Hybrid magic state distillation for universal fault-tolerant quantum computation

Wenqiang Zheng,[1,*] Yafei Yu,[2,*] Jian Pan,[1] Jingfu Zhang,[3] Jun Li,[1] Zhaokai Li,[1] Dieter Suter,[3] Xianyi Zhou,[1,†] Xinhua Peng,[1,4,‡] and Jiangfeng Du[1,4,§]

[1]*Hefei National Laboratory for Physical Sciences at Microscale and Department of Modern Physics, University of Science and Technology of China, Hefei, Anhui 230026, China*

[2]*Laboratory of Nanophotonic Functional Materials and Devices, LQIT and SIPSE, South China Normal University, Guangzhou 510006, China*

[3]*Fakultät Physik, Technische Universität Dortmund, 44221 Dortmund, Germany*

[4]*Synergetic Innovation Center of Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei, Anhui 230026, China*

(Received 30 September 2014; revised manuscript received 23 December 2014; published 17 February 2015)

A set of stabilizer operations augmented by some special initial states known as "magic states" gives the possibility of universal fault-tolerant quantum computation. However, magic state preparation inevitably involves nonideal operations that introduce noise. The most common method to eliminate the noise is magic state distillation (MSD) by stabilizer operations. Here we propose a hybrid MSD protocol by connecting a four-qubit H-type MSD with a five-qubit T-type MSD in order to overcome the shortcomings of the previous MSD protocols. The hybrid MSD protocol further integrates distillable ranges of different existing MSD protocols and provides considerable improvement in qubit cost. Moreover, we experimentally demonstrate the four-qubit H-type MSD protocol using nuclear magnetic resonance technology, together with the previous five-qubit MSD experiment, to show the feasibility of the hybrid MSD protocol.

DOI: 10.1103/PhysRevA.91.022314 PACS number(s): 03.67.Lx, 03.67.Pp, 76.60.−k

I. INTRODUCTION

Decoherence and control errors are some of the major obstacles for the implementation of scalable quantum information processing. To overcome these obstacles, quantum fault-tolerance theory has been developed [1,2], in which the information is encoded in a subspace of a larger Hilbert space. The subspace is fixed by a subgroup of the Pauli group, consisting of some Hermitian tensor products of Pauli operators which are defined as the stabilizer of the subspace. Logical operations are *transversally* performed on the encoded information [3,4], with the aim of preventing the propagation of errors within the code block and further avoiding correlated errors in the course of quantum error correction. Unfortunately, only a limited set of operations, known as *stabilizer operations* (consisting of Clifford group unitaries [5], preparation of $|0\rangle$, and measurement in the computational basis), can be implemented in such a fault-tolerant manner, which cannot provide a universal quantum computation according to the Gottesman-Knill theorem [6,7]. This dilemma can be solved by introducing a nonstabilizer state (not eigenstates of Pauli operators) as an ancilla and then implementing a gate outside the Clifford group through gate teleportation [1].

Preparation of a nonstabilizer state would inevitably involve nonstabilizer operations [8,9], which are not fault tolerant and induce noise in the nonstabilizer state. The most common method for reducing noise is to distill noisy copies of these resource nonstabilizer states to an almost pure nonstabilizer state with only stabilizer operations [8,10,11]. The pure nonstabilizer states that can be prepared through distillation with only stabilizer operations are called *magic states*, and the fault-tolerant distillation for magic states is called *magic state distillation* (MSD). So far, two types of states have been found to be "magic," and they are called T-type and H-type magic states [10]. Consequently, MSD enables universal fault-tolerant quantum computing, and it also opens a framework to observe what kind of quantum states can provide universal fault-tolerant computational power [12–14]. However, MSD poses a big challenge to quantum computation because it consumes a majority of qubit resources in architectures [15]. Much effort has been devoted to developing economical methods to get pure magic states by concatenating MSD protocols for the same H-type magic state [16] and to building up effective instructions to compile magic states and the stabilizer operations for implementing non-Clifford operations [17]. In addition, various MSD protocols suffer from different shortcomings. For example, for five-qubit T-type MSD [10], there is a gap between the distillable unstabilizer states and stabilizer states for seven-qubit H-type MSD [11], the polarization of the output state increases only polynomially in the number of noisy copies at the range of high polarization, and for four-qubit H-type MSD [18], it cannot yield nearly pure magic states.

Here we propose a MSD protocol by hybridizing one H-type MSD protocol [18] with one T-type MSD protocol [10]. This hybrid protocol not only overcomes the shortcomings in the previous MSD protocols mentioned above but also has two additional advantages: the integration of distillable ranges of previous individual MSD protocols and a great reduction in qubit resource consumption. Moreover, up to now, the only experiment on the five-qubit T-type MSD was implemented in a NMR system [19]. Here we report an experimental demonstration for the four-qubit MSD by NMR to show the feasibility of this hybrid protocol.

*These authors contributed equally to this work.
†zhouxy@ustc.edu.cn
‡xhpeng@ustc.edu.cn
§djf@ustc.edu.cn

II. PRIOR MSD PROTOCOLS

An arbitrary one-qubit state can be represented as $\rho = (\mathbf{1} + x\sigma_x + y\sigma_y + z\sigma_z)/2$, where σ_x, σ_y, σ_z, and $\mathbf{1}$ denote the Pauli matrices and identity operator and $\vec{a} = (x, y, z)$ is a dimensionless vector of length $\leqslant 1$ that specifies the position of the state in the Bloch sphere. One the single-qubit state ρ with $|x| + |y| + |z| \leqslant 1$ (forming a stabilizer octahedron $\mathcal{O}_s$) cannot be distilled to nonstabilizer states with only Clifford operations [10]. Prior MSD protocols show that some states outside $\mathcal{O}_s$ can be distilled towards eigenstates of Clifford gates, such as the Hadamard H gate and the T gate [13,20]. These eigenstates are magic states: the H type has $\vec{a}_H = (0, \pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}}), (\pm\frac{1}{\sqrt{2}}, 0, \pm\frac{1}{\sqrt{2}}), (\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}}, 0)$, and the T type has $\vec{a}_T = (\pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}})$. Without loss of generality, here we focus on two of them:

$$\begin{aligned} H \text{ type: } & \rho_{\mathrm{H}} = [\mathbf{1} + (\sigma_x + \sigma_z)/\sqrt{2}]/2, \\ T \text{ type: } & \rho_{\mathrm{T}} = [\mathbf{1} + (\sigma_x + \sigma_y + \sigma_z)/\sqrt{3}]/2. \end{aligned} \tag{1}$$

The polarization of an arbitrary state ρ in the direction of the magic states (H direction or T direction) is defined as

$$\begin{aligned} H \text{ type: } & p_{\mathrm{H}}(\rho) = 2\mathrm{Tr}(\rho\rho_{\mathrm{H}}) - 1 = (x + z)/\sqrt{2}, \\ T \text{ type: } & p_{\mathrm{T}}(\rho) = 2\mathrm{Tr}(\rho\rho_{\mathrm{T}}) - 1 = (x + y + z)/\sqrt{3}, \end{aligned} \tag{2}$$

which quantifies how close state ρ is to the magic states. Given the resource of these pure magic states, one can implement gates outside the Clifford group (e.g., the $\pi/12$ phase gate and the $\pi/8$ phase gate) to enable universal quantum computation [10].

Bravyi and Kitaev proposed a T-type MSD protocol based on the five-qubit error-correcting code [10]. Provided noisy copies with $p_{\mathrm{T}}(\rho_{\mathrm{in}}) > \sqrt{3/7} \approx 0.655$, it is possible to obtain the output with $p_{\mathrm{T}}(\rho_{\mathrm{out}}) \to 1$ by iteration. We denote the distillable range of this T-type MSD protocol as $\mathcal{A}_{\mathrm{T}}$, represented by the gray and orange regions (i.e., 1 and 2 regions) in Fig. 1. A gap exists between the region $\mathcal{A}_{\mathrm{T}}$ and the stabilizer octahedron $\mathcal{O}_{\mathrm{s}}$. In contrast, Reichardt proposed an H-type MSD protocol based on the seven-qubit Steane code [11,18]. It is possible to obtain the output with $p_{\mathrm{H}}(\rho_{\mathrm{out}}) \to 1$ using this protocol, provided noisy copies with $p_{\mathrm{H}}(\rho_{\mathrm{in}}) > 1/\sqrt{2} \approx 0.707$. We denote this distillable range of

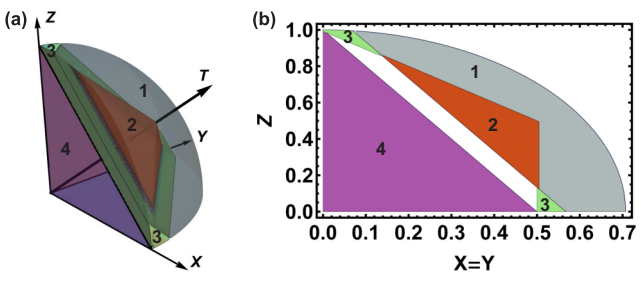

FIG. 1. (Color online) Distillable ranges of the five-qubit T-type MSD protocol ($\mathcal{A}_{\mathrm{T}}$, denoted as the gray and orange regions, i.e., 1 and 2 regions) and the seven-qubit H-type MSD protocol ($\mathcal{A}_{\mathrm{H}}$, denoted as the green and gray regions, i.e., 3 and 1 regions). The purple region, i.e., 4 region shows the interior of the stabilizer octahedron $\mathcal{O}_{\mathrm{s}}$. (a) One octant of the Bloch sphere. (b) A cross section of one octant of the Bloch sphere through the plane $x = y$.

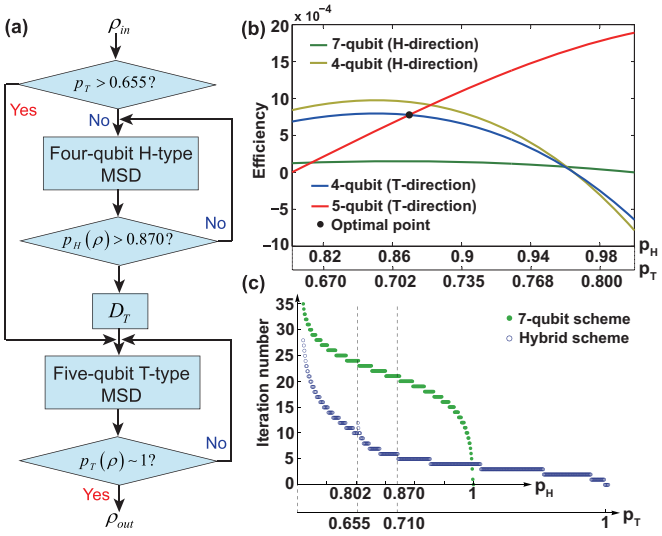

FIG. 2. (Color online) (a) Flowchart of the hybrid MSD protocol. $D_{\mathrm{T}}(\rho)=(\rho + T\rho T^\dagger + T^\dagger\rho T)/3$ is the T-projection operation. (b) The efficiency ν of the different MSD protocols along their target directions. The efficiency of four-qubit MSD protocol is also plotted along the T direction by transforming the polarization from the H direction to the T direction by D_{T} for comparison to the five-qubit MSD protocol. (c) The necessary iteration number to distill noisy states for an almost pure magic state (T type or H type) with target fidelity above 0.999. The second horizontal axis shows the corresponding p_{T} after the operation D_{T}.

the seven-qubit H-type MSD protocol as $\mathcal{A}_{\mathrm{H}}$, represented by the green and gray regions (i.e., 3 and 1 regions) in Fig. 1. The range $\mathcal{A}_{\mathrm{H}}$ is tight (no gap) in the directions crossing the octahedron edges, which means a transition from universal quantum computation to classical efficient simulation [13]. Alternatively, states can be distilled in the H direction by using a four-qubit Clifford circuit [18], with the price of a smaller distillable range $\mathcal{A}'_{\mathrm{H}} \subset \mathcal{A}_{\mathrm{H}}$, where the ultimate polarization $p_{\mathrm{H}}(\rho_{\mathrm{out}})$ is approximately equal to 0.964, not 1. From Fig. 1, we can see that the distillable ranges of the H-type and T-type MSD protocols do not overlap completely. It should be observed that $\mathcal{A}_{\mathrm{H}} - (\mathcal{A}_{\mathrm{T}} \cap \mathcal{A}_{\mathrm{H}}) = \mathcal{A}'_{\mathrm{H}} - (\mathcal{A}_{\mathrm{T}} \cap \mathcal{A}'_{\mathrm{H}})$.

In general, MSD protocols require measurements of the code's stabilizers. When all measurement outcomes are "zero," this round of distillation is a successful one. Assuming θ is the success probability, n/θ is the average qubit consumption in each iteration, where n is the code size of the protocol. To access the efficiency of different MSD protocols, we define a quantity $\nu = \Delta p\theta/n$ as the increased polarization per consumed qubit in each iteration. Here Δp is the increased polarization in the target direction after one iteration, i.e., the efficiency for each iteration. Figure 2(b) shows the efficiency of the different MSD protocols along their respective target directions.

III. HYBRID MSD PROTOCOL

Our hybrid MSD protocol is based on these two observations that $\mathcal{A}'_{\mathrm{H}} \cap \mathcal{A}_{\mathrm{T}} \neq \emptyset$ and the different efficiencies of these MSD protocols. A flowchart of the protocol is shown in Fig. 2(a). Having got an ensemble of many copies of

noisy state $\rho_{\rm in}$ with the polarization $p_{\rm H}(\rho_{\rm in}) > 0.707$, we first choose a few copies and measure their polarizations in the T direction to check whether $p_{\rm T}(\rho_{\rm in}) > 0.655$. If the answer is yes, these noisy copies are directly distilled by the five-qubit T-type MSD module. Otherwise, we send them into the four-qubit H-type MSD module for the higher polarization $p_{\rm H}(\rho_{\rm int})$. Comparing the efficiencies of the four-qubit and five-qubit protocols shown in Fig. 2(b), we can see that in the range $0.707 < p_H < 0.870$, the four-qubit protocol is more efficient, while beyond this range the five-qubit protocol has higher efficiency. Hence once $p_{\rm H}(\rho_{\rm int})$ reaches the optimal turning point $p_{\rm H}^{\rm op}(\rho_{\rm int}) = 0.870$, the intermediate states are then projected to the T direction by the twirling operation $D_{\rm T}$ [10]. $D_{\rm T}$ converts the polarization from the H direction to the T direction while reducing the polarization by a factor of $\sqrt{2/3}$. Next, these states are sent into the five-qubit T-type MSD module for further distillation. The hybrid protocol ultimately outputs almost pure T-type magic states. The first criterion $[p_{\rm T}(\rho_{\rm in}) > 0.655]$ is based on the numerical result that in the region $\mathcal{A}_{\rm T} \cap \mathcal{A}_{\rm H}$, the five-qubit protocol is less efficient for only 1% of the distillable states. We can see that both ranges $\mathcal{A}_{\rm T}$ and $\mathcal{A}_{\rm H}$ are distillable to T-type magic states using the hybrid protocol. One interesting conclusion is that the distillable range is also tight in directions crossing the octahedron edges for the T-type magic state, not just for H-type magic state.

Compared with the seven-qubit MSD protocol, the hybrid protocol can greatly reduce qubit cost for almost all of the distillable region. Figure 2(b) shows that the seven-qubit protocol performs with much lower efficiency in each round of distillation. In addition, the hybrid protocol has a considerable advantage in the necessary iteration number [Fig. 2(c)]. For the region $\mathcal{A}_{\rm H} - (\mathcal{A}_{\rm T} \cap \mathcal{A}_{\rm H})$ [i.e., $p_{\rm H}(\rho_{\rm in}) > 0.707$ and $p_{\rm T}(\rho_{\rm in}) < 0.655$], the qubit cost can be evaluated as $(4/\bar{\theta}_4)^{N_4}(5/\bar{\theta}_5)^{N_5}$ for the hybrid protocol and $(7/\bar{\theta}_7)^{N_7}$ for the seven-qubit protocol. Here $N_{4,5,7}$ are the iteration numbers of the 4-, 5-, and 7-qubit MSD schemes and $\bar{\theta}_{4,5,7}$ are the average success probabilities ($\bar{\theta}_4 = 0.244, \bar{\theta}_5 = 0.124$, and $\bar{\theta}_7 = 0.046$). The hybrid protocol can reduce the qubit cost by a roughly estimated factor of 10^{35} with respect to the seven-qubit MSD protocol (Appendix C), with a target polarization above 0.999 (this corresponds to implementing one non-Clifford operation with a theoretical fidelity of 0.9995 [21]). The same observation can be extracted for the major part of the region $\mathcal{A}_{\rm T} \cap \mathcal{A}_{\rm H}$ thanks to the property of five-qubit protocol that the increase of the out polarization is exponentially fast in the number of noisy copies when the polarization is high enough [10]. Just for a small region around the pure magic state $\rho_{\rm H}$, the seven-qubit protocol can be slightly more efficient. The exceptional region occupies about 0.57% of the distillable range.

IV. EXPERIMENTAL DEMONSTRATION OF THE FOUR-QUBIT H-TYPE MSD PROTOCOL

Now, we look in detail at the four-qubit H-type MSD scheme, whose quantum circuit is shown in Fig. 3(a): (i) first, prepare four copies of a noisy magic state $\rho_{\rm in}^{\otimes 4}$ as the input state, (ii) perform the parity checking in pairs, and (iii) if all measurements give a zero result, i.e., the three measured qubits are in the state $|000\rangle$, the protocol succeeds, and one applies the H-projection operation D_H to the qubit that has not been measured. The output state is $\rho_{\rm dis} \otimes |000\rangle\langle 000|$, with the success probability $\theta = (2 + 2p_0^2 + p_0^4)/16$, where $\rho_{\rm dis}$ has the output polarization:

$$p_{\rm H}(\rho_{\rm dis}) = \frac{6p_0^2 + p_0^4}{\sqrt{2}\{2 + 2p_0^2 + p_0^4\}}. \tag{3}$$

It gives $p_{\rm H}(\rho_{\rm dis}) > p_0$ when $0.7071 < p_0 < 0.9617$. Here $p_0 = p_{\rm H}(\rho_{\rm in})$ is the initial polarization of the input states.

We experimentally demonstrate the four-qubit distillation protocol. The physical system is iodotrifluroethylene (C_2F_3I) dissolved in d-chloroform. One ^{13}C nucleus and three ^{19}F nuclei are used as a four-qubit quantum information processor [22]. The natural Hamiltonian of the coupled spin system is $\mathcal{H} = \sum_i \mathcal{H}_z^i + \sum_{i<j} \mathcal{H}_c^{ij}$, where $\mathcal{H}_z^i = \pi \upsilon_i \sigma_z^i$ is the Zeeman term, υ_i is the Larmor frequency of spin i, and $\mathcal{H}_c^{ij} = (\pi/2) J_{ij} \sigma_z^i \sigma_z^j$ describes the interaction between spins i and j, with J_{ij} being the scalar coupling strength. Experiments

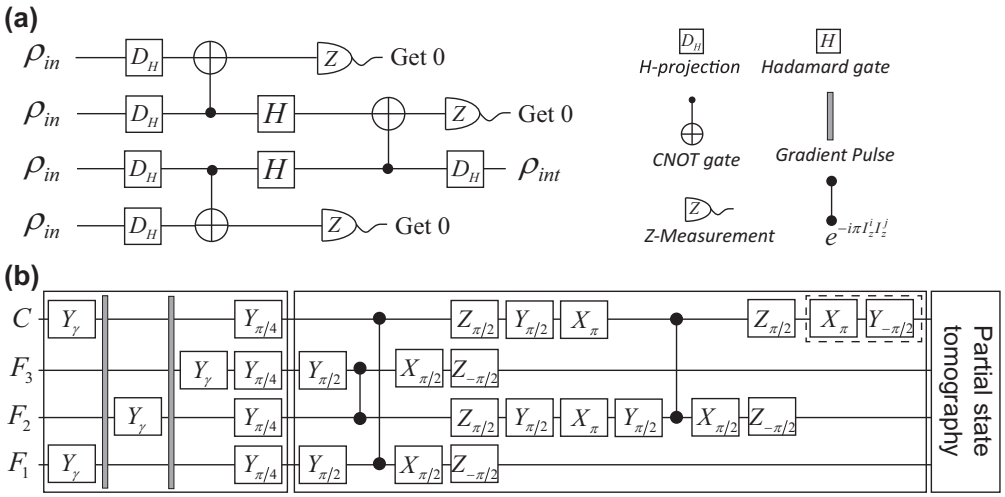

FIG. 3. (a) Quantum circuit and (b) the corresponding pulse sequence for the four-qubit H-type MSD protocol, where $X_\alpha = e^{-iaI_x}$, $Y_\alpha = e^{-iaI_y}$, and $Z_\alpha = e^{-iaI_z}$ denote single-qubit rotations. The H-projection operation $D_{\rm H}(\rho) = (\rho + H\rho H^\dagger)/2$ was realized by accumulating signals of two experiments: one with the Hadamard gate implemented by the pulses shown in the dashed box in (b) and the other without it.

were performed on a Bruker AV-500 spectrometer at room temperature. All of the relevant parameters along with the molecular structure are shown in Appendix G.

Figure 3(b) shows the pulse sequence of the experiment, corresponding to the quantum circuit in Fig. 3(a). We first initialized the system in a pseudopure state (PPS) [23] $\rho_{0000} = (1-\epsilon)\mathbf{1}/16 + \epsilon|0000\rangle\langle 0000|$ by using the line-selective method [22,24], where $\epsilon \approx 10^{-5}$ is the polarization. Instead of the first four H-projection operations in Fig. 3(a), four copies of noisy H-type magic states were prepared using the depolarization procedure shown in the first box of Fig. 3(b). A rotation by an angle γ around the y axis transforms σ_z to $\sigma_z \cos\gamma + \sigma_x \sin\gamma$. The following gradient field destroys the x component. By changing the rotation angle γ, we experimentally prepared five sets of noisy magic states, and each set has a different average polarization: $p_{\rm H}(\rho_{\rm in}^1) = 0.661, p_{\rm H}(\rho_{\rm in}^2) = 0.826, p_{\rm H}(\rho_{\rm in}^3) = 0.857, p_{\rm H}(\rho_{\rm in}^4) = 0.885, p_{\rm H}(\rho_{\rm in}^5) = 0.999$.

The three parity-check gates of Fig. 3(a) were implemented through the distillation procedure shown in the second box of Fig. 3(b). It consists of Clifford operations. At the output side, the ^{13}C nucleus carries the distilled magic state $\rho_{\rm dis}$. To avoid the error accumulation and to exhibit a near-perfect distillation step, we used one high-fidelity shaped pulse found with the gradient ascent pulse engineering (GRAPE) algorithm [25–27] to implement this sequence. The GRAPE pulses have durations of 16.8 ms, with theoretical fidelities above 0.996.

The distilled output state $\rho_{\rm dis}$ can be written as (the ^{13}C nucleus is labeled as qubit 1)

$$\rho_{\rm out}^{\rm expt} = \sum_{i=0}^{7} \theta_i \rho_i \otimes |i\rangle\langle i| + \sum_{i\neq j=0}^{7} \theta_{ij}\rho_{ij} \otimes |i\rangle\langle j|, \qquad \rho_i = \frac{1}{2}(\mathbf{1} + x_i\sigma_x + z_i\sigma_z), \tag{4}$$

where θ_i is the probability of the measurement outcome, corresponding to the resulting state $|i\rangle$ of the other three redundant qubits, and $|i\rangle = |000\rangle, |001\rangle, \ldots, |111\rangle$ for $i = 0,1,2,\ldots,7$. Measuring the outcome $|000\rangle$ indicates a successful purification. In NMR quantum information processing, since only ensemble measurements are available, we directly measure the expectation value of an observable, without projective measurements. In spite of this, we can obtain all θ_i, x_i, and z_i using partial quantum state tomography [28] to see the purification effect. Five readout operations are sufficient to determine all the desired parameters. First, by directly reading the signal of ^{13}C, we can obtain all $\theta_i x_i$; second, by reading the signal of ^{13}C after the application of a $\pi/2$ pulse, we can get all $\theta_i z_i$. The additional three readout operations consist of applying a $\pi/2$ pulse on F_1, F_2, and F_3, then reading the signals. The spectra of the four nuclei after applying a $\pi/2$ pulse are shown in Fig. 4(a). They are sufficient to determine all diagonal terms of the density matrix, which means we obtain all $\theta_i = m_{i+1} + m_{i+9}$, where m_i represents the ith diagonal term. Since this sample is unlabeled, we must transfer the polarizations of the ^{19}F spins to the ^{13}C spin using SWAP gates and then read the information of the ^{19}F spin through the ^{13}C spectrum [29]. The experimental results are shown in Fig. 4(b). The corresponding measured output polarizations are $p_{\rm H}(\rho_{\rm out}^1) = 0.640, p_{\rm H}(\rho_{\rm out}^2) = 0.838, p_{\rm H}(\rho_{\rm out}^3) = 0.867, p_{\rm H}(\rho_{\rm out}^4) = 0.894, p_{\rm H}(\rho_{\rm out}^5) = 0.979$. We see that the H polarization of the noisy magic states $\rho_{\rm in}^2, \rho_{\rm in}^3, \rho_{\rm in}^4$ have been

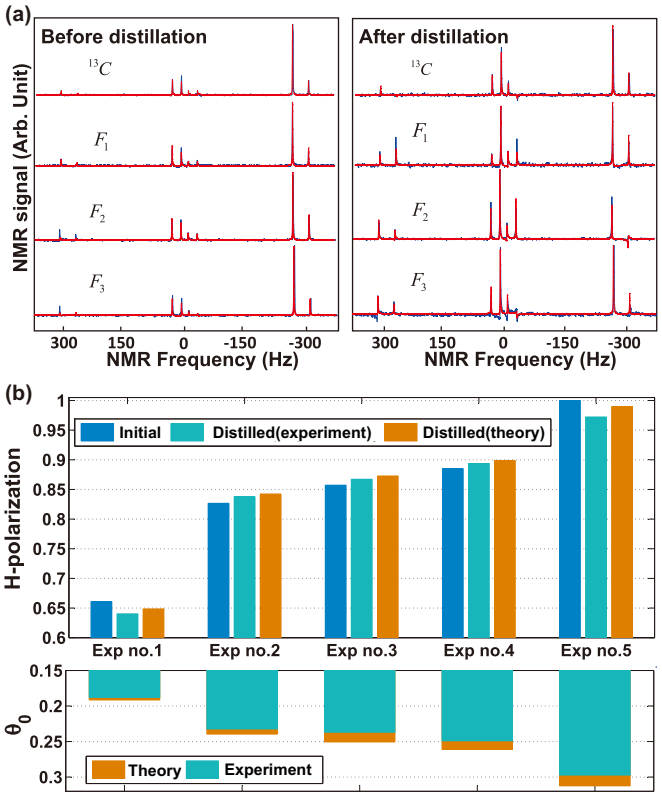

FIG. 4. (Color online) (a) Experimental spectra of the four nuclei before and after distillation for $p_{\rm H}(\rho_{\rm in}) = 0.826$. The experimentally measured and simulated spectra are shown as the blue and red curves, respectively. (b) Experimental results of the distillation. Top: polarizations along the H direction before and after distillation. Bottom: the success probability θ_0 of distillation vs the input polarization.

experimentally improved by the four-qubit MSD protocol because their input polarizations are in the distillable range. This result shows that one can enhance the H polarization of the quantum states in $\mathcal{A}_{\rm H}$ with the initial T polarization $p_{\rm T} \leqslant 0.655$ to $p_{\rm H} \geqslant 0.870$ using the four-qubit MSD and send them into the next step of the hybrid protocol, five-qubit MSD, to converge to the pure magic state $\rho_{\rm T}$.

The total experimental time of the distillation procedure including the readout procedure ranges from 25 to 40 ms. It is short compared to the transversal relaxation times of the nuclei (the minimum T_2^* of the four nuclei is about 250 ms), so the signal attenuation caused by the spin-spin relaxation effect is small. We numerically optimized all GRAPE pulses so that they are robust to 5% inhomogeneity of the rf field. By doing this, the influence of the rf field inhomogeneity is largely eliminated. For the four input copies, the largest polarization deviation of the individual spin from the average polarization is 0.029. A detailed numerical analysis of the robustness of the distillation algorithm to these differences in input polarizations is presented in Appendix F, which shows that the distillation algorithm is strongly robust to the imperfect copies of the initial state. The relative deviations between the experimental results and the theoretical expectation are 0.6%–1.2%. They mainly come from the imperfections of GRAPE pulses, experimental parameters, and data processing.

V. CONCLUSION

In conclusion, we presented a hybrid MSD protocol, which aims to take advantage of different MSD protocols. It further integrates all of the currently known distillable ranges and extends the T-type distillable range to the stabilizer octahedron edges. Moreover, the hybrid scheme has the remarkable advantage of saving qubit resources. The hybrid construction exhibits the ability to establish a unified framework for different MSD protocols; that is, all states in the integrated distillable range, in fact, can provide the quantum power for universal quantum computation along both the H and T directions. This might imply that the quantum power given to the distillable states may have some common source independent of the magic directions, e.g., contextuality [14] or negative Wigner representation [30]. We also experimentally demonstrate the four-qubit scheme by the NMR technology. The present experimental results, together with the previous NMR experiment for the five-qubit protocol, confirm the feasibility of the hybrid MSD scheme. It is expected that as more MSD methods are put forward, more and more distinguished combinations will come out according to the hybrid formalism.

ACKNOWLEDGMENTS

Y.Y. thanks R. Laflamme's group for discussions about magic state distillation. This work is supported by the National Key Basic Research Program of China (Grants No. 2013CB921800 and No. 2014CB848700), the National Natural Science Funds for Distinguished Young Scholar (Grant No. 11425523), the National Natural Science Foundation of China (Grants No. 11375167, No. 11227901, and No. 91021005), the Chinese Academy of Sciences, Strategic Priority Research Program (B) of CAS (Grant No. XDB01030400).

APPENDIX A: THE PERFORMANCE OF THE FOUR-QUBIT PROTOCOL IN THE ASYMPTOTIC REGIME

In the four-qubit protocol, the iterative function of the error probability $\varepsilon = \frac{1-p_H}{2}$ is

$$\varepsilon_{\text{out}} = f(\varepsilon) = 0.5 - \frac{6(1-2\varepsilon)^2 + (1-2\varepsilon)^4}{\sqrt{8}[2 + 2(1-2\varepsilon)^2 + (1-2\varepsilon)^4]}. \quad \text{(A1)}$$

Its first-order Taylor expansion near the polarization $p_H = 0.962$ (i.e., $\varepsilon = 0.019$) is

$$\begin{aligned}\varepsilon_{\text{out}} &= f(0.019) + f'(0.019)(\varepsilon - 0.019) + o[(\varepsilon - 0.019)^2] \\ &= 0.019 + 0.75(\varepsilon - 0.019) + o[(\varepsilon - 0.019^2], \end{aligned} \quad \text{(A2)}$$

where $\varepsilon^* = 0.019$ is the convergence value and the convergence rate is linear, which is different from the five-qubit protocol's quadratic convergence in the asymptotic regime. For the input ε_0, after the first iteration, $\varepsilon_1 \approx 0.019 + 0.75\,(\varepsilon_0 - 0.019)$. After the second iteration, $\varepsilon_2 \approx 0.019 + 0.75^2\,(\varepsilon_0 - 0.019)$. After k iterations, $\varepsilon_n \approx 0.019 + 0.75^k\,(\varepsilon_0 - 0.019)$. Near ε^*, we can approximate the success probability $\theta_s \approx \theta\,(\varepsilon^*) = 0.294$. The total number of initial noisy magic states $N = (4/\theta_s)^n \approx (13.6)^k$. So

$$\begin{aligned}\varepsilon_{\text{out}} &\approx 0.019 + 0.75^{\log_{13.6} N}(\varepsilon_0 - 0.019) \\ &= 0.019 + N^{(\frac{1}{\log_{0.75} 13.6})}(\varepsilon_0 - 0.019). \end{aligned} \quad \text{(A3)}$$

The error rate in the distilled magic states is reduced polynomially with respect to the number of noisy input magic states.

Nevertheless, in our hybrid scheme, before the polarization reaches the asymptotic regime of the four-qubit protocol, we switch to the five-qubit protocol, which gives an exponential decay of the error rate. The four-qubit protocol has the ability to reduce the qubit cost, which is mainly rooted in its lower qubit cost in every round of distillation, which results in an efficiency higher than that of the five-qubit protocol before the optimal turning point ($p_H = 0.870$).

APPENDIX B: THE INTEGRATION BETWEEN THE FOUR-QUBIT H-TYPE MSD AND THE FIVE-QUBIT T-TYPE MSD

In the hybrid magic state distillation (MSD) scheme, the noisy copies first enter into H-type MSD modules as four qubits in one group. With the increase in the iteration number, the modules output states closer to the H-type magic state. Certainly, their polarizations in the T direction p_T also increase. Once p_H is higher than the optimal turning point $p_H = 0.87$, which corresponds to $p_T = 0.71$ after operation D_T, the state enters into the T-type MSD module. Then, after

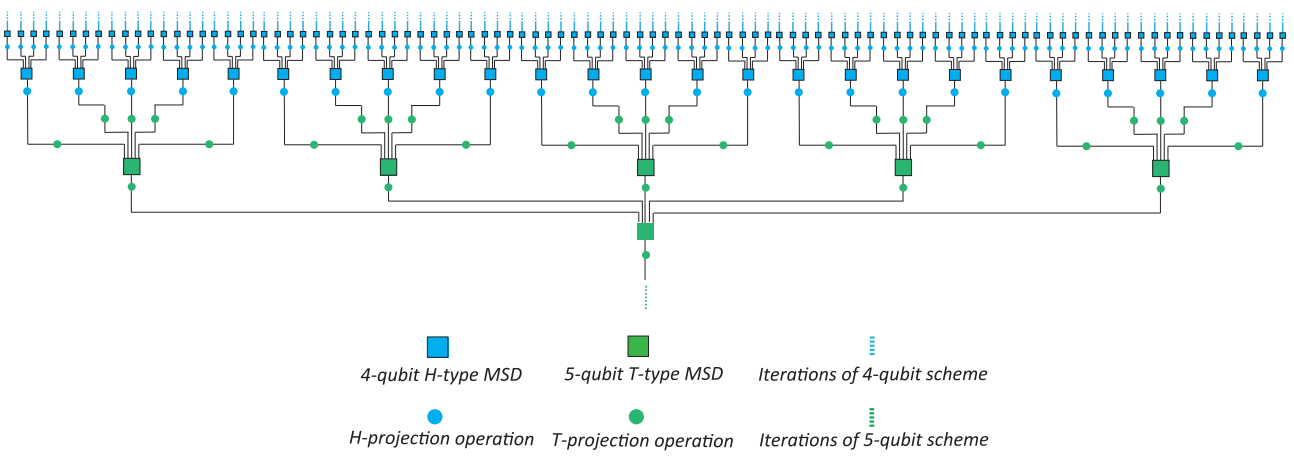

FIG. 5. (Color online) The integration between the four-qubit H-type MSD and the five-qubit T-type MSD.

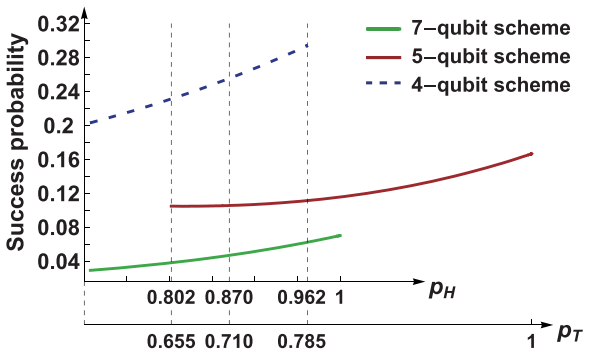

FIG. 6. (Color online) The success probability of the three individual MSD protocols.

several iterations of T-type MSD, we obtain states with $p_T \sim 1$. Figure 5 shows the integration between the two schemes.

APPENDIX C: THE QUBIT COST IN THE LOW-POLARIZATION RANGE $\mathcal{A}_{\mathrm{H}} - (\mathcal{A}_{\mathrm{T}} \cap \mathcal{A}_{\mathrm{H}})$

Figure 6 shows the success probability of the three individual MSD protocols vs the input polarization. From this, we can obtain the average success probability for one iteration for different protocols: $\bar{\theta}_4 = 0.244$, $\bar{\theta}_5 = 0.124$, and $\bar{\theta}_7 = 0.046$. For the input states in the region $\mathcal{A}_{\mathrm{H}} - (\mathcal{A}_{\mathrm{T}} \cap \mathcal{A}_{\mathrm{H}})$, the average necessary iteration number to achieve a T-type or H-type magic state with a target polarization above 0.999 is about 26 using the seven-qubit protocol. Then the qubit cost is evaluated as $(7/\bar{\theta}_7)^{26} \sim 10^{56}$. The hybrid protocol needs about 11 iterations of the four-qubit protocol and 5 iterations of the five-qubit protocol. The qubit cost is evaluated as $(4/\bar{\theta}_4)^{11}(5/\bar{\theta}_5)^5 \sim 10^{21}$. Hence the qubit cost is reduced by a factor of about 10^{35} using the hybrid protocol for the low-polarization states in $\mathcal{A}_{\mathrm{H}} - (\mathcal{A}_{\mathrm{T}} \cap \mathcal{A}_{\mathrm{H}})$.

As an example, Table I shows the numerical calculations of the performance of the hybrid protocol and the seven-qubit protocol when they are used to distill one noisy state whose polarization is $p_H = 0.780$ in the H direction, which is equivalent to $p_T = 0.636$ in the T direction. Although the target states of the two schemes are different ($p_H = 0.999$ in the seven-qubit case, $p_T = 0.999$ in the hybrid case), either of the target states can be used to implement one non-Clifford operation with a theoretical fidelity of 0.9995. We can see that the hybrid protocol requires fewer iterations and possesses a much higher successful probability than the seven-qubit protocol, which leads to the great savings in qubit cost.

APPENDIX D: THE EFFICIENCY OF THE WHOLE PROCEDURE

The quantity $\nu = \Delta p_{T(H)}\theta/n$ represents the increased polarization per consumed qubit in each iteration, i.e., the efficiency for each iteration. The efficiency of the whole MSD procedure with k iterations is defined as

$$V = \frac{\Delta p_1 + \Delta p_2 + \cdots + \Delta p_k}{n_1 n_2 \cdots n_k/\theta_1\theta_2\cdots\theta_k}. \tag{D1}$$

Both ν and V are determined by two parameters: the increment $\Delta p_{T(H)}$ gained from one iteration and the average qubit consumption n/θ in one iteration. Given a target polarization, the optimal turning point may be slightly different from the one calculated by the efficiency for each iteration ν. Setting

TABLE I. Comparison of the performance of the hybrid MSD protocol and the seven-qubit protocol when they are used to distill one noisy state in the H direction with $p_H = 0.78$. Since the target direction of the hybrid protocol is the T direction, the polarizations gained by the subprocedure of the four-qubit H-type protocol have been converted to the T direction, i.e., multiplied by the factor $\sqrt{2/3}$.

Itera-tion	Output polarization		Successful probability		Qubit cost	
No.	7-qubit scheme(p_h)	Hybrid scheme(p_t)	7-qubit scheme	Hybrid scheme	7-qubit scheme	Hybrid scheme
1	0.8001	0.6471	0.0359	0.2242	195	17.84
2	0.8226	0.6584	0.0380	0.2282	3592	312
3	0.8465	0.6706	0.0407	0.2327	10^6	5376
4	0.8703	0.6833	0.0437	0.2377	10^8	10^4
5	0.8928	0.6962	0.0470	0.2432	10^{11}	10^6
6	0.9129	0.7090	0.0504	0.2489	10^{13}	10^7
7	0.9301	0.7213	0.0536	0.2548	10^{15}	10^8
8	0.9445	0.7723	0.0566	0.0907	10^{17}	10^{10}
9	0.9562	0.8490	0.0613	0.0996	10^{19}	10^{12}
10	0.9656	0.9356	0.0632	0.1166	10^{21}	10^{13}
11	0.9730	0.9890	0.0646	0.1423	10^{23}	10^{15}
12	0.9789	0.9997	0.0658	0.1622	10^{25}	10^{16}
13	0.9835		0.0668		10^{27}	
14	0.9872		0.0676		10^{29}	
15	0.9900		0.0682		10^{31}	
16	0.9922		0.0686		10^{33}	
17	0.9939		0.0690		10^{35}	
18	0.9953		0.0693		10^{37}	
19	0.9963		0.0695		10^{39}	
20	0.9971		0.0697		10^{41}	
21	0.9978		0.0698		10^{43}	
22	0.9983		0.0699		10^{45}	
23	0.9987		0.0700		10^{47}	
24	0.9990		0.0701		10^{49}	

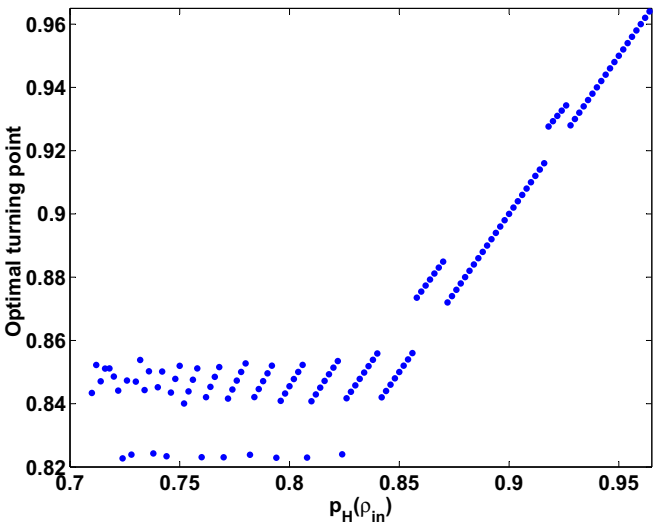

FIG. 7. (Color online) The optimal turning points for different input polarizations with the target polarization above 0.999. The optimal turning points gather around $p_H = 0.85$. If the input states' polarizations are higher than 0.86, it is better to distill them directly using the five-qubit T-type MSD protocol.

the target $P_T = 0.999$, for different initial polarizations, we numerically calculated the optimal turning points, which are shown in Fig. 7. We can see the optimal turning points gather around $p_H = 0.85$, which is slightly different from the hybrid protocol's turning point $p_H = 0.87$.

APPENDIX E: THE HYBRID PROTOCOL

In the hybrid protocol, the T-projection operation D_T is performed to the state when the polarization reaches the turning point $p_H = 0.87$. Then we get a state with the polarization of the T direction $p_T > 0.71$.

(1) If $p_T(\rho_{\text{in}}) > 0.71$, we directly distill the state with five-qubit protocol.

(2) If $0.655 < p_T(\rho_{\text{in}}) < 0.71$ and $p_H(\rho_{\text{in}}) < 0.707$, the state cannot benefit from the four-qubit protocol. We should distill it with the five-qubit protocol.

(3) If $p_T(\rho_{\text{in}}) < 0.655$ and $p_H(\rho_{\text{in}}) > 0.707$, the state cannot directly benefit from the five-qubit protocol. We should distill it using the present hybrid protocol, i.e., first distill it to $p_H(\rho_{\text{in}}) = 0.87$ with the four-qubit protocol and then distill it to the almost pure T-type magic state using the five-qubit protocol.

(4) If $0.655 < p_T(\rho_{\text{in}}) < 0.71$ and $0.707 < p_H(\rho_{\text{in}}) < 0.87$, for about 95% of the states in this range, it is better to directly distill them with the five-qubit protocol.

APPENDIX F: ROBUSTNESS OF THE FOUR-QUBIT H-TYPE MSD PROTOCOL

MSD protocols theoretically require some identical copies of the input noisy states, i.e., exactly the same polarization. However, it is impossible to prepare such ideal copies in the experiments. Therefore these experimental copies have slightly different polarizations. Table II shows the initial polarizations prepared in our experiment; the largest deviation from the average polarization for the individual qubit is around 0.029. It is important to analyze the robustness of the distillation protocol against the deviations of polarizations of input states.

For a polarization p_a, limited by the largest deviation δ, we randomly choose 400 input polarizations from the range $p_a - \delta$ to $p_a + \delta$. We divide them into 100 groups. The ith group has four input copies with polarizations p_i^1, p_i^2, p_i^3, and p_i^4. The average of the polarizations of the ith group is denoted as p_i^{in}. We calculate the output polarization p_i^{out} of the four-qubit MSD protocol for the ith group. The quantity $p_i^{\text{out}} - p_i^{\text{in}}$ represents the distillation effect of the ith group. We use the average distillation effect of these 100 groups [i.e., $\sum_{i=1}^{100}(p_i^{\text{out}} - p_i^{\text{in}})/100$] to evaluate the distillation effect of the protocol at the point (p_a, δ). Figure 8(a) shows the average distillation effect influenced by p_a and δ. It shows that this distillation protocol is strongly robust against the deviations of polarization between the input qubits. Figure 8(b) shows the output polarization corresponding to a Gaussian distribution of the input polarization. We can see that the output distributions have a larger average and smaller variance compared to the input ones.

TABLE II. The experimental polarizations of the prepared initial states.

nuclear spin	H-polarization of Initial state	Average polarization of Initial state
C	0.6436	0.6607
F1	0.6602	
F2	0.6719	
F3	0.6667	

nuclear spin	H-polarization of Initial state	Average polarization of Initial state
C	0.8135	0.8266
F1	0.8330	
F2	0.8444	
F3	0.8152	

nuclear spin	H-polarization of Initial state	Average polarization of Initial state
C	0.8476	0.8572
F1	0.8538	
F2	0.8591	
F3	0.8664	

nuclear spin	H-polarization of Initial state	Average polarization of Initial state
C	0.8603	0.8849
F1	0.9016	
F2	0.9138	
F3	0.8632	

nuclear spin	H-polarization of Initial state	Average polarization of Initial state
C	0.9927	0.9998
F1	1.0009	
F2	1.0077	
F3	0.9984	

APPENDIX G: SAMPLE INFORMATION

We used molecules of iodotrifluroethylene (C_2F_3I) dissolved in d-chloroform as the input system. As ^{13}C and ^{19}F are spin-$\frac{1}{2}$ nuclei, four qubits can be encoded using this sample for NMR quantum information processing. The natural Hamiltonian of the coupled spin system is $\mathcal{H} = \sum_i \mathcal{H}_i^z + \sum_{i<j} \mathcal{H}_{ij}^c$, where $\mathcal{H}_i^z = \pi \upsilon_i \sigma_z^i$ is the Zeeman Hamiltonian, υ_i is the Larmor frequency of spin i, and $\mathcal{H}_{ij}^c = (\pi/2) J_{ij} \sigma_z^i \sigma_z^j$ describes the interaction between spins i and j, with J_{ij} being the scalar coupling strength. All of the relevant parameters along with the molecular structure are shown in Fig. 9.

APPENDIX H: PPS PREPARATION

The system is originally in the thermal equilibrium state $\rho_{eq} = \mathbf{1}/16 + \sum_{i=1}^{4} \varepsilon_i I_z^i$, where $\varepsilon_i \sim 10^{-5}$ and $\boldsymbol{I} = (I_x, I_y, I_z)$ is the spin vector operator. For PPS $\rho_{0000} = (1 - \epsilon_0)\mathbf{1}/16 + \epsilon_0 |0000\rangle\langle 0000|$, the populations of all energy levels must be equalized except the first level. For this purpose, we numerically found an array $\{x_\alpha\}_{\alpha=1,\ldots,14}$, which determines a unitary operator

$$U_1 = \sum_\alpha \exp\left[-i x_\alpha I_x^{(\alpha+1,\alpha+2)}\right], \tag{H1}$$

where $I_x^{(\alpha+1,\alpha+2)}$ is the single quantum transition operator between levels $\alpha + 1$ and $\alpha + 2$. U_1 satisfies the following

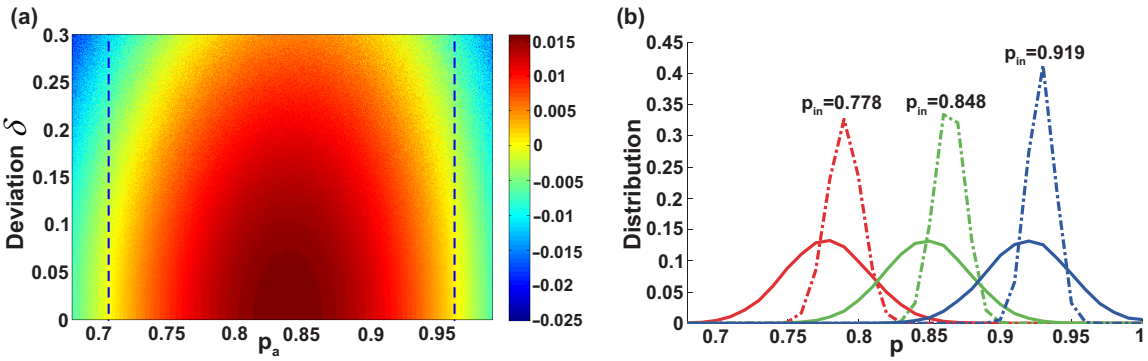

FIG. 8. (Color online) (a) Distillation effect influenced by various p_a and the biggest deviations δ. For example, the point (0.82, 0.08) in the δ and p_{in} plane means that we randomly choose the input polarizations from the range $0.82 - 0.08$ to $0.82 + 0.08$. The blue perpendicular lines represent the theoretical distillable boundary. (b) Distributions of output polarization corresponding to Gaussian distribution inputs. The solid curves represent the input polarization of Gaussian distributions with centers at 0.778, 0.848, and 0.919. The dashed curves show the distributions of the output.

requirement: $\text{diag}[U_1 \rho_{eq} U_1^\dagger] = \text{diag}[\rho_{0000}]$. That is, this unitary operator achieves saturation of the latter 15 energy levels, while the population of the first level remains unchanged. Then one gradient field pulse destroys all the coherences except for the homonuclear zero coherences of ^{19}F nuclei. The other specially designed unitary operator U_2 applies to the system and transforms these redundant zero coherences to others that can be eliminated with the gradient field pulse. Then, applying another gradient field pulse, we prepare the PPS ρ_{0000}. As U_1 is obtained with numerical search and U_2 is actually a combination of some controlled-NOT (CNOT) gates between two selected levels, it is hard to find a conventional pulse sequence to implement these two unitary operators. We engineered each operator as an individual shaped pulse using the gradient ascent pulse engineering (GRAPE) algorithm [25]. These two GRAPE pulses have a duration of around 25 ms with a theoretical fidelity above 0.994, and they are also designed to be robust against the rf inhomogeneity.

	C	F_1	F_2	F_3	T_2
C	0	-296.9	-274.7	39.5	7.9s
F_1		14409.7	66.6	50.5	6.8s
F_2			2636.6	-129.0	4.4s
F_3				-14404.2	4.8s

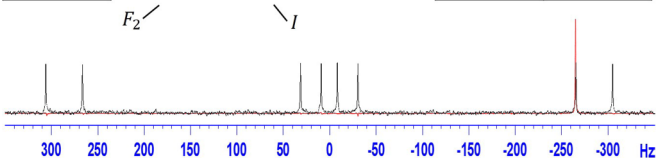

FIG. 9. (Color online) Characteristics of the four-qubit quantum register. The inset shows the structure, where the four qubits are labeled ^{13}C, F_1, F_2, F_3. The chemical shifts and scalar coupling constants (in Hz) are on and above the diagonal in the table, respectively. The last column shows the transversal relaxation time T_2 of each nucleus measured by CPMG sequences. Shown below are spectra of ^{13}C obtained by $\pi/2$ readout pulses when the system is prepared in the thermal equilibrium state (black) and PPS ρ_{0000} [red (gray)].

APPENDIX I: STATE TOMOGRAPHY

For input state tomography, the input state can be written as

$$\rho_0 = \frac{1}{2}[I + x_1(\sigma_x + \sigma_z)] \otimes \frac{1}{2}[I + x_2(\sigma_x + \sigma_z)] \otimes \cdots \frac{1}{2}[I + x_3(\sigma_x + \sigma_z)] \otimes \frac{1}{2}[I + x_4(\sigma_x + \sigma_z)] = \sum_{i=0}^{7}\sum_{j=0}^{7} \frac{1}{2}[I + x_1(\sigma_x + \sigma_z)] \otimes m_{ij} |i\rangle \langle j|,$$

where $\sum_{k=0}^{7} m_{kk} = 1$, $|i\,(j)\rangle = |000\rangle, |001\rangle, \ldots, |111\rangle$. So we can get x_1 by summing all signals of qubit 1. Similarly, we can get all x_i. The average of x_i is viewed as the polarization x_{in} of the input state.

For distilled state tomography, the state after distillation can be written as (assuming qubit 1 carries the distilled magic state)

$$\rho_{\text{out}}^{\text{exp}} = \sum_{i=0}^{7} \theta_i \rho_i \otimes |i\rangle \langle i| + \sum_{i \neq j=0}^{7} \theta_{ij} \rho_{ij} \otimes |i\rangle \langle j|,$$

$$\rho_i = \frac{1}{2}(I + x_i \sigma_x + z_i \sigma_z),$$

where θ_i is the probability of the measurement outcome, corresponding to the resulting state $|i\rangle$ of the other three redundant qubits, and $|i\rangle = |000\rangle, |001\rangle, \ldots, |111\rangle$ for $i = 0,1,2,\ldots,7$. Measuring outcome $|000\rangle$ indicates a successful purification. We can determine all the desired parameters using the following steps:

(1) *Read out on each qubit after the application of a $\pi/2$ pulse.* With this step, we can get all diagonal elements $\{m_i\}_{i=1,\ldots,16}$ of $\rho_{\text{out}}^{\text{exp}}$. For example, after operating a $\pi/2$ pulse on qubit 1, the intensity of the spectral line, which corresponds to the transition $|0000\rangle$ to $|1000\rangle$, is proportional

to the difference between the corresponding populations, i.e., $m_1 - m_9$. Then we get all $\theta_i = m_i + m_{i+8}$.

(2) *Read out on qubit 1 and read out on qubit 1 after the application of a* $\pi/2$ *pulse*. These two readout operations are sufficient to measure all $\theta_i \rho_i$.

Table III shows the experimental results of the four-qubit H-type MSD. The middle three inputs are in the distillable range. We observe higher polarizations corresponding to these three inputs.

TABLE III. Experimental results of the four-qubit H-type MSD.

Initial polarization	Polarization after distillation	Success probability
0.661	0.640	0.188
0.827	0.838	0.232
0.857	0.867	0.237
0.884	0.894	0.249
0.999(8)	0.979	0.297

[1] E. Knill, R. Laflamme, and W. H. Zurek, Science **279**, 342 (1998).

[2] P. Shor, in *Procedings of the 37th Annual Symposium on Foundations of Computer Science, Los Alamitos, CA, 1996* (IEEE Computer Society Press, New York, 1996), pp. 56–65.

[3] A. M. Steane, Fortschr. Phys. **46**, 443 (1998).

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

[5] D. Gottesman, Phys. Rev. A **57**, 127 (1998).

[6] D. Gottesman, Ph.D. thesis, Caltech, Pasadena, 1997.

[7] S. Aaronson and D. Gottesman, Phys. Rev. A **70**, 052328 (2004).

[8] S. Bravyi, Phys. Rev. A **73**, 042313 (2006).

[9] M. Howard and J. Vala, Phys. Rev. A **85**, 022304 (2012).

[10] S. Bravyi and A. Kitaev, Phys. Rev. A **71**, 022316 (2005).

[11] B. W. Reichardt, Quantum Inf. Process. **4**, 251 (2005).

[12] E. T. Campbell and D. E. Browne, Phys. Rev. Lett. **104**, 030503 (2010).

[13] W. van Dam and M. Howard, Phys. Rev. Lett. **103**, 170504 (2009).

[14] M. Howard, J. J. Wallman, V. Veitch, and J. Emerson, Nature (London) **510**, 351 (2014).

[15] N. C. Jones, R. Van Meter, A. G. Fowler, P. L. McMahon, J. Kim, T. D. Ladd, and Y. Yamamoto, Phys. Rev. X **2**, 031007 (2012).

[16] C. Jones, Phys. Rev. A **87**, 042305 (2013); S. Bravyi and J. Haah, *ibid.* **86**, 052329 (2012); A. M. Meier, B. Eastin, and E. Knill, Quantum Inf. Comput. **13**, 195 (2013).

[17] A. J. Landahl and C. Cesare, arXiv:1302.3240; B. Eastin, Phys. Rev. A **87**, 032321 (2013); E. T. Campbell, *ibid.* **83**, 032317 (2011).

[18] B. Reichardt, Quantum Inf. Comput. **9**, 1030 (2009).

[19] A. M. Souza, J. Zhang, C. A. Ryan, and R. Laflamme, Nat. Commun. **2**, 169 (2011).

[20] W. van Dam and M. Howard, Phys. Rev. A **83**, 032310 (2011).

[21] E. M. Fortunato, M. A. Pravia, N. Boulant, G. Teklemariam, T. F. Havel, and D. G. Cory, J. Chem. Phys. **116**, 7599 (2002).

[22] X. Peng, Z. Luo, W. Zheng, S. Kou, D. Suter, and J. Du, Phys. Rev. Lett. **113**, 080404 (2014).

[23] N. A. Gershenfeld and I. L. Chuang, Science **275**, 350 (1997).

[24] X. Peng, X. Zhu, X. Fang, M. Feng, K. Gao, X. Yang, and M. Liu, Chem. Phys. Lett. **340**, 509 (2001).

[25] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, J. Magn. Reson. **172**, 296 (2005).

[26] D. Lu, N. Xu, R. Xu, H. Chen, J. Gong, X. Peng, and J. Du, Phys. Rev. Lett. **107**, 020501 (2011).

[27] N. Xu, J. Zhu, D. Lu, X. Zhou, X. Peng, and J. Du, Phys. Rev. Lett. **108**, 130501 (2012).

[28] J.-S. Lee, Phys. Lett. A **305**, 349 (2002).

[29] X. Peng, S. Wu, J. Li, D. Suter, and J. Du, Phys. Rev. Lett. **105**, 240405 (2010).

[30] V. Veitch, C. Ferrie, D. Gross, and J. Emerson, New J. Phys. **14**, 113011 (2012); V. Veitch, S. A. Hamed Mousavian, D. Gottesman, and J. Emerson, *ibid.* **16**, 013009 (2014).