function of p_L and $\frac{\tau}{t_L C}$ (the normalized effective radiative lifetime). Lower plots: the effective error probability γ as a function of p_L. We assume $1 - F = p_I = p_M = 5\%$ (left) and 1% (right), and $\eta = 0.2$.

FIG. 19: (Color online) Plots of clock cycle time t_C and effective error probability γ, for one-level pumping against dephasing error. Upper plots: contours of the normalized clock cycle time t_C/t_L (upper) as a function of p_L and $\frac{\tau}{t_L C}$ (the normalized effective radiative lifetime). Lower plots: the effective error probability γ as a function of p_L. We assume $1 - F = p_I = p_M = 5\%$ (left) and 1% (right). And $\eta = 0.2$.

In the limit of negligible radiative decay time, we obtain the lower bound for the normalized clock cycle time

$$\lim_{\frac{\tau}{t_L C} \to 0} t_C/t_L = \begin{cases} N_{\text{tot}} & \text{for } m = 0 \\ (2m + 2) N_{\text{tot}} & \text{for } m \geq 1 \end{cases}, \qquad (36)$$

where for $m \geq 1$ there is a time overhead $2m + 2$ associated with local operation and robust measurement; while there is no such overhead for $m = 0$.

B. Estimated numbers for experimental setups

Suppose the parameters are $(t_L, \tau, \eta, C) = (0.1\ \mu s, 10\ \text{ns}, 0.2, 10)$ [62, 63, 64] and $(1 - F, p_I, p_M, p_L, \varepsilon_M) = (5\%, 5\%, 5\%, 10^{-4}, 8 \times 10^{-4})$ for our quantum registers (based on ion-traps or NV centers). For depolarizing errors, two-level pumping can achieve $(t_C, \gamma) = (200\ \mu s, 2.7 \times 10^{-3})$. For some entanglement generation schemes [3, 32, 55] in principle only dephasing error exists, because they have intrinsic purification against bit-flip errors. If all bit-flip errors are suppressed, then one-level pumping is sufficient and $(t_C, \gamma) = (42\ \mu s, 2.2 \times 10^{-3})$ [74]. In Table I, we have listed (t_C, γ) for parameters $1 - F = p_M = p_I = 5\%$ or 1%, and $p_L = 10^{-3}, 10^{-4}, 10^{-5}$, or 10^{-6}. As expected, we find that t_C gets longer, if the fidelity F is lower and/or the error probability (p_M or p_I) is higher; t_C significantly reduces if the error for the unpurified Bell pairs changes from depolarizing error to dephasing error.

We remark that t_C should be much shorter than the memory time of the storage qubit, t_{mem}. Because the
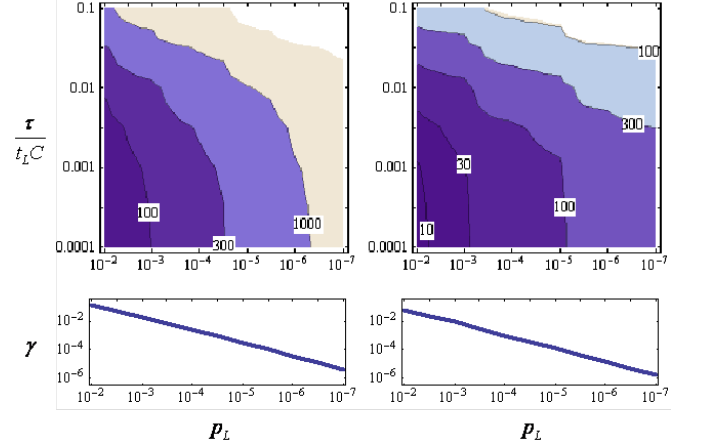
memory error probability for each clock cycle is approximately t_C/t_{mem}, which should be small (say 10^{-4}) in order to achieve fault-tolerant quantum computation. This is indeed the case for both trapped ions (where $t_{mem} \sim 10$ s has been demonstrated [43, 44]), as well as for proximal nuclear spins of NV centers (where t_{mem} approaching a second can be inferred [12]). So far, we have justified the feasibility of the hybrid distributed quantum computation scheme. In the next subsection, we will provide a criterion for hybrid distributed quantum computation.

C. Constraints from finite memory life time

Above we have mostly ignored the effect of finite memory time, and with the various sequences of purification of imperfections the final fidelity of the operations have then been limited only be the local operation. All of these purifications, however, increase the time of the operations and eventually the system may become limited by the finite life time of the memory. In this subsection we shall evaluate this constraint set by the finite memory lifetime.

To simplify the discussion we assume that we have a very short radiative lifetime τ or that we are able to achieve a very large Purcell factor so that τ/C becomes negligible. All the time scales are then proportional to the local gate time t_L. With a finite memory time, i.e., some fixed t_{mem}/t_L, there is a limit to have many operations we can do before we are limited by the memory error. To get an estimate for this limit we assume that the ideal number of operations is roughly given by the

	Depolarizing				Dephasing			
	$F = 0.95$		$F = 0.99$		$F = 0.95$		$F = 0.99$	
	$t_C(\mu s)$	γ	$t_C(\mu s)$	γ	$t_C(\mu s)$	γ	$t_C(\mu s)$	γ
$p_L = 10^{-3}$	65	1.9×10^{-2}	19	9.1×10^{-3}	20	1.7×10^{-2}	8	8.7×10^{-3}
$p_L = 10^{-4}$	200	2.7×10^{-3}	49	1.2×10^{-3}	42	2.2×10^{-3}	17	9.9×10^{-4}
$p_L = 10^{-5}$	387	3.5×10^{-4}	65	1.3×10^{-4}	80	2.8×10^{-4}	22	1.2×10^{-4}
$p_L = 10^{-6}$	997	4.5×10^{-5}	162	1.7×10^{-5}	140	3.4×10^{-5}	39	1.4×10^{-5}

TABLE I: We list the values of t_C and γ as a function of p_L (rows) and F (columns) for depolarizing and dephasing errors of the unpurified Bell pairs. We also assume $p_M = p_I = 1 - F$, and $(t_L, \tau, \eta, C) = (0.1\ \mu s, 10\ ns, 0.2, 10)$. Note that t_C estimated by using average infidelity is approximately $1.3 \sim 1.6$ times less than the numbers listed here.

point, where the memory error probability is the same as the effective error probability for the non-local coupling gate:

$$t_C/t_{mem} = \gamma. \qquad (37)$$

Then according to Eq. (36), we have

$$\frac{t_{mem}}{t_L} = \frac{t_{mem}}{t_C}\frac{t_C}{t_L} = \gamma^{-1}\left(2m + 2 - \delta_{m,0}\right)N_{\text{tot}}, \qquad (38)$$

where the variables $\{\gamma, m, N_{\text{tot}}\}$ are all determined by the imperfection parameters $\{1 - F, p_M, p_I, p_L\}$. We further reduce the imperfection parameters by assuming $1 - F = p_M = p_I$, and get the contour plot of t_{mem}/t_L in terms of the imperfection parameter p_L and $1 - F$ in Fig. 20. In the plot, we consider both the situation of depolarizing or dephasing error during entanglement generation.

For given t_{mem}/t_L, we may use Fig. 20 to find the valid region in the parameter space of p_L and $1 - F$, and then identify the achievable effective error probability γ. For example, with ion-trap systems it may be possible to achieve $t_{mem}/t_L \sim 10^8$ [43, 44, 62], and the region left of the shaded contour line ($\log_{10} t_{mem}/t_L = 8$) can then be accessed, which enables us to obtain a wide range effective error probability γ depending on the practical values of p_L and $1 - F$. For NV centers, it should be feasible to achieve $t_{mem}/t_L \sim 10^7$ by having $t_{mem} \approx 10$ s and $t_L \sim 10^{-6}$ s [12]; the region on the left side of the shaded contour line ($\log_{10} t_{mem}/t_L = 7$) still covers a large portion of the parameter space. For a given experiment situation with a finite memory time as well as other imperfections, we can thus use Fig. 20 to determine the achievable performance of hybrid distributed quantum computation.

IX. APPROACHES TO FAULT TOLERANCE

The entanglement based approach discussed in this paper provides a method to make gates between any quantum registers and this can be used to implement arbitrary quantum circuits, once the errors in the gates are sufficiently small. The errors can be further suppressed by using quantum error correction. For example, as shown in Table I, $(p_L, F) = \left(10^{-4}, 0.95\right)$ can achieve $\gamma \approx 2.7 \times 10^{-3}$, well below the 1% threshold for fault tolerant computation based on approaches such as the C_4/C_6 code [65] or 2D toric codes [66]; $(p_L, F) = \left(10^{-6}, 0.99\right)$ can achieve $\gamma \approx 1.7 \times 10^{-5}$, which allows efficient codes such as the BCH $[[127,43,13]]$ code to be used without concatenation.

Following Ref. [67] we estimate 20 registers per logical qubit to be necessary for a calculation involving $K = 10^4$ logical qubits and $Q = 10^6$ logical operations, assuming the memory failure rate and effective error probability are $t_C/t_{mem} \approx \gamma \approx 1.7 \times 10^{-5}$ (e.g., achieved by $t_{mem} \approx 10$ s, $t_C \approx 162\ \mu s$). (This estimate is based on Fig. 10b of Ref. [67].) Assuming that error correction is applied after each logical operation, and that logical operation and following recovery take approximately $2 - 16$ clock cycles depending on the type of operation and the coding scheme (see section II.A of Ref. [67]), the total running time of this computation would then be approximately $400 - 3000$ s.

We remark that one important property of distributed quantum computation is that the measurement time is relatively fast compared with the non-local coupling gate, because the measurement does not rely on the time-consuming processes of entanglement generation and purification while the non-local coupling gate does. This property is different from the conventional model of quantum computation, where the measurement is usually a slow process that induces extra overhead in both time and physical resources [67]. Thus, instead of reconciling slow measurements [68], it might also be interesting to study possible improvement using fast measurements for fault-tolerant quantum computation.

The above estimates have been performed assuming that our hybrid register based approach is mapped directly to the standard circuit model. In some situations this may, however, not be the most advantageous way to proceed, since the register architecture may allow for more efficient performance of certain tasks. As a particular example, we now briefly discuss an alternative new approach to fault-tolerant preparation of GHZ states (e.g., $|00\cdots0\rangle + |11\cdots1\rangle$), which are a critical component both for syndrome extraction and construction of universal gates in quantum error correcting codes [65, 69]. This
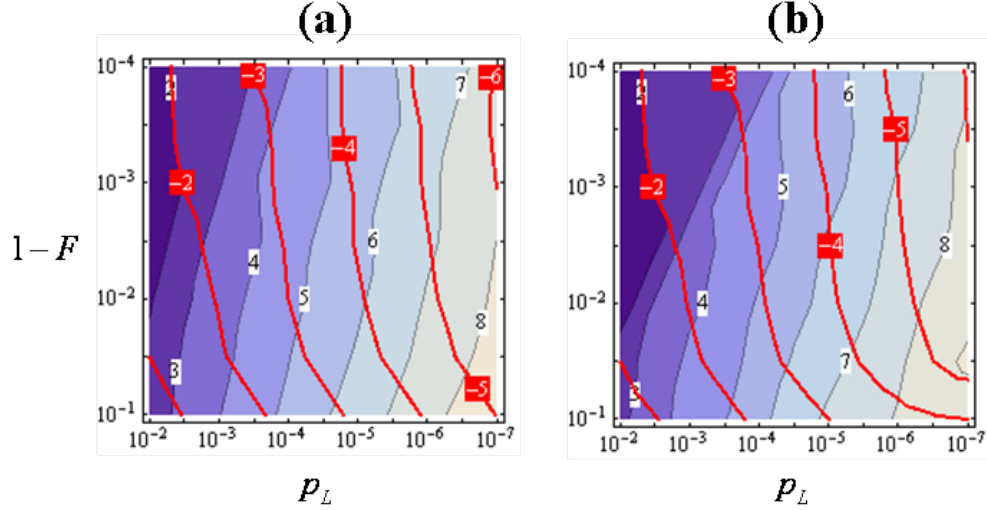
FIG. 20: (Color online) Contours for t_{mem}/t_L (shaded contours) and γ (red contours), as a function of p_L and $1 - F$. The label for the contour values are in logarithmic scale with base 10. We consider both situations of (a) depolarizing error and (b) dephasing error. The other parameters are $p_M = p_I = 1 - F$ and $\tau/C = 0$.
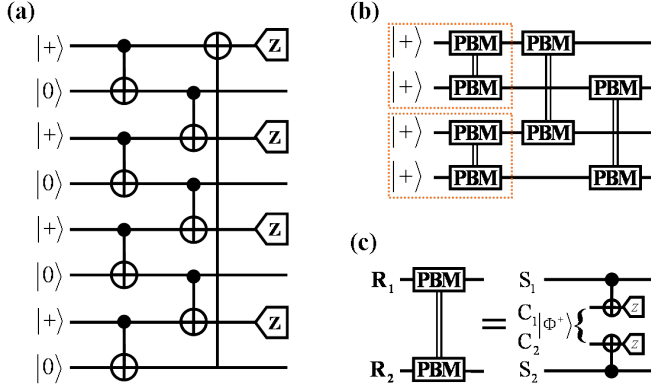

FIG. 21: (Color online) Circuits for fault-tolerant preparation of GHZ state. (a) Conventional circuit uses 8 quantum registers to generate a 4-qubit GHZ state. The state $|+\rangle$ is $(|0\rangle + |1\rangle)/\sqrt{2}$. (b) New circuit requires only 4 registers, by using partial Bell measurement (PBM) between registers, detailed in (c).

new approach relies upon the observation that the EPR pairs from entanglement generation can be used for *deterministic* partial Bell measurement (PBM), which is achieved by applying local coupling gates and projective measurements as shown in Fig. 21c. After the PBM, the two storage qubits are projected to the subspace spanned by the Bell states $|\Phi^\pm\rangle = (|00\rangle \pm |11\rangle)/\sqrt{2}$ if the measurement outcomes are the same, or they are projected to the subspace spanned by the Bell states $|\Psi^\pm\rangle = (|01\rangle \pm |10\rangle)/\sqrt{2}$ if the measurement outcomes are different. For the latter case, we may further flip one of the storage qubits, so that they are projected to the

subspace spanned by $|\Phi^\pm\rangle$. By using PBMs, we can perform fault-tolerant preparation of GHZ state efficiently (up to single qubit rotations) as detailed below.

Fault-tolerant state preparation requires that the probability to have errors in more than one qubit in the prepared state is $O\left(p^2\right)$, with the error probability for each input qubit or quantum gate being $O\left(p\right)$; that is multiple errors only occur at the second or higher orders. The regular circuit to prepare a four-qubit GHZ state fault-tolerantly [65] is shown in Fig. 21a. If this circuit should be implemented with quantum registers, the CNOT gates in Fig. 21a should be created by using the circuit detailed in Fig. 3, and eight quantum registers would be required.

By using PBMs, however, only *four* quantum registers are needed in order to generate GHZ states fault-tolerantly as shown in Fig. 21b. The fault-tolerance comes from the last (redundant) PBM between the second and fourth register (Fig. 21b), which detects bit-errors from earlier PBMs. The advantage of the PBMs is that it propagates neither bit- nor phase-errors. The circuit of Fig. 21c indicates that the only way to propagate error from one input to another (say, S_1 to S_2) is via some initial error in the entangled pair between C_1 and C_2. However, for Bell states $|\Phi^\pm\rangle$, we have the following identities

$$X_{C_2} \left|\Phi^\pm\right\rangle_{C_1,C_2} = \pm X_{C_1} \left|\Phi^\pm\right\rangle_{C_1,C_2} \qquad (39)$$

$$Z_{C_2} \left|\Phi^\pm\right\rangle_{C_1,C_2} = Z_{C_1} \left|\Phi^\pm\right\rangle_{C_1,C_2}. \qquad (40)$$

Similar identities also exist for Bells states $|\Psi^\pm\rangle$. Suppose S_1 has an error, because of the above identities, we can always treat the imperfection of the Bell pair $|\Phi^+\rangle_{C_1,C_2}$ as an error in C_1 (the qubit from the same register as S_1). Therefore, only the first register has errors and they
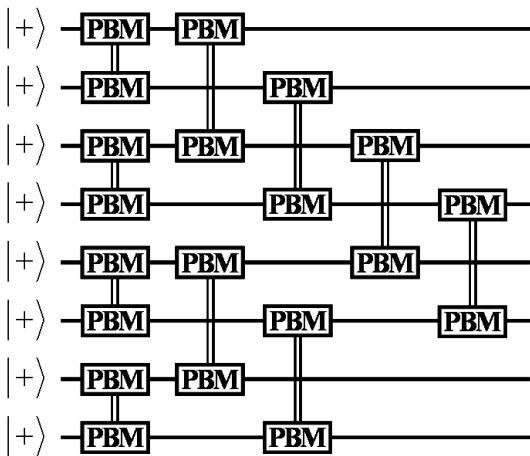
FIG. 22: Circuit for fault-tolerant preparation of the 8-qubit GHZ state with only 8 quantum registers.

FIG. 23: (Color online) The contours of the infidelity $\varepsilon_{E,\text{infid}}^{(n_b, n_p)}$ as a function of n_b and n_p. We use $\log_{10} \varepsilon_{E,\text{infid}}^{(n_b, n_p)}$ to label the contours. We assume a depolarizing error with initial fidelity $F = 0.95$, and $\tilde{\varepsilon}_M = p_L = 0$. The final infidelity can be arbitrarily small for sufficiently large n_b and n_p. This indicates that the bit-phase two-level entanglement pumping scheme can create pairs with arbitrarily high fidelity.

never propagate to the second one. (PBM may induce errors to unconnected but entangled qubits.)

The present scheme may be expanded to larger numbers of qubits and generally we may fault-tolerantly prepare 2^n-qubit GHZ state with only 2^n quantum registers, by recursively using Fig. 21b with the two dashed boxes replaced by two 2^{n-1}-qubit GHZ states. The circuit for fault-tolerant preparation of the 8-qubit GHZ state is shown in Fig. 22. Note that we can perform PBMs acting on different registers in parallel. Suppose each PBM takes one clock cycle, the preparation time is 2 (clock cycles) for a 4-qubit GHZ state shown in Fig. 21b. The two PBMs in the orange boxes are performed in the first clock cycle, and the rest for the second clock cycle. Generally, for a 2^n-qubit GHZ state with $n \geq 3$ (see discussion in Appendix C), the preparation time is only 3 (clock cycles), and the error probability for each register is only approximately $3p/2$. Therefore, the PBM-based scheme for fault-tolerant preparation of the GHZ state is efficient in both time and physical-resources.

X. CONCLUSION

In conclusion, we have proposed an efficient register-based, hybrid quantum computation scheme. Our scheme requires only five qubits (or less) per register, and it is robust against various kinds of imperfections, including imperfect initialization/measurement and low fidelity entanglement generation. We presented a Markov chain model to analyze the time overhead associated with the robust operations of measurement and entanglement generation. We found reasonable time overhead and considered practical implementation of quantum registers with ion traps or NV centers. We also provided an example using partial Bell measurement to prepare GHZ states for fault-tolerant quantum computation. It might be possi-
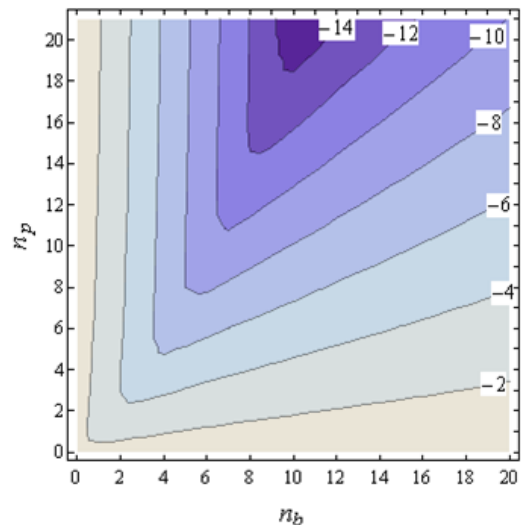
ble to further facilitate fault-tolerant quantum computation with systematic optimization using dynamic programming [70].

APPENDIX A: BIT-PHASE TWO-LEVEL PUMPING SCHEME

In this Appendix we show that our bit-phase two-level entanglement pumping scheme can create pairs with fidelity arbitrarily close to unity, if we have perfect local operations. Numerical indication of this is shown in Fig. 23, and in the following we will provide a rigorous proof to this claim.

We assume that the initial state is a mixed state that has only diagonal terms in the Bell basis

$$\rho = a \left|\Phi^+\right\rangle\left\langle\Phi^+\right| + b \left|\Phi^-\right\rangle\left\langle\Phi^-\right| + c \left|\Psi^+\right\rangle\left\langle\Psi^+\right| + d \left|\Psi^-\right\rangle\left\langle\Psi^-\right|,$$
(A1)

where $|\Phi^{\pm}\rangle = (|00\rangle \pm |11\rangle)/\sqrt{2}$, $|\Psi^{\pm}\rangle = (|01\rangle \pm |10\rangle)/\sqrt{2}$, and the coefficients are non-negative and sum to unity. (This assumption is only made to simplify the presentation. For a general density matrix, only the diagonal elements given in Eq. (A1) are important [34].) After the purification the density matrix retains this form but with new coefficients. Therefore, we only need four coefficients for each state using Bell basis, denoted as the fidelity vector $\vec{F} = (a, b, c, d)$. We use the lower index to keep track of the pumping steps, so that the fidelity vector for the unpurified state is $\vec{F}_0 = (a_0, b_0, c_0, d_0)$, and the vector for the purified state after n steps of entanglement pumping is $\vec{F}_n = (a_n, b_n, c_n, d_n)$.

Suppose we use the $\vec{F}_0$ state to pump the state $\vec{F}_n$ against bit-errors, the success probability is

$$p_{n+1} = (a_0 + b_0)(a_n + b_n) + (c_0 + d_0)(c_n + d_n), \quad \text{(A2)}$$

and the fidelity vector is

$$\vec{F}_{n+1} \qquad \qquad \text{(A3)}$$
$$= \frac{1}{p_{n+1}}(a_0 a_n + b_0 b_n, a_0 b_n + b_0 a_n, c_0 c_n + d_0 d_n, c_0 d_n + d_0 c_n), \quad \text{(A4)}$$

for perfect local operations (measurement and CNOT gate).

Similarly, for pumping against phase-errors, the success probability is

$$p'_{n+1} = (a_0 + c_0)(a_n + c_n) + (b_0 + d_0)(b_n + d_n), \quad \text{(A5)}$$

and the fidelity vector is

$$\vec{F}'_{n+1} \qquad \qquad \text{(A6)}$$
$$= \frac{1}{p'_{n+1}}(a_0 a_n + c_0 c_n, b_0 b_n + d_0 d_n, a_0 c_n + c_0 a_n, b_0 d_n + d_0 b_n). \quad \text{(A7)}$$

In general, any state can be turned into a so-called Werner state with the same fidelity, and as a worst case scenario we shall assume the unpurified Bell state to be a Werner state:

$$\vec{F}_0 = \left(F_0, \frac{1-F_0}{3}, \frac{1-F_0}{3}, \frac{1-F_0}{3}\right). \quad \text{(A8)}$$

with $F_0 > 1/2$ to ensure that it contains distillable entanglement. For convenience of later discussion, we rewrite $\vec{F}_0$ as

$$\vec{F}_0 = (1 - 3\alpha/2, \alpha/2, \alpha/2, \alpha/2) \quad \text{(A9)}$$

where $\alpha = \frac{2}{3}(1 - F_0) < \frac{1}{3}$.

1. First level pumping

For the first level of pumping against bit-errors, we may parameterize the fidelity vector

$$\vec{F}_n = (a_n, b_n, c_n, d_n) = \left(\frac{1}{2} + \delta_n, \frac{1}{2} - \delta_n - 2\eta_n, \eta_n, \eta_n\right) \quad \text{(A10)}$$

in terms of two variables δ_n and η_n, and obtain some bounds for these variables. Since we are pumping against bit-errors, η_n decreases with n

$$\eta_{n+1} < \eta_n < \cdots < \eta_0 = \alpha/2 < 1/6. \quad \text{(A11)}$$

The success probability for the $(n+1)$th step of pumping is

$$p_{n+1} = (1 - \alpha)(1 - 2\eta_n) + 2\alpha\eta_n \quad \text{(A12)}$$
$$< 1 - \alpha - 2\alpha\eta_n < 1 - \alpha, \quad \text{(A13)}$$

and on the other hand we have

$$p_{n+1} > (1 - \alpha)(1 - 2\eta_n) > \frac{2}{3}(1 - \alpha) \quad \text{(A14)}$$

where the second inequality follows from $\eta_n < 1/6$ [74]. The third element of the fidelity vector is

$$\eta_{n+1} = c_{n+1} = \frac{c_0 \eta_n + d_0 \eta_n}{p_{n+1}} = \frac{\alpha}{p_{n+1}}\eta_n. \quad \text{(A15)}$$

Since $p_{n+1} > \frac{2}{3}(1 - \alpha)$, we have

$$\eta_{n+1} < \frac{\alpha}{\frac{2}{3}(1-\alpha)}\eta_n < \cdots < \left(\frac{3\alpha}{2(1-\alpha)}\right)^{n+1} c_0 \quad \text{(A16)}$$
$$< \left(\frac{3}{4}\right)^{n+1} c_0$$

which indicates that η_n approaches zeros exponentially with respect to n. When n is large enough, η_n is negligible and $p_n \approx 1 - \alpha - O(\eta_n)$.

Similarly, we obtain the recursive relation for δ_n.

$$\frac{1}{2} + \delta_{n+1} = a_{n+1} = \frac{1}{p_{n+1}}(a_0 a_n + b_0 b_n) \quad \text{(A17)}$$
$$= \frac{1}{2}\frac{1 - \alpha - 2\alpha\eta_n}{p_{n+1}} + \frac{1 - 2\alpha}{p_{n+1}}\delta_n \quad \text{(A18)}$$
$$> \frac{1}{2} + \frac{1 - 2\alpha}{1 - \alpha}\delta_n. \quad \text{(A19)}$$

Thus

$$\delta_{n+1} > \frac{1 - 2\alpha}{1 - \alpha}\delta_n > \cdots > \left(\frac{1 - 2\alpha}{1 - \alpha}\right)^{n+1}\delta_0 \quad \text{(A20)}$$
$$= \left(\frac{1 - 2\alpha}{1 - \alpha}\right)^{n+1}\frac{1 - 3\alpha}{2}$$

and $\delta_0 = \frac{1-3\alpha}{2}$. We have therefore

$$\frac{\eta_n}{\delta_n} < \left(\frac{2 - 4\alpha}{3\alpha}\right)^n \frac{\alpha}{1 - 3\alpha} \quad \text{(A21)}$$

2. Second level pumping

After n_b steps of first level pumping, we have the fidelity vector

$$\vec{F}_0' = \left(\frac{1}{2} + \delta_0', \frac{1}{2} - \delta_0' - 2\eta_0', \eta_0', \eta_0'\right) \quad (A22)$$

where $\delta_0' = \delta_{n_b} > \left(\frac{1-2\alpha}{1-\alpha}\right)^{n+1}\delta_0$ and $\eta_0' = \eta_{n_b} < \left(\frac{3\alpha}{2(1-\alpha)}\right)^{n+1}c_0$.

The fidelity vector after n_p steps of pumping against phase errors is

$$\vec{F}_n' = \left(\frac{1}{2} + \delta_n', \frac{1}{2} - \delta_n' - \eta_n' - d_n, \eta_n', d_n\right), \quad (A23)$$

where $\delta_n' < 1/2$ and $\eta_n' < 1/2$.

The success probability for the $(n+1)$th step is

$$\begin{aligned}
p_{n+1}' &= \left(\frac{1}{2} + \delta_0' + \eta_0'\right)\left(\frac{1}{2} + \delta_n' + \eta_n'\right) \\
&\quad + \left(\frac{1}{2} - \delta_0' - \eta_0'\right)\left(\frac{1}{2} - \delta_n' - \eta_n'\right) \\
&= \frac{1}{2} + 2\left(\delta_0' + \eta_0'\right)\left(\delta_n' + \eta_n'\right) \quad (A24) \\
&> \frac{1}{2}
\end{aligned}$$

We now consider the elements of $\vec{F}_n'$. On one hand, the erroneous admixture of $|\Psi^-\rangle$ described by $\{d_n\}$ keep decreasing with n, since $|\Psi^-\rangle$ errors are also purified during the second level pumping. On the other hand, the erroneous admixture of $|\Psi^+\rangle$ described by $\{\eta_n'\}$ may increase with n, but it is upper-bounded by the following relation:

$$\begin{aligned}
\eta_{n+1}' &= \frac{a_0 c_n + c_0 a_n}{p_{n+1}'} \\
&< 2\left[\left(\frac{1}{2} + \delta_0'\right)\eta_n' + \eta_0'\left(\frac{1}{2} + \delta_n'\right)\right] \\
&< \left(1 + 2\delta_0'\right)\eta_n' + 2\eta_0' \quad (A25) \\
&< \left(1 + 2\delta_0'\right)^{n+1}\eta_0',
\end{aligned}$$

and one can also show the lower bound for η_{n+1}' by induction:

$$\eta_{n+1}' > a_0 c_n + c_0 a_n > \frac{1}{2}\left(\eta_n' + \eta_0'\right) > \eta_0'. \quad (A26)$$

However, δ_n' approaches $1/2$ much faster:

$$\begin{aligned}
\frac{1}{2} + \delta_{n+1}' &= a_{n+1} = \frac{a_0 a_n + c_0 c_n}{p_{n+1}'} \\
&= \frac{\left(\frac{1}{2} + \delta_0'\right)\left(\frac{1}{2} + \delta_n'\right) + \eta_0'\eta_n'}{\frac{1}{2} + 2\left(\delta_0' + \eta_0'\right)\left(\delta_n' + \eta_n'\right)} \quad (A27) \\
&= \frac{1}{2} + \frac{\delta_n' + \delta_0' - 2\left(\delta_0'\eta_n' + \delta_n'\eta_0' + \eta_0'\eta_n'\right)}{1 + 4\delta_0'\delta_n' + 4\left(\delta_0'\eta_n' + \delta_n'\eta_0' + \eta_0'\eta_n'\right)}
\end{aligned}$$

If $\frac{1}{2} > \delta_n' > \delta_0' > \eta_n' > \eta_0'$, we have

$$\delta_0'\eta_n' + \delta_n'\eta_0' + \eta_0'\eta_n' < \delta_0'\eta_n' + \eta_0' < \delta_0'\zeta, \quad (A28)$$

where we in the second inequality introduced a number ζ such that $\delta_0'\eta_n' + \eta_0' < \delta_0'\zeta$. In the next subsection, we will show that we may choose ζ small (i.e., $\zeta = 2\varepsilon$) such that

$$\begin{aligned}
\delta_{n+1}' &> \frac{\delta_n' + \delta_0'\left(1 - 2\zeta\right)}{1 + 4\delta_0'\delta_n'\left(1 + \zeta\right)} \quad (A29) \\
\lambda_{n+1}' &> \frac{\lambda_n' + \lambda_0'\left(1 - 2\zeta\right)}{1 + 4\lambda_0'\left(1 - 2\zeta\right)\lambda_n'} \\
&> \lambda_n' + \lambda_0'\left(1 - 2\zeta\right) - 2\lambda_0'\left(1 - 2\zeta\right)\lambda_n' \quad (A30)
\end{aligned}$$

where $\lambda_n' = \sqrt{\frac{1+\zeta}{1-2\zeta}}\delta_n' \approx \left(1 + 3\zeta/2\right)\delta_n'$ for $n = 0, 1, \cdots$, and the third equality uses $\frac{x+y}{1+4xy} > x + y - 2xy$ for $0 < x, y < 1/2$. Finally, we have

$$\begin{aligned}
1/2 - \lambda_{n+1}' &< \left(1 - 2\lambda_0'\left(1 - 2\zeta\right)\right)\left(1/2 - \lambda_n'\right) \\
&< \left(1 - 2\lambda_0'\left(1 - 2\zeta\right)\right)^{n+1}\left(1/2 - \lambda_0'\right) \quad (A31)
\end{aligned}$$

So far, we have introduced inequalities to bound elements of fidelity vectors at different stages of pumping. In the next subsection, we will use these constraints to show that we are able to achieve fidelity arbitrarily close to unity by carefully choosing the numbers of pumping steps (n_b, n_p) for the bit-phase two-level pumping scheme.

3. $\varepsilon - N$ argument

For $\forall \varepsilon > 0$ and $\alpha < 2/7$, we may choose n_b

$$\boxed{n_b \geq \max\left\{\frac{\ln\varepsilon - \ln\frac{\alpha}{1-3\alpha}}{\ln\frac{3\alpha}{2-4\alpha}}, \frac{3\ln\varepsilon - \ln\frac{\alpha}{2}}{\ln\frac{3\alpha}{2(1-\alpha)}}\right\},} \quad (A32)$$

such that

$$\eta_0' < \left(\frac{3\alpha}{2(1-\alpha)}\right)^{n_b}\frac{\alpha}{2} < \varepsilon^3 \quad (A33)$$

$$\frac{\eta_0'}{\delta_0'} = \frac{\eta_{n_b}}{\delta_{n_b}} < \left(\frac{3\alpha}{2-4\alpha}\right)^{n_b}\frac{\alpha}{1-3\alpha} < \varepsilon \quad (A34)$$

and

$$\delta_0' > \left(\frac{1-2\alpha}{1-\alpha}\right)^{n_b}\frac{1-3\alpha}{2} \quad (A35)$$

Then we choose for n_p

$$\boxed{\frac{2\ln\varepsilon}{\ln\left(1-2\delta_0'\right)} > n_p > \frac{\ln\varepsilon}{\ln\left(1-2\lambda_0'\left(1-2\zeta\right)\right)}.} \quad (A36)$$

Such n_p always exists, for $\frac{2\ln\varepsilon}{\ln\left(1-2\delta_0'\right)} \approx \frac{2\ln\varepsilon}{-2\delta_0'} >$ $\frac{\ln\varepsilon}{-2\delta_0'\sqrt{(1+\zeta)(1-2\zeta)}} \approx \frac{\ln\varepsilon}{\ln\left(1-2\lambda_0'(1-2\zeta)\right)}$. Thus, we have

$$\eta_{n_p}' < (1+2\delta_0')^{n_p}\,\eta_0' < (1-2\delta_0')^{-n_p}\,\eta_0' < \varepsilon^{-2}\eta_0' < \delta_0' \tag{A37}$$

and

$$1/2 - \lambda_{n_p}' < \left(1-2\delta_0'\left(1-2\zeta\right)\right)^{n_p}\left(1/2-\lambda_0'\right) < \varepsilon/2 \tag{A38}$$

We now verify the two required relations which are underlined in the previous discussion. The relation $\frac{1}{2} > \delta_n' > \delta_0' > \eta_n' > \eta_0'$ is satisfied for all $0 < n \le n_p$. And the relation $\underline{\delta_0'\eta_{n_p}' + \eta_0' < \delta_0'\zeta}$ is also implied, since $\eta_{n_p}' + \eta_0'/\delta_0' < \varepsilon^{-2}\eta_0' + \varepsilon <\underline{2\varepsilon \equiv: \zeta}$.

Finally, the achievable fidelity for the above choice of n_b and n_p is

$$1 - F_{n_p,n_b} = 1/2 - \delta_{n_p}' = 1/2 - \left(\frac{1-2\zeta}{1+\zeta}\right)^{1/2}\lambda_{n_p}'$$

$$< 1/2 - \left(\frac{1-2\zeta}{1+\zeta}\right)^{1/2}\left(1/2-\varepsilon\right) < 4\varepsilon. \tag{A39}$$

The bit-phase two-level pumping thus allows us to approach $F = 1$ with arbitrary good precision.

APPENDIX B: MARKOV CHAIN MODEL FOR TWO-LEVEL PUMPING

Here we present the Markov chain model for two-level entanglement pumping.

The state transition diagram for two-level entanglement pumping is shown in Fig. 11. We assume that the required pumping steps are n_b and n_p for the two levels, respectively. Since two entangled pairs are stored, we need two labels to track the intermediate state for two-level entanglement pumping.

Here we use "0,0" to denote the initial state with no Bell pairs, "0, $j+1$" for the state with one purified pair surviving j steps of pumping at the first level, "$k+1, j+1$" for the state with one purified pair surviving k steps of pumping at the second level and one purified pair surviving j steps of pumping at the first level, "$k + 1, *$" for the state with one purified pair surviving k steps of pumping at the second level and one purified pair surviving n_b steps of pumping at the first level, and "$*, 0$" for the final state with one purified pair surviving n_p steps of pumping at the second level.

For the first level pumping, the (success) transition probability from state "k, j" to state "$k, j+1$" is q_j, while the (failure) transition probability from state "k, j" to state "$k, 0$" is $1 - q_j$, for $j = 0, 1, \cdots, n_b$. For the second level pumping, the (success) transition probability from state "$k, *$" to state "$k+1, 0$" is Q_k, while the (failure) transition probability from state "$k, *$" to state "$0, 0$" is $1 - Q_k$, for $k = 0, 1, \cdots, n_p$. Here the transition probabilities $\{q_j\}$ and $\{Q_k\}$ can be calculated according to the density matrices of the intermediate purified Bell pairs [36]. The final state is self-trapped, and goes back to itself with unit probability. Altogether there are $(n_b + 2)(n_p + 1) + 1$ states.

In order to fulfill the requirement that each transition attempt consumes one unpurified Bell pair, we need to contract the states of "$k, *$" and "$k+1, 0$" into one state, since this transition does not consume any unpurified Bell pair. After the contraction, there are $(n_b + 1)(n_p + 1)+1$ states remaining.

Therefore, we may use a (column) vector $\vec{P}$ with $(n_b + 1)(n_p + 1) + 1$ elements to characterize the probability distribution among all $(n_b + 1)(n_p + 1) + 1$ states. From the t-th attempt to the $(t+1)$th attempt, the probability vector evolves from $\vec{P}(t)$ to $\vec{P}(t + 1)$ according to the following rule

$$\vec{P}(t+1) = \mathbf{M}\,\vec{P}(t), \tag{B1}$$

and the transition matrix is

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_1 + q_{n_b}(1-Q_0)\mathbf{N}_1 & q_{n_b}(1-Q_1)\mathbf{N}_1 & \cdots & q_{n_b}(1-Q_2)\mathbf{N}_1 & 0 \\ q_{n_b}Q_0\mathbf{N}_1 & \mathbf{M}_1 & & & \\ & q_{n_b}Q_1\mathbf{N}_1 & \cdots & & \\ & & \cdots & \mathbf{M}_1 & \\ 0 & & & q_{n_b}Q_{n_p}\mathbf{N}_1 & 1 \end{pmatrix}, \tag{B2}$$

with sub-matrices:

$$\mathbf{M}_1 = \begin{pmatrix} 0 & 1-q_1 & \cdots & 1-q_{n_b} & 0 \\ 1 & 0 & & & \\ & q_1 & \cdots & & \\ & & \cdots & 0 & \\ & & & q_{n_b} & 0 \end{pmatrix}_{(n_b+1)\times(n_b+1)}, \tag{B3}$$

and

$$\mathbf{N}_1 = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ & & 0 & 0 \\ & \cdots & & \cdots \\ 0 & & & 0 \end{pmatrix}_{(n_b+1)\times(n_b+1)} . \tag{B4}$$

APPENDIX C: FAULT-TOLERANT PREPARATION OF 2^n-QUBIT GHZ STATE

We consider fault-tolerant preparation of a 2^n-qubit GHZ state with $n \geq 3$. We label 2^n registers by $0, 1, \cdots, 2^n - 1$. The GHZ state can be prepared in just 3 (clock cycles): in the first clock cycle, we perform PBM for register pairs $(2m, 2m + 1)$ with integer m; in the second clock cycle, we perform PBM for pairs $(2m, 2m + 2)$ and $(2m + 1, 2m + 3)$; in the last clock cycle, we perform PBM for pairs $(8m + 3, 8m + 5)$, $(8m + 4, 8m + 6)$, $(16m' + 7, 16m' + 9)$, $(16m' + 8, 16m' + 10)$, $\cdots$, $(2^{n-1} + 2^{n-2} - 1, 2^{n-1} + 2^{n-2} + 1)$, and $(2^{n-1} + 2^{n-2}, 2^{n-1} + 2^{n-2} + 2)$ in parallel.

In order to prepare the specific GHZ state $|00\cdots0\rangle +$ $|11\cdots1\rangle$, we still need to perform bit-flip operations for individual registers, which are determined by the measurement outcomes for all PBMs. Suppose the error probability for each PBM is p. Since the redundancy checks of the PBMs impose consistency requirements for measurement outcomes (error detection), the probability for undetected errors in measurement outcomes has been suppressed to $O(p^2)$ for each PBM. To the leading order of p, we only need to consider the phase errors from PBM that are not detected by the redundancy check. Thus, the total error probability scales as $2^{n-1}3p$, and the error probability for each register is only approximately $3p/2$. Therefore, we have demonstrated a PBM-based scheme to prepare the GHZ state fault-tolerantly, which is efficient in both time and physical-resources.

[1] H. Mabuchi and A. C. Doherty, Science **298**, 1372 (2002).

[2] A. Wallraff, D. I. Schuster, A. Blais, L. Frunzio, R. S. Huang, J. Majer, S. Kumar, S. M. Girvin, and R. J. Schoelkopf, Nature (London) **431**, 162 (2004).

[3] C. Simon and W. T. M. Irvine, Phys. Rev. Lett. **91**, 110405 (2003).

[4] L. M. Duan, B. B. Blinov, D. L. Moehring, and C. Monroe, Quantum Inf. Comput. **4**, 165 (2004).

[5] Y. L. Lim, S. D. Barrett, A. Beige, P. Kok, and L. C. Kwek, Phys. Rev. A **73**, 012304 (2006).

[6] M. D. Barrett, J. Chiaverini, T. Schaetz, J. Britton, W. M. Itano, J. D. Jost, E. Knill, D. Leibfried, R. Ozeri, and D. J. Wineland, Nature (London) **429**, 737 (2004).

[7] M. Riebe, H. Haffner, C. F. Roos, W. Hansel, J. Benhelm, G. P. T. Lancaster, T. W. Korber, C. Becher, F. Schmidt-Kaler, D. F. V. James, et al., Nature (London) **429**, 734 (2004).

[8] R. Reichle, D. Leibfried, E. Knill, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, Nature (London) **443**, 838 (2006).

[9] O. Mandel, M. Greiner, A. Widera, T. Rom, T. W. Hansch, and I. Bloch, Nature (London) **425**, 937 (2003).

[10] F. Jelezko, T. Gaebel, I. Popa, M. Domhan, A. Gruber, and J. Wrachtrup, Phys. Rev. Lett. **93**, 130501 (2004).

[11] L. Childress, M. V. G. Dutt, J. M. Taylor, A. S. Zibrov, F. Jelezko, J. Wrachtrup, P. R. Hemmer, and M. D. Lukin, Science **314**, 281 (2006).

[12] M. V. G. Dutt, L. Childress, L. Jiang, E. Togan, J. Maze, F. Jelezko, A. S. Zibrov, P. R. Hemmer, and M. D. Lukin, Science **316**, 1312 (2007).

[13] T. Yamamoto, Y. A. Pashkin, O. Astafiev, Y. Nakamura, and J. S. Tsai, Nature (London) **425**, 941 (2003).

[14] R. McDermott, R. W. Simmonds, M. Steffen, K. B. Cooper, K. Cicak, K. D. Osborn, S. Oh, D. P. Pappas, and J. M. Martinis, Science **307**, 1299 (2005).

[15] W. Dür and H. J. Briegel, Phys. Rev. Lett. **90**, 067901 (2003).

[16] D. K. L. Oi, S. J. Devitt, and L. C. L. Hollenberg, Phys. Rev. A **74**, 052313 (2006).

[17] R. Van Meter, W. J. Munro, K. Nemoto, and K. M. Itoh, e-print arXiv: **quant-ph/0607160** (2006).

[18] R. Van Meter, K. Nemoto, and W. J. Munro, IEEE Transactions on Computers **56**, 1643 (2007).

[19] D. Leibfried, B. DeMarco, V. Meyer, D. Lucas, M. Barrett, J. Britton, W. M. Itano, B. Jelenkovic, C. Langer, T. Rosenband, et al., Nature (London) **422**, 412 (2003).

[20] L. M. K. Vandersypen and I. L. Chuang, Rev. Mod. Phys. **76**, 1037 (2004).

[21] J. I. Cirac, P. Zoller, H. J. Kimble, and H. Mabuchi, Phys. Rev. Lett. **78**, 3221 (1997).

[22] K. M. Svore, B. M. Terhal, and D. P. DiVincenzo, Phys. Rev. A **72**, 022317 (2005).

[23] D. Bouwmeester, A. K. Ekert, and A. Zeilinger, *The physics of quantum information : quantum cryptography, quantum teleportation, quantum computation* (Springer, Berlin; New York, 2000).

[24] D. Gottesman and I. L. Chuang, Nature (London) **402**, 390 (1999).

[25] A. Sørensen and K. Mølmer, Phys. Rev. A **58**, 2745 (1998).

[26] X. Zhou, D. W. Leung, and I. L. Chuang, Phys. Rev. A **62**, 052316 (2000).

[27] J. Eisert, K. Jacobs, P. Papadopoulos, and M. B. Plenio, Phys. Rev. A **62**, 052317 (2000).

[28] H. J. Briegel and R. Raussendorf, Phys. Rev. Lett. **86**, 910 (2001).

[29] R. Raussendorf and H. J. Briegel, Phys. Rev. Lett. **86**, 5188 (2001).

[30] S. C. Benjamin, D. E. Browne, J. Fitzsimons, and J. J. L. Morton, New Journal of Physics **8**, 141 (2006).

[31] L. Childress, J. M. Taylor, A. S. Sørensen, and M. D. Lukin, Phys. Rev. A **72**, 052330 (2005).

[32] L. M. Duan and H. J. Kimble, Phys. Rev. Lett. **90**, 253601 (2003).

[33] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Phys. Rev. Lett. **76**, 722 (1996).

[34] D. Deutsch, A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu, and A. Sanpera, Phys. Rev. Lett. **77**, 2818 (1996).

[35] H. J. Briegel, W. Dür, J. I. Cirac, and P. Zoller, Phys. Rev. Lett. **81**, 5932 (1998).

[36] W. Dür, H. J. Briegel, J. I. Cirac, and P. Zoller, Phys. Rev. A **59**, 169 (1999).

[37] L. Jiang, J. M. Taylor, A. Sørensen, and M. D. Lukin, e-print arXiv: **quant-ph/0703029** (2007).

[38] C. Kim, C. Knoernschild, B. Liu, and J. Kim, Ieee Journal of Selected Topics in Quantum Electronics **13**, 322 (2007).

[39] B. B. Blinov, D. L. Moehring, L. M. Duan, and C. Monroe, Nature (London) **428**, 153 (2004).

[40] W. Rosenfeld, S. Berner, J. Volz, M. Weber, and H. Weinfurter, Phys. Rev. Lett. **98**, 050504 (2007).

[41] P. Maunz, D. L. Moehring, S. Olmschenk, K. C. Younge, D. N. Matsukevich, and C. Monroe, Nat Phys **3**, 538 (2007).

[42] D. L. Moehring, P. Maunz, S. Olmschenk, K. C. Younge, D. N. Matsukevich, L. M. Duan, and C. Monroe, Nature (London) **449**, 68 (2007).

[43] C. Langer, R. Ozeri, J. D. Jost, J. Chiaverini, B. DeMarco, A. Ben-Kish, R. B. Blakestad, J. Britton, D. B. Hume, W. M. Itano, et al., Phys. Rev. Lett. **95**, 060502 (2005).

[44] H. Haffner, F. Schmidt-Kaler, W. Hansel, C. F. Roos, T. Korber, M. Chwalla, M. Riebe, J. Benhelm, U. D. Rapol, C. Becher, et al., Appl. Phys. B **81**, 151 (2005).

[45] D. Leibfried, E. Knill, S. Seidelin, J. Britton, R. B. Blakestad, J. Chiaverini, D. B. Hume, W. M. Itano, J. D. Jost, C. Langer, et al., Nature (London) **438**, 639 (2005).

[46] H. Haffner, W. Hansel, C. F. Roos, J. Benhelm, D. Chekal kar, M. Chwalla, T. Korber, U. D. Rapol, M. Riebe, P. O. Schmidt, et al., Nature (London) **438**, 643 (2005).

[47] S. Tomljenovic-Hanic, M. J. Steel, C. M. de Sterke, and J. Salzman, Optics Express **14**, 3556 (2006).

[48] J. S. Hodges, J. C. Yang, C. Ramanathan, and D. G. Cory, e-print arXiv: **quant-ph/0707.2956** (2007).

[49] N. Khaneja, e-print arXiv: **0707.1572** (2007).

[50] P. Cappellaro, L. Jiang, N. Khaneja, and M. D. Lukin (in preparation).

[51] L. Jiang, M. V. G. Dutt, E. Togan, L. Childress, P. Cappellaro, J. M. Taylor, and M. D. Lukin, e-print arXiv: **0707.1341** (2007).

[52] M. H. Levitt, Progress in Nuclear Magnetic Resonance Spectroscopy **18**, 61 (1986).

[53] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbruggen, and S. J. Glaser, J. Magn. Res. **172**, 296 (2005).

[54] V. B. Braginsky, Y. I. Vorontsov, and K. S. Thorne, Science **209**, 547 (1980).

[55] S. D. Barrett and P. Kok, Phys. Rev. A **71**, 060310(R) (2005).

[56] D. B. Hume, T. Rosenband, and D. J. Wineland, Phys. Rev. Lett. **99**, 120502 (2007).

[57] W. Dür, Master thesis, Innsbruck (1998).

[58] S. P. Meyn and R. L. Tweedie, *Markov chains and stochastic stability* (Springer-Verlag, London; New York, 1993).

[59] E. T. Campbell, Phys. Rev. A **76**, 040302(R) (2007).

[60] P. Michler, A. Kiraz, C. Becher, W. V. Schoenfeld, P. M. Petroff, L. D. Zhang, E. Hu, and A. Imamoglu, Science **290**, 2282 (2000).

[61] E. M. Purcell, Physical Review **69**, 681 (1946).

[62] J. J. Garcia-Ripoll, P. Zoller, and J. I. Cirac, Phys. Rev. Lett. **91**, 157901 (2003).

[63] T. Steinmetz, Y. Colombe, D. Hunger, T. W. Hansch, A. Balocchi, R. J. Warburton, and J. Reichel, Applied Physics Letters **89**, 111110 (2006).

[64] M. Keller, B. Lange, K. Hayasaka, W. Lange, and H. Walther, Nature (London) **431**, 1075 (2004).

[65] E. Knill, Nature (London) **434**, 39 (2005).

[66] R. Raussendorf and J. Harrington, Phys. Rev. Lett. **98**, 190504 (2007).

[67] A. M. Steane, Phys. Rev. A **68**, 042322 (2003).

[68] D. P. DiVincenzo and P. Aliferis, Phys. Rev. Lett. **98**, 020501 (2007).

[69] M. A. Nielsen and I. Chuang, *Quantum computation and quantum information* (Cambridge University Press, Cambridge, U.K.; New York, 2000).

[70] L. Jiang, J. M. Taylor, N. Khaneja, and M. D. Lukin, Proc. Natl. Acad. Sci. U. S. A. **104**, 17291 (2007).

[71] The error model introduced in Eq. (2) can be regarded as the worst case error, since it in principle includes all possible errors that can happen to the system.

[72] The verification-based scheme requires $k \geq 0$ verifications for a good initialization and suppresses the initialization error to $\varepsilon_I \approx (p_I + p_M)^{k+1} + \frac{2k+1}{2} p_L$, a smaller error probability than that of the measurement-based scheme. However, the verification-based scheme may fail at some intermediate step and require reiteration of the verification from scratch, which is very similar to the one-level entanglement pumping, described by the Markov chain model in Sec. VII.

[74] For depolarizing or dephasing errors, the estimated t_C is approximately 1.3 times less by using the average infidelity.

[74] Note that the lower bound for p_{n+1} can be arbitrarily close to $1 - \alpha$ for sufficiently large n. Here, for clarity we use the loose bound $\frac{2}{3}(1 - \alpha) < p_{n+1}$, which later imposes a constraint $\alpha < 2/7$ (i.e. $F > 4/7 \approx 0.57$). In principle, by choosing tighter lower bounds for p_{n+1}, the following proof works for all $\alpha < 1/3$.