Remote Entanglement of Superconducting Qubits via Solid-State Spin Quantum Memories

Hodaka Kurokawa[1,*] Moyuki Yamamoto[2] Yuhei Sekiguchi[1] and Hideo Kosaka[1,2,†]

[1]*Institute of Advanced Sciences, Yokohama National University, 79-5 Tokiwadai, Hodogaya, Yokohama 240-8501, Japan*

[2]*Department of Physics, Graduate School of Engineering Science, Yokohama National University, 79-5 Tokiwadai, Hodogaya, Yokohama 240-8501, Japan*

Quantum communication between remote superconducting systems is being studied intensively to increase the number of integrated superconducting qubits and to realize a distributed quantum computer. Since optical photons must be used for communication outside a dilution refrigerator, the direct conversion of microwave photons to optical photons has been widely investigated. However, the direct conversion approach suffers from added photon noise, heating due to a strong optical pump, and the requirement for large cooperativity. Instead, for quantum communication between superconducting qubits, we propose an entanglement distribution scheme using a solid-state spin quantum memory that works as an interface for both microwave and optical photons. The quantum memory enables quantum communication without significant heating inside the refrigerator, in contrast to schemes using high-power optical pumps. Moreover, introducing the quantum memory naturally makes it possible to herald entanglement and parallelization using multiple memories.

I. INTRODUCTION

A superconducting qubit is one of the most promising building blocks for a quantum computer [1]. Recent developments in superconducting qubits enabled the integration of several tens of qubits onto a single chip and well-controlled quantum operation on them [2]. However, for practical calculations, integration on the order of one million qubits is estimated to be required [3], which is an extremely challenging task considering the current technologies. Although quantum memories may alleviate this requirement by several orders of magnitude [4], tens of thousands of qubits are still needed. To mitigate the severe demands concerning the integration of qubits, distributed quantum computing is suggested as an alternative [5–7]. Similar to a distributed classical computer, a distributed quantum computer is realized by linking small quantum computers using classical and quantum communication channels.

To realize quantum communication between superconducting qubits in remote dilution refrigerators, an optical photon must be used for communication outside the dilution refrigerators. Thus, the direct conversion of microwave photons to optical photons has been intensively

investigated in the last decade [8,9]. However, in the direct conversion scheme, high conversion efficiency is generally not compatible with the low optical pump power necessary to avoid heating up inside the refrigerator and to avoid generating noise. Instead, schemes using entangled photons to generate entanglements between superconducting qubits have recently gained attention [10–12]. These schemes use entangled-photon pairs, which have already been widely adapted for entanglement distribution at telecom wavelengths [13–15]. For quantum communication with optical photons, the application of a quantum memory is another choice to distribute entanglement between remote nodes [16–20]. Thus, using an analogy with communication at telecom wavelengths, we propose a quantum memory as an interface for microwave and optical photons, which works as a quantum repeater to generate entanglements between superconducting qubits in different dilution refrigerators.

In this paper, we introduce a scheme to generate entanglement between superconducting qubits in different refrigerators via a quantum memory. To analyze the scheme's performance, we consider a specific situation: generating entanglement between remote superconducting qubits using a quantum interface [21] consisting of a color center in a diamond, a microwave resonator, and a photonic and phononic crystal cavity. Assuming the current state-of-the-art technologies, we estimate the entanglement

---

*kurokawa-hodaka-hm@ynu.ac.jp
†kosaka-hideo-yp@ynu.ac.jp

generation rate of the memory-based scheme. Since it is not trivial whether the memory-based scheme has any advantages over other schemes, we compare its performance with those of the direct conversion scheme and the entangled-photon scheme. We confirm that as long as the conversion efficiency of the direct conversion scheme does not exceed a certain threshold, it is beneficial to introduce the memory for the entanglement generation. Besides, though the entanglement generation rates of the memory-based and entangled-photon schemes are estimated to be on the same order of magnitude, heat generation inside the refrigerator could be reduced more in the memory-based scheme than in the entangled-photon scheme. Since decreasing heating in the refrigerator is crucial for scaling quantum devices, using a memory could not only deliver entanglement but also help the thermal design inside a refrigerator, contributing to the establishment of a distributed quantum computer.

II. ENTANGLEMENT GENERATION PROTOCOL BETWEEN REMOTE SUPERCONDUCTING QUBITS

This section overviews the entanglement generation protocol between superconducting systems in remote dilution refrigerators using a quantum interface based on the solid-state spin quantum memory. Figure 1(a) shows a schematic of a quantum interface, consisting of a color center with nuclear spin memories in a diamond and three resonators (microwave, mechanical, and optical), that is connected to an optical fiber at the room-temperature region and to a superconducting qubit via a waveguide. The device works as an interface for both microwave and optical photons. To obtain a single microwave (optical) photon coupling that is strong enough to transfer the quantum state efficiently, a microwave resonator, a phononic crystal cavity, and a photonic crystal cavity are
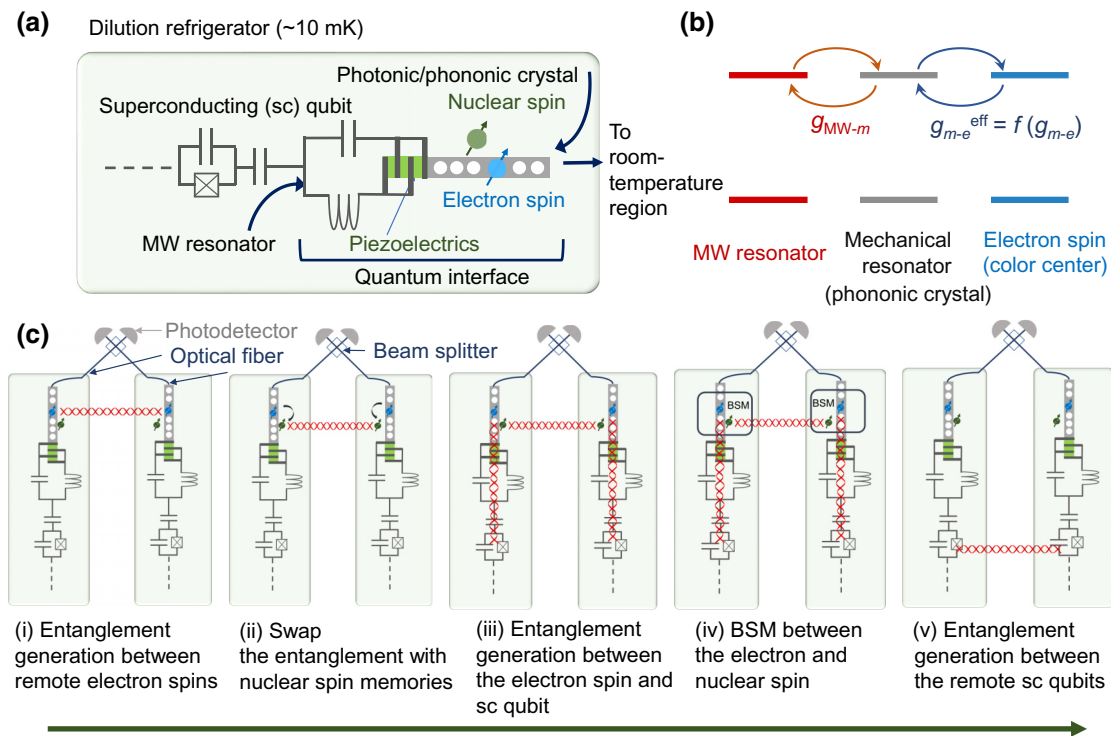

FIG. 1. (a) Schematic of a quantum interface inside a dilution refrigerator. The interface consists of a color-center electron spin, a nuclear spin, a microwave (MW) resonator, and a photonic and phononic crystal cavity. The microwave resonator is connected to a superconducting (sc) qubit via a waveguide. The photonic crystal is linked to an optical fiber, and an optical photon is transmitted to room-temperature region. (b) Schematic of the energy level of cascaded resonators (qubits) with the electromechanical interaction, $g_{MW\text{-}m}$, and the effective mechanical-spin coupling, $g_{m\text{-}e}^{\text{eff}}$. We use $g_{m\text{-}e}^{\text{eff}}$ instead of the bare mechanical-spin coupling, $g_{m\text{-}e}$, since the effective coupling between the mechanical mode and the electron spin depends on the choice of a color center and a detailed entanglement generation procedure. (c) Entanglement generation protocol between remote superconducting qubits. (i) Entanglement is generated between remote electron spins. (ii) The entanglement is swapped to nearby nuclear spins. (iii) Entanglement is generated between the electron spin and the superconducting qubit in each node. (iv) Bell-state measurement (BSM) is performed between the electron and the nuclear spin. (v) Entanglement is generated between remote superconducting qubits.

integrated into the interface. Figure 1(b) shows a simplified model of the coupled resonators and the electron spin. Since a color center with the orbital degree of freedom strongly interacts with the strain (phonon) due to the spin-orbit interaction [22,23], a microwave photon is converted to a microwave phonon via the piezoelectric effect. Then, the phonon confined inside the one-dimensional phononic crystal cavity interacts with a color center through the strain (mechanical)-electron spin coupling. The small mode volume ($< 0.1\lambda^3$) [24,25] of the one-dimensional (1D) phononic crystal is essential to realize sufficiently strong strain-spin coupling, where λ is the wavelength of elastic wave [26]. Here, we consider N-V^-, N-V^0, and Si-V^- as candidates that have high strain sensitivity in their ground or excited state. Similarly, coupling between a color center and an optical photon is enhanced by the confinement of the optical field inside the photonic crystal cavity.

Figure 1(c) shows a protocol for the remote entanglement generation between superconducting qubits using the quantum interface. We consider nodes inside dilution refrigerators separated by several meters. First, entanglement between remote electron spins is generated using a beam splitter that erases optical path information. Then, the entanglement is swapped to nearby nuclear spin memories. Next, entanglement between the superconducting qubit and the electron spin is generated in each node. Finally, Bell-state measurement (BSM) is performed between the electron and nuclear spins in each node, generating entanglement between remote superconducting qubits. In the following, we analyze and discuss each part of the protocol in terms of the entanglement generation rate.

III. REMOTE ENTANGLEMENT BETWEEN SOLID-STATE SPINS

Entanglement generation between remote color centers in diamonds has already been demonstrated using two-photon and single-photon protocols [19,27–29] (see Appendix A for details on single- and two-photon protocols). The two-photon protocol projects electron spins within different nodes (*A* and *B*) on one of the two Bell states,

$$\frac{1}{\sqrt{2}}(|0\rangle_e^A |1\rangle_e^B \pm |1\rangle_e^A |0\rangle_e^B), \quad (1)$$

where the sign depends on whether the same detector ($+$) or different detector ($-$) is clicked in the two-photon emission rounds [19]. $|0\rangle$ ($|1\rangle$) denotes the spin state in the computational basis, the subscript e indicates the electron spin, and the superscripts *A* and *B* correspond to the two nodes. Since the two-photon protocol is tolerant to the error caused by photon loss and does not require optical path

length stabilization, in the following we consider only the two-photon protocol.

Then, we consider a practical limit of the entanglement generation rate between color centers separated by several meters. The entanglement generation rate between remote color centers, $R_{e\text{-}e}$, is expressed as

$$R_{e\text{-}e} = \frac{1}{2} r_{e\text{-}e} (\eta_e^{\text{opt}})^2, \quad (2)$$

where η_e^{opt} is the detection efficiency of the optical photon emitted from a color center considering all losses during the transmission and $r_{e\text{-}e}$ is the entanglement generation attempt rate of the electron spins. The factor 2 in the denominator of Eq. (2) indicates that only two of the four Bell states can be heralded using a beam splitter. Note that $\eta_e^{\text{opt}} \sim 1$ is already feasible considering the current state-of-the-art technologies (see Appendix B).

For color centers separated by 10 m, the propagation time of an optical photon in an optical fiber is 25 ns from a color center to a photodetector placed at the midpoint of the refrigerators (5 m). Then, the ideal entanglement attempt rate will be 40 MHz. However, considering the current technology, the actual entanglement attempt rate is limited by the gate operation time for the emission of a coherent photon. In a two-photon protocol that used the N-V^- center, the typical operation time to generate a coherent photon is approximately 10 μs including the initialization time [19,27], corresponding to $r_{e\text{-}e} = 100$ kHz. Thus, assuming $\eta_e^{\text{opt}} = 0.9 - 1$, $R_{e\text{-}e}$ would be 80–100 kHz.

IV. ENTANGLEMENT GENERATION BETWEEN THE SUPERCONDUCTING QUBIT AND THE SPIN

In this section, we propose two protocols to generate entanglement between the superconducting qubit and the electron spin in a diamond. We then consider the entanglement generation rate based on these protocols. In the following, we consider that $|0\rangle_e$ is the electron spin ground state in the computational basis that is excited by the microwave time-bin qubit.

Figure 2(a) shows the first protocol, which uses an optical photon to herald the successful entanglement. The protocol can be applied to a color center that has high strain susceptibility in its ground or excited state. In Appendix D, we discuss the feasibility of the protocol considering N-V^- and N-V^0 as an example.

First, the superconducting qubit, which is prepared in the state, $(1/\sqrt{2})(|g\rangle_{\text{sc}} + |e\rangle_{\text{sc}})$, emits a coherent microwave time-bin qubit [30,31], resulting in the state,

$$\frac{1}{\sqrt{2}}(|g\rangle_{\text{sc}} |L\rangle_{\text{MW}} + |e\rangle_{\text{sc}} |E\rangle_{\text{MW}}). \quad (3)$$

Here, $|g\rangle_{\text{sc}}$ and $|e\rangle_{\text{sc}}$ denote the superconducting qubit basis states, and $|L\rangle_{\text{MW}}$ ($|E\rangle_{\text{MW}}$) is the microwave late
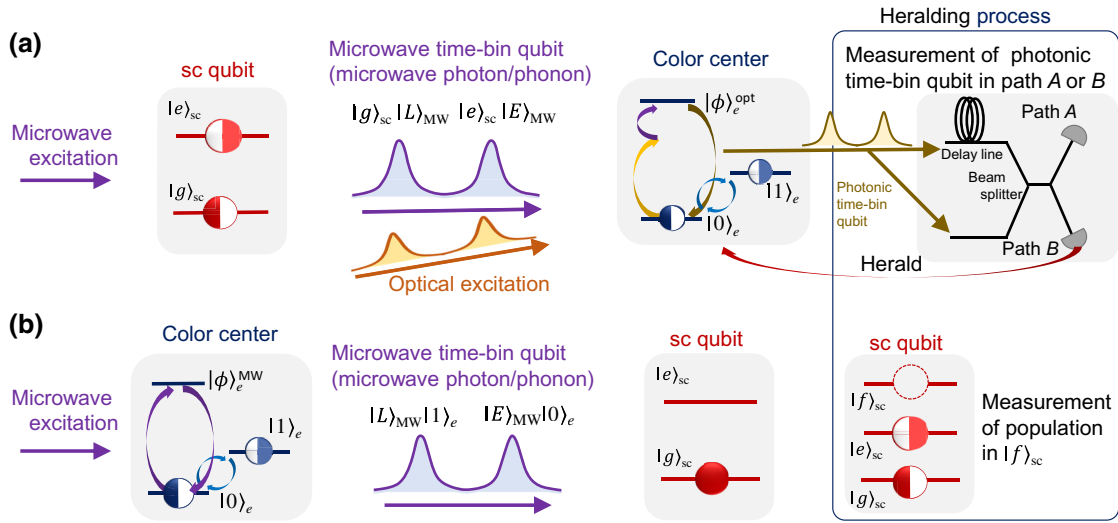
FIG. 2. Protocols to generate entanglement between a superconducting (sc) qubit and a color-center electron spin. (a) The superconducting qubit, which is in the superposition state, $(1/\sqrt{2})(|g\rangle_{\text{sc}} + |e\rangle_{\text{sc}})$, generates a microwave time-bin qubit. Here, for simplicity, the superconducting qubit is depicted as a two-level system, though another energy level and a coupled resonator are actually used to generate the microwave time-bin qubit [30]. The electron spin is prepared in the state $(1/\sqrt{2})(|0\rangle_e + |1\rangle_e)$. The electron spin, which is in $|0\rangle_e$, receives the microwave early time bin with a simultaneous optical excitation. The successful excitation to the optically excited state, $|\phi\rangle_e^{\text{opt}}$, results in the emission of an optical early time bin. The repetition of the procedure with bit flip between $|0\rangle_e$ and $|1\rangle_e$ generates the optical time-bin qubit, which is measured using a delay line and a beam splitter. The detection of the optical photon heralds the successful absorption of the microwave time-bin qubit. Applying a correction to the computational basis of the electron spin depending on the measurement outcome teleports the state of the microwave time-bin qubit to the electron spin, generating entanglement between the superconducting qubit and the electron spin. (b) The electron spin is prepared in the state $(1/\sqrt{2})(|0\rangle_e + |1\rangle_e)$. A microwave driving pulse is applied to excite $|0\rangle_e$ to $|\phi\rangle_e^{\text{MW}}$, which couples to a mechanical mode of the phononic crystal. $|\phi\rangle_e^{\text{MW}}$ emits the microwave early time bin via the phononic crystal and the microwave cavity. Repeating the emission with the bit flip between $|0\rangle_e$ and $|1\rangle_e$ generates a microwave time-bin qubit. The superconducting qubit receives the microwave time-bin qubit, generating entanglement between the electron spin and the superconducting qubit. The successful absorption of the microwave time-bin qubit can be heralded by confirming the absence of population in the second excited state, $|f\rangle_{\text{sc}}$, of the superconducting qubit [30].

(early) time-bin qubit. The detail of the emission of a time-bin qubit from a superconducting qubit is described in Appendix C. The electron spin is prepared in the state $(1/\sqrt{2})(|0\rangle_e + |1\rangle_e)$. Then, at the same time as the excitation of $|0\rangle_e$ by $|E\rangle_{\text{MW}}$, the electron spin is subjected to an optical pump (see Appendix D for details). Here, the detuning of the optical pump is equal to the frequency of $|E\rangle_{\text{MW}}$. As a result of the successful absorption of $|E\rangle_{\text{MW}}$ and the optical photons, the electron spin is excited to an optically excited state, $|\phi\rangle_e^{\text{opt}}$. Then, the electron spin emits a photonic early time bin, $|E\rangle_{\text{opt}}$. After the absorption of $|E\rangle_{\text{MW}}$, $|0\rangle_e$ and $|1\rangle_e$ are flipped. The procedure of photon emission is repeated for $|L\rangle_{\text{MW}}$. The electron spin emits a photonic late time bin, $|L\rangle_{\text{opt}}$, conditioned on the successful absorption of $|L\rangle_{\text{MW}}$. After that, $|0\rangle_e$ and $|1\rangle_e$ are flipped again. Consequently, the electron spin entangles with $|E\rangle_{\text{opt}}$ and $|L\rangle_{\text{opt}}$ as

$$\frac{1}{\sqrt{2}}(|g\rangle_{\text{sc}} |L\rangle_{\text{opt}} |1\rangle_e + |e\rangle_{\text{sc}} |E\rangle_{\text{opt}} |0\rangle_e). \quad (4)$$

Using a delay line and a beam splitter, photonic time-bin qubits are measured on the basis $(1/\sqrt{2})(|0\rangle_{\text{opt}}^A |1\rangle_{\text{opt}}^B \pm$

$|1\rangle_{\text{opt}}^A |0\rangle_{\text{opt}}^B)$. Detection of a single photon at the photodetectors heralds the successful interaction between the microwave time-bin qubit and the electron spin. The electron spin state is then corrected depending on the measurement outcome. Consequently, the state of the microwave time-bin qubit is teleported into the electron spin state with the heralding of the optical photon as

$$\frac{1}{\sqrt{2}}(|g\rangle_{\text{sc}} |1\rangle_e + |e\rangle_{\text{sc}} |0\rangle_e). \quad (5)$$

It is noteworthy that different protocols have been suggested to transfer the state of a superconducting qubit to the electron spin (N-V^-) in a diamond [32,33]. The scheme proposed in Ref. [32] uses a superconducting flux qubit coupled to a single N-V^-. To maximize the interaction between the superconducting qubit and N-V^-, N-V^- is placed in close proximity to the flux qubit (approximately 15 nm). However, such a configuration may hinder integration of an optical cavity, since the electrode that is too close to the N-V^- center contributes to the loss of the optical photon, degrading the quality factor of the optical cavity. Thus, the protocol is not compatible with our scheme,

which uses a photonic cavity to enhance spin-photon interaction. In another approach, a microwave photon emitted from a superconducting qubit is converted to an optical photon using an electro-optic effect and driving laser [33–37]. Using the converted optical photon and a driving laser, the ground state of N-V^- is controlled through the interaction between a Λ-type three-level system ($|m_s = \pm 1\rangle$ and $|A_2\rangle$ of N-V^-). Since their scheme already integrates whispering-gallery-mode photonic cavity into the device structure, it could be another choice to transfer the state of a superconducting qubit to the electron spin.

Figure 2(b) shows the second protocol for the entanglement generation, which uses a microwave time-bin qubit emitted from the electron spin. In this scheme, a color center having high strain susceptibility in the ground state is required (e.g., N-V^0 and Si-V^-). In the Appendix D 2, we consider the detail of the protocol, using N-V^0 as an example.

The initial state of the electron spin is prepared in $(1/\sqrt{2})(|0\rangle_e + |1\rangle_e)$. Then, $|0\rangle_e$ is excited to $|\phi\rangle_e^{\mathrm{MW}}$ by a microwave driving pulse supplied from an external signal generator. Based on the model shown in Fig. 1(b), the excited electron spin emits a microwave phonon to the phononic crystal cavity (see Appendix E for the detail). Repeating the emission with bit flip between $|0\rangle_e$ and $|1\rangle_e$ generates the state,

$$\frac{1}{\sqrt{2}}(|L\rangle_{\mathrm{MW}}|1\rangle_e + |E\rangle_{\mathrm{MW}}|0\rangle_e). \qquad (6)$$

The microwave time-bin qubit is transferred to the superconducting qubit via a waveguide coupled to the microwave cavity. The superconducting qubit, which is prepared in $|g\rangle_{\mathrm{sc}}$, absorbs the microwave time-bin qubit (see Appendix C for details), resulting in the entangled state,

$$\frac{1}{\sqrt{2}}(|g\rangle_{\mathrm{sc}}|1\rangle_e + |e\rangle_{\mathrm{sc}}|0\rangle_e). \qquad (7)$$

By measuring the population in the second excited state, $|f\rangle_{\mathrm{sc}}$, of the superconducting qubit, we can confirm whether the superconducting qubit successfully absorbs the microwave time-bin qubit or not, heralding the generation of entanglement [30].

A difference between these two protocols is whether the superconducting qubit or the electron spin is the emitter (receiver) of the microwave time-bin qubit. Efficient ($> 90\%$) microwave time-bin qubit generation and absorption by the superconducting qubit have been demonstrated [30,38], suggesting that the superconducting qubit can be used in both ways with similar efficiency. Thus, the efficiency of microwave emission or absorption by the color-center spins determines which protocols to choose. However, this efficiency depends on various parameters:

the choice of color center, coupling between the electron spin and the phonon, coupling between the microwave phonon and the microwave photon, and the quality factor of the phononic crystal and the microwave resonator. Thus, in the following, we treat the microwave photon generation or absorption efficiency of the electron spin as a parameter. Regarding the efficiency based on feasible device parameters, we calculate and discuss it in Appendix D. We should note that measurement of the electron spin on $|+\rangle_e = (1/\sqrt{2})(|0\rangle_e + |1\rangle_e)$ basis directly converts a microwave time-bin qubit to an optical time-bin qubit in the first protocol, enabling entanglement generation between the superconducting qubits by the interference of optical photons. However, here, we focus only on the protocol using the electron and nuclear spins as quantum memories.

On the basis of both schemes, we estimate an actual entanglement generation rate between the superconducting qubit and the electron spin. For a superconducting qubit coupled to a cavity, the generation or absorption of a microwave time-bin qubit takes several hundreds of nanoseconds to a few microseconds [30,38]. Here, we assume that it takes 1 μs to generate or absorb a microwave time-bin qubit for the superconducting qubit. Similarly, for the electron spin, approximately 1 μs is thought to be needed for the operation based on the simulation of the master equation (see Appendix D). Then, ignoring the propagation time in the waveguide between the superconducting qubit and the electron spin, it is possible to attempt entanglement generation at a rate of $r_{\mathrm{sc}\text{-}e} \sim 500$ kHz, where $r_{\mathrm{sc}\text{-}e}$ is the entanglement trial rate. Considering the success probability of microwave photon emission or absorption, the successful entanglement generation rate between the superconducting qubit and the electron spin, $R_{\mathrm{sc}\text{-}e}$, can be expressed as

$$R_{\mathrm{sc}\text{-}e} = r_{\mathrm{sc}\text{-}e}\eta_{\mathrm{sc}}^{\mathrm{MW}}\eta_e^{\mathrm{MW}}, \qquad (8)$$

where $\eta_{\mathrm{sc}}^{\mathrm{MW}}$ (η_e^{MW}) is the probability of successful microwave emission or absorption operation by the superconductor (spin). Here, we can confirm the generation of entanglement with the heralding techniques mentioned above. Figure 3 shows $R_{\mathrm{sc}\text{-}e}$ as a function of η_e^{MW}. Although we calculate $R_{\mathrm{sc}\text{-}e}$ for $\eta_e^{\mathrm{MW}} = 0.9, 1.0$, it does not seriously affect $R_{\mathrm{sc}\text{-}e}$. Assuming the entanglement generation rate between remote color centers to be 100 kHz based on the discussion in the previous section, $R_{\mathrm{sc}\text{-}e} \sim 100$ kHz do not seriously degrade the whole entanglement generation rate between the superconducting qubits. Thus, $\eta_e^{\mathrm{MW}} = 0.2$, corresponding to $R_{\mathrm{sc}\text{-}e} \sim 100$ kHz, could be sufficient to generate entanglement between the superconducting qubit and the electron spin. Even $\eta_e^{\mathrm{MW}} \sim 0.02$, corresponding to $R_{\mathrm{sc}\text{-}e} \sim 10$ kHz, could work when the whole entanglement generation rate is limited to a few tens of kilohertz by the other operations (e.g., initialization of the nuclear
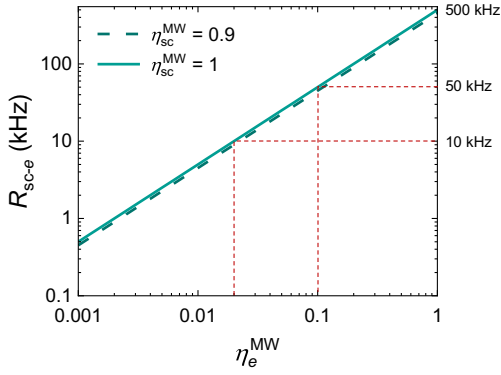
FIG. 3. The entanglement generation rate between the superconducting qubit and electron spin as a function of the absorption (emission) efficiency of a microwave time-bin qubit by the electron spin. The rate is calculated for different absorption (emission) efficiencies of a microwave time-bin qubit by the superconducting qubit, $\eta_{sc}^{MW} = 0.9, 1.0$. The red dashed lines in the graph are shown to clarify the entanglement generation rates of 10 kHz, and 50 kHz, which are considered necessary so as not to seriously degrade the whole entanglement generation rate between the superconducting qubits.

spin, BSM for the electron spin). Although the interaction between a microwave photon and a single electron spin still needs proof-of-concept experiments, our numerical simulation indicates that $\eta_e^{MW} > 0.1$ is feasible with 1D phononic crystal with small mode volume ($<0.1\lambda^3$) and high strain sensitivity of a specific color center (see Appendix D). Besides, it should be noted that the short lifetime of the excited state of N-V^- and N-V^0 do not affect our protocol though it can degrade η_e^{MW}. Increasing $g_{m\text{-}e}$ compensates for the decrease of η_e^{MW} (see Appendix D).

V. ENTANGLEMENT GENERATION BETWEEN REMOTE SUPERCONDUCTING QUBITS

After the successful entanglement generation between the superconducting qubit and the electron spin in nodes A and B, BSM is performed between the electron and nuclear spins in each node (Fig. 4). Depending on the BSM result, the quantum state of the superconducting qubits in nodes A and B is projected to the states, $(1/\sqrt{2})(|g\rangle_{sc}^A |g\rangle_{sc}^B \pm |e\rangle_{sc}^A |e\rangle_{sc}^B)$ or $(1/\sqrt{2})(|g\rangle_{sc}^A |e\rangle_{sc}^B \pm |e\rangle_{sc}^A |g\rangle_{sc}^B)$. The overall gate sequence of the entanglement generation is shown in Fig. 4. Also, in Table I, we summarize the parameters needed to control and read out the quantum states based on the above discussions and references shown in the caption. Here, we consider three situations: (a) moderate hyperfine coupling and weak microwave drive power, (b) strong hyperfine coupling and moderate microwave drive power, and (c) strong hyperfine coupling and strong microwave drive power. As can be seen from Fig. 4 , the time for BSM, τ^{BSM}, is $\tau^{C_n NOT_e} + \tau_e^{\pi/2} + \tau^{C_e NOT_n} + 2\tau_e^{SS}$, where

$\tau^{C_n NOT_e}$ is the time to perform a $C_n NOT_e$ gate on the electron spin, $\tau_e^{\pi/2}$ is the time to perform a $\pi/2$ gate on the electron spin, $\tau^{C_e NOT_n}$ is the time to perform a $C_e NOT_n$ gate on the nuclear spin, τ_e^{SS} is the time to perform a single-shot readout of the electron spin. Based on parameters in Table I, τ^{BSM} is 20–30 μs. To efficiently generate entanglement, the time for the BSM must be shorter than the coherence time of the superconducting qubits. State-of-the-art superconducting qubits have sufficient performance, since their coherence time reaches 100 μs [1] or even approximately 1 ms [39], which is longer than the BSM time. Of note, a much quicker BSM is still required to generate the entanglement with high fidelity.

Considering the success probability of entanglement generation between the remote electron spins and that between the superconducting qubit and the electron spin, the average entanglement generation rate between the superconducting qubits, $R_{sc\text{-}sc}^{mem}$, is

$$R_{sc\text{-}sc}^{mem} = \frac{1}{\tau^{init} + \tau^{e\text{-}e} + \tau_n^{SWAP} + \tau^{sc\text{-}e} + \tau^{BSM}}, \quad (9)$$

where $\tau^{init} = \tau_e^{init} + \tau_n^{init}$ is the time to initialize the electron (τ_e^{init}) and the nuclear spins (τ_n^{init}), $\tau^{e\text{-}e} = 2\tau_e^{ent}/(\eta_e^{opt})^2$ is the time to generate entanglement between remote electron spins without initialization, τ_e^{ent} is the time to emit an entangled photon from the electron spin, τ_n^{SWAP} is the time to swap the state of the electron spin with that of the nuclear spin, and $\tau^{sc\text{-}e}$ is the average time to generate entanglement between the electron spin and the superconducting qubit in both nodes. Considering the expected number of entanglement generation attempts until success in both nodes, $\tau^{sc\text{-}e}$ can be calculated by

$$\tau^{sc\text{-}e} = \left\{ p^2 + \sum_{i=0}^{\infty}(i+2)p^2(1-p)^{i+1} \times [(1-p)^{i+1} + 2\sum_{j=0}^{i}(1-p)^j] \right\} / r_{sc\text{-}e}, \quad (10)$$

where $p = \eta_{sc}^{MW}\eta_e^{MW}$ is the success probability of the entanglement generation attempts, $r_{sc\text{-}e} = 1/(\tau_{MW}^{emi} + \tau_{MW}^{abs})$ is the rate of the entanglement generation attempts, τ_{MW}^{emi} is time to emit a microwave time-bin qubit at the electron spin (superconducting qubit), and τ_{MW}^{abs} is time to absorb a microwave time-bin qubit at the superconducting qubit (electron spin). Here, the success probability of the first trial is p^2 and that of the nth ($n > 1$) trial is $p^2(1-p)^{n-1} \times [(1-p)^{n-1} + 2\sum_{j=0}^{n-2}(1-p)^j]$, which is derived from a simple calculation. In Table I, assuming the highest efficiency, $\eta_e^{opt} = 1$ and $\eta_{sc}^{MW} = \eta_e^{MW} = 1$, we estimate $R_{sc\text{-}sc}^{mem}$ to be 10–30 kHz at maximum.

If we can use multiple quantum spin memories [21] as shown in Fig. 5, we can prepare the remote entanglement
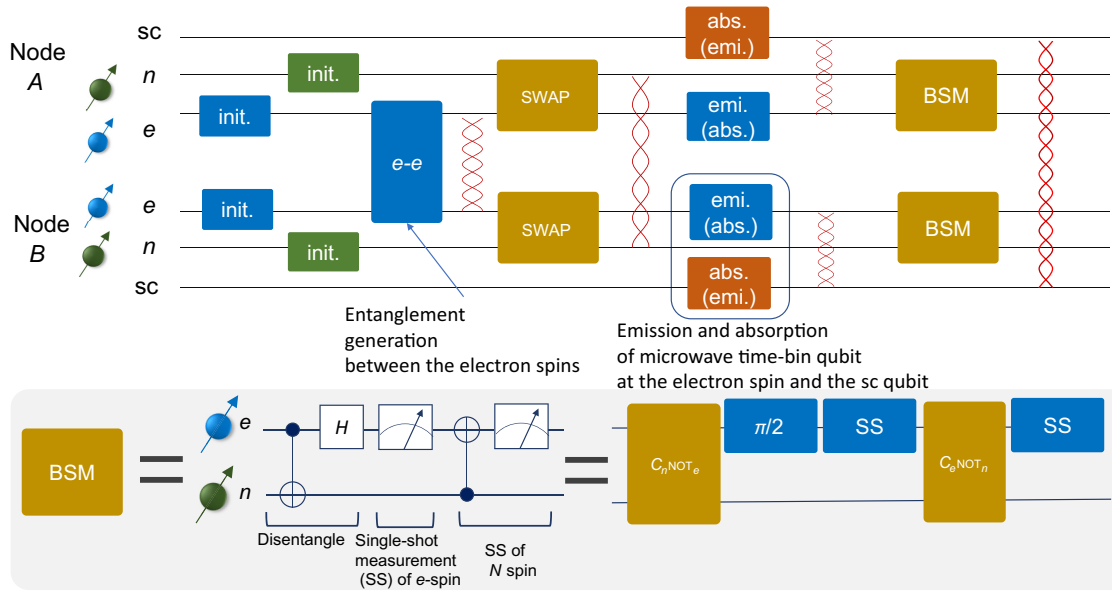
FIG. 4.   A gate sequence for the entanglement generation between remote superconducting qubits. Initialization of the electron (e) and nuclear (n) spins (init.) is followed by entanglement generation between the electron spins in the different nodes (e-e). Entanglement between the electron spins is then swapped to the nuclear spins (SWAP). Entanglement between the superconducting (sc) qubit and the electron spin is generated via the absorption (abs.) and emission (emi.) of the microwave time-bin qubit. Finally, BSM is performed for the electron and the nuclear spins to generate entanglement between the superconducting qubits. A BSM sequence is shown in the gray box below the whole gate sequence to estimate the BSM gate time. The state of the electron spin is measured by a single-shot measurement (SS). The times for each operation are summarized in Table I.

between electron spins during the BSM of another spin memory. Based on the assumption shown in Table I, a system with an additional parallel memory is sufficient to maintain the entanglement between the remote spins and to supply it to the superconducting qubits. The overhead time to generate electron spin-spin entanglement is again eliminated, resulting in the entanglement attempt rate of approximately 40 kHz at maximum, which is limited by the speed of the BSM of the electron and nuclear spins.

VI. DISCUSSION

We estimate the entanglement generation rate between remote superconducting qubits using the quantum spin memory. Here, we compare the performance of the memory-based scheme to those of other protocols without memory.

A. Comparison to direct conversion scheme

Direct conversion of a microwave to an optical photon is a straightforward approach transferring a quantum state of a superconducting qubit or microwave photon itself outside a dilution refrigerator. Direct conversion makes it possible to implement an entanglement generation protocol between remote superconducting qubits using the two-photon protocol. Here, though the single-photon protocol can be used, we consider only the two-photon protocol,

in which stabilization of the optical path is not necessary. Figure 6 shows a schematic of the protocol. First, the superconducting qubits coupled to the cavity emit the coherent time-bin qubit,

$$\frac{1}{\sqrt{2}}(|g\rangle_{\text{sc}}|L\rangle_{\text{MW}} + |e\rangle_{\text{sc}}|E\rangle_{\text{MW}}). \qquad (11)$$

The microwave time-bin qubit is then converted to the photonic time-bin qubit at the transducer. The photonic time-bin qubits from nodes A and B interfere at a beam splitter and are measured by a photodetector. Consequently, the superconducting qubits in different nodes are projected to the entangled state as

$$\frac{1}{\sqrt{2}}(|e\rangle_{\text{sc}}^{A}|g\rangle_{\text{sc}}^{B} \pm |g\rangle_{\text{sc}}^{A}|e\rangle_{\text{sc}}^{B}), \qquad (12)$$

where the sign depends on which detector is clicked. The same as in Eq. (2), the entanglement generation rate in this direct conversion case, $R_{\text{sc-sc}}^{\text{DC}}$, can be expressed as

$$R_{\text{sc-sc}}^{\text{DC}} = \frac{1}{2}r_{\text{sc-sc}}\eta_{\text{DC}}^2(\eta_{\text{loss}}^{\text{DC}})^2, \qquad (13)$$

where $r_{\text{sc-sc}}$ is the entanglement attempt rate corresponding to the speed with which the microwave time-bin qubit is generated at a superconducting qubit, η_{DC} is

TABLE I. Parameters related to the entanglement generation protocol in three different situations. (a) Moderate hyperfine coupling to a nuclear spin and weak microwave driving power. (b) Strong hyperfine coupling to a nuclear spin and moderate microwave driving power. (c) Strong hyperfine coupling to a nuclear spin and strong microwave driving power. The parameters are the hyperfine coupling between the electron spin and the nuclear spin, A, the Rabi frequency of the electron spin, Ω_e^{Rabi}, the Rabi frequency of the nuclear spin, Ω_n^{Rabi}, time to initialize the electron spin, τ_e^{init}, time to initialize the nuclear spin, τ_n^{init}, time to generate an entangled photon from the electron spin, τ_e^{ent}, time to swap the electron spin state to that of the nuclear spin, τ_n^{SWAP}, time to emit a microwave time-bin qubit from the superconducting qubit (spin), $\tau_{\text{MW}}^{\text{emi}}$, time to receive a microwave time-bin qubit at the spin (superconducting qubit), $\tau_{\text{MW}}^{\text{abs}}$, time to perform a Hadamard gate, $\tau_e^{\pi/2}$, time to perform a $C_e\text{NOT}_n$ gate, $\tau^{C_e\text{NOT}_n}$, time to perform a $C_n\text{NOT}_e$ gate, $\tau^{C_n\text{NOT}_e}$, time to perform a single-shot measurement of the electron spin, τ_e^{SS}, and the maximum entanglement generation rate, $R_{\text{sc-sc}}^{\text{mem}}$, assuming $\eta^{\text{opt}} = 1$, $\eta_{\text{sc}}^{\text{MW}} = \eta_e^{\text{MW}} = 1$. The gate times for the electron and nuclear spins ($\tau_e^{\pi/2}$, $\tau^{C_e\text{NOT}_n}$, and $\tau^{C_n\text{NOT}_e}$) are calculated based on Ω_e^{Rabi} and Ω_n^{Rabi} except in the case of $A \sim 1$ MHz. We consider to use a π pulse to perform $C_e\text{NOT}_n$ and $C_n\text{NOT}_e$ like Ref. [40]. When $A \sim 1$ MHz, $\tau^{C_e\text{NOT}_n}$ is determined according to Ref. [41] and $\tau^{C_n\text{NOT}_e}$ is calculated from Ω_e^{Rabi} assuming the use of a geometric gate (2π pulse) [42,43]. τ_e^{init} is determined by referring to experimentally used values [19,27]. (For measurement-based initialization, $\tau_e^{\text{init}} = 10 - 30$ μs [27,29,44].) $\tau_n^{\text{init}} = \tau^{C_e\text{NOT}_n} + \tau^{C_n\text{NOT}_e} + \tau_e^{\text{init}}$, $\tau_n^{\text{SWAP}} = \tau^{C_e\text{NOT}_n} + \tau^{C_n\text{NOT}_e}$. We refer to Ref. [19,27] for τ_e^{ent} and to Ref. [27,29,45] for τ_e^{SS}.

	(a)	(b)	(c)
	Approximately		
A	1 MHz	Approximately	100 MHz [40]
Ω_e^{Rabi}	0.5 MHz	10 MHz	100 MHz
Ω_n^{Rabi}		0.4 MHz	4 MHz
τ_e^{init} [19,27]	5 μs	5 μs	5 μs
τ_n^{init}	17 μs	6.3 μs	5.13 μs
τ_e^{ent} [19,27]	2 μs	2 μs	2 μs
τ_n^{SWAP}	12 μs	1.3 μs	0.13 μs
$\tau_{\text{MW}}^{\text{emi}}$	1 μs	1 μs	1 μs
$\tau_{\text{MW}}^{\text{abs}}$	1 μs	1 μs	1 μs
$\tau_e^{\pi/2}$	0.5 μs	0.025 μs	0.0025 μs
$\tau^{C_e\text{NOT}_n}$	10 μs [41]	1.25 μs	0.125 μs
$\tau^{C_n\text{NOT}_e}$	2 μs	0.05 μs	0.005 μs
τ_e^{SS} [27,29,45]	10 μs	10 μs	10 μs
$R_{\text{sc-sc}}^{\text{mem}}$	14 kHz	25 kHz	27 kHz

the microwave-to-optical conversion efficiency, and $\eta_{\text{loss}}^{\text{DC}}$ includes all losses during the propagation of the photon. For the calculation of $R_{\text{sc-sc}}^{\text{DC}}$, we assume that the superconducting qubit coupled to the microwave cavity can generate a microwave time-bin qubit at a rate of $r_{\text{sc-sc}} = 1$ MHz [30,38]. Here, for simplicity, we do not consider the detail of the transducer. Using Eq. (13) and ignoring any transmission loss in the waveguide or fiber ($\eta_{\text{loss}}^{\text{DC}} = 1$), we
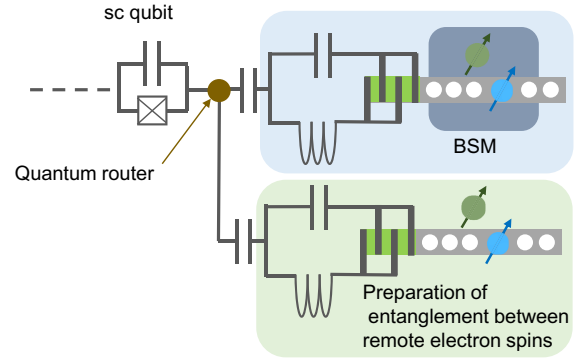

FIG. 5. Schematic of the quantum interface with two parallel memories. A quantum router [46] is used to route a microwave photon to each interface.

calculate an upper limit of $R_{\text{sc-sc}}^{\text{DC}}$. Figure 7 shows $R_{\text{sc-sc}}^{\text{DC}}$ as a function of η_{DC}. To compare the protocol with memory, $R_{\text{sc-sc}}^{\text{mem}}$ is also shown in the same graph as a function of η_e^{MW}, since both η_{DC} and η_e^{MW} are considered to be major limiting factors for $R_{\text{sc-sc}}^{\text{mem}}$. We calculate $R_{\text{sc-sc}}^{\text{mem}}$ using Eq. (9) and assuming $\eta_e^{\text{opt}} = \eta_{\text{sc}}^{\text{MW}} = 0.9$ or 1 for the three cases shown in Table I. Since putting $\eta_e^{\text{opt}} = \eta_{\text{sc}}^{\text{MW}} = 0.9$ decreased $R_{\text{sc-sc}}^{\text{mem}}$ by less than 2 % when $\eta_e^{\text{MW}} > 0.1$, we show only the result of $\eta_e^{\text{opt}} = \eta_{\text{sc}}^{\text{MW}} = 1$ in Fig. 7. Figure 7 shows that when microwave-to-optical transduction efficiency (η_{DC}) is below 0.1 and microwave-to-spin conversion efficiency (η_e^{MW}) is approximately 0.1, the memory-based scheme exceeds the performance of the DC scheme. Besides, in practice, it is generally difficult to reach the regime, $\eta_{\text{DC}} > 0.1$, while maintaining low optical pump power, low thermal noise, and sufficient external microwave (optical) coupling [10]. Although using a color center in a diamond as a single-microwave phonon emitter (receiver) still needs a proof-of-concept experiment, the numerical simulation (see Appendix D) suggests that $\eta_e^{\text{MW}} \sim 0.1$ can be realized in the feasible design of the cavity and color center, promising the usefulness of this memory-based protocol.

B. Comparison to the entangled-photon scheme without memory

To avoid the difficulties involved in the realization of a high-efficiency microwave-to-optical transducer, entangled-photon schemes have recently attracted attention for their ability to generate entanglement between a microwave photon inside a dilution refrigerator and an optical photon outside of it [10–12]. In these schemes, the parametric down-conversion of the blue-detuned optical pump light generates entangled photons at the microwave-optical transducers (e.g., electro-optic transducers, optomechanical transducers). Figure 8 shows a schematic of the protocol. When the transducers at nodes A and B are simultaneously subjected to the blue-detuned
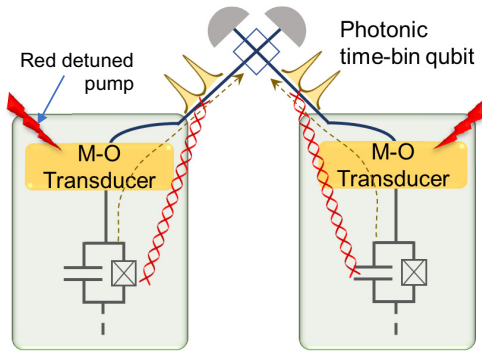
FIG. 6. Schematic to generate entanglement between remote superconducting qubits using a microwave-to-optical (MO) transducer. The two-photon protocol with a red detuned pump (the direct conversion scheme) is shown. The microwave time-bin qubit from the superconducting qubits is converted to a photonic time-bin qubit at the transducer. A beam splitter and photodetectors are used to generate and herald the entanglement between the superconducting qubits.

FIG. 8. Schematic of the generation of entanglement between remote superconducting qubits using microwave-to-optical (MO) transducers. The protocol with a blue detuned pump (the entangled-photon scheme) [11] is shown. An entangled-photon pair is generated at the transducer by the spontaneous parametric down-conversion. As with Fig. 6, a beam splitter and photodetectors are used to generate and herald the entanglement between the microwave photons.

light, an entangled microwave-optical photon pair is generated in either node. The measurement of an optical photon using a beam splitter and photodetectors heralds the
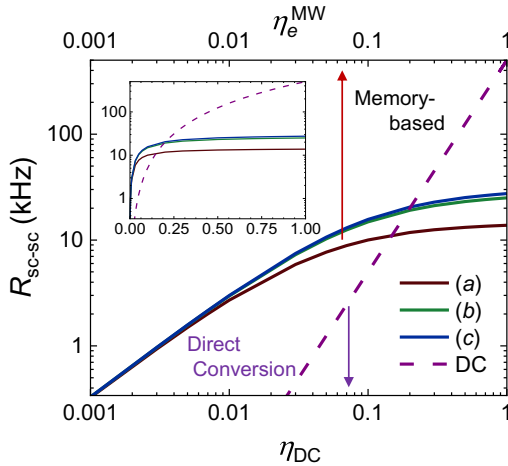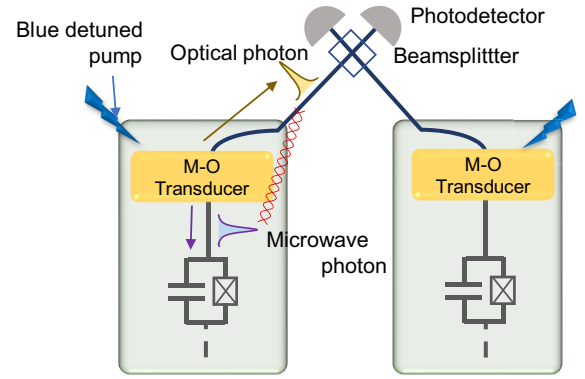

FIG. 7. Entanglement generation rate between remote superconducting qubits with the direct conversion scheme ($R_{\text{sc-sc}}^{\text{DC}}$) and that of the memory-based scheme ($R_{\text{sc-sc}}^{\text{mem}}$) calculated using Eqs. (9) and (13). The inset shows the same plot with a linear horizontal axis scale. The dashed line corresponds to the direct conversion (DC) scheme. It is plotted as a function of microwave-to-optical transduction efficiency, η_{DC}. The solid lines are those of the memory-based scheme. They are each plotted as a function of the absorption (emission) efficiency of the microwave time-bin qubit by the electron spin, η_e^{MW}. (a), (b), and (c) in the legend correspond to the three cases summarized in Table I: (a) weak hyperfine coupling and weak microwave drive, (b) strong hyperfine coupling and moderate microwave drive, (c) strong hyperfine coupling and strong microwave drive.

generation of the entanglement between microwave photons in each node,

$$\frac{1}{\sqrt{2}}(|0\rangle_{\text{MW}}^A |1\rangle_{\text{MW}}^B \pm |1\rangle_{\text{MW}}^A |0\rangle_{\text{MW}}^B), \qquad (14)$$

where the sign depends on the side on which the detector is clicked.

Since the number of microwave photons in each node entangles, the absorption of a microwave photon at the superconducting qubits, which are initially at $|g\rangle_{\text{sc}}$, generates the entanglement as

$$\frac{1}{\sqrt{2}}(|g\rangle_{\text{sc}}^A |e\rangle_{\text{sc}}^B \pm |e\rangle_{\text{sc}}^A |g\rangle_{\text{sc}}^B). \qquad (15)$$

In Ref. [11], the entanglement generation rate is estimated to be 1–10 kHz with in-refrigerator heating of 10 μW and 10–100 kHz with in-refrigerator heating of 100 μW based on state-of-the-art device parameters. This heating comes from the strong optical pump generating entangled photons at a high rate. Since heating on the order of 10 μW cannot be negligible considering the cooling power of the 10 mK region in a dilution refrigerator, the operation of higher-temperature plates with greater cooling power is proposed [10,11].

In the memory-based protocol, in-refrigerator heating mainly comes from the dissipation of the microwave drive pulse and laser light to control and read out the electron spin. To achieve an electron Rabi frequency of 10 MHz, a flux density of approximately 3 G is needed, considering the gyromagnetic ratio of the electron, $\gamma/2\pi \sim$ 3 MHz/G. From Ampere's law, approximately 1 mA current is needed to induce the required magnetic field

at a color center whose distance from the electrode is approximately 1 μm. Then, the input power at the mixing chamber plate should be 150 μW with 50 Ω termination outside the dilution refrigerator. The attenuation of the commercially available superconducting coaxial waveguide is approximately 0.01 dB/m at a few GHz below 4 K [47]. Thus, considering a superconducting waveguide with 1 m length below the mixing chamber (approximately 10 mK) and microwave pulses with a duty cycle of 0.1 [the ratio of microwave (RF) gate time to the whole entanglement generation sequence time], the dissipated power is estimated to be 30 nW, which is more than 2 orders of magnitude smaller than the typical cooling power at the mixing chamber, approximately 10 μW (e.g., LD-series, Bluefors). Even if we use 15-mW microwave pulses, a tenfold increase in the gate speed owing to the high power decreases the duty cycle to 0.01, resulting in dissipation of about 400 nW. Besides, the loss at the cable is suggested to be lowered by 2 orders of magnitude using a waveguide with a well-prepared surface [47], which may further reduce the heat generation. We should note that, we assume here that the contact resistance at each joint or connector can be negligible. It should also be noted that, for color centers whose ground state is the eigenstate of both spin and orbit (e.g., N-V^0, group-IV vacancy color centers), much more microwave power is required than the above estimation, since flipping both spin and orbital quantum numbers using only a magnetic field is forbidden in the first order [48].

The typical optical power needed to control and read out a color center is below 100 nW. Moreover, a readout of the electron spin state in a single-photon regime (0.001–0.1 average photon in a pulse) was demonstrated to be feasible by using a photonic crystal structure [29,44,45]. Thus, the optical input power could also be well below the typical cooling power of the mixing chamber plate.

Although it is difficult to predict heat generation exactly at low temperatures, based on the above rough estimation, the total heat generation below the mixing chamber is 2 to 3 orders of magnitude smaller than that generated in the entangled-photon scheme without memory. In addition, the entanglement generation rate of the memory-based scheme (several tens of kilohertz) is comparable to that of the entangled-photon scheme without memory. Namely, introducing the quantum memory interfacing microwave and optical photon can decrease heat generation, maintaining an entanglement generation rate in the tens of kilohertz at the cost of fabrication and control overhead.

C. Perspectives

Though the memory-based entanglement generation scheme could be a promising option for operation in the 10-mK region, there are still several challenges to overcome. We must integrate all of the necessary

functions (e.g., gate speed, memory time, strain sensitivity), which were demonstrated individually or at different color centers. Moreover, as mentioned above, the interaction between a single electron spin and a microwave photon needs a proof-of-principle experiment.

In contrast to the application to long-distance communication, gate speed to control the electron spin limits the whole entanglement generation rate. Thus, a quantum memory enabling fast initialization, control, and readout with a nuclear spin memory that has moderate memory time (approximately 10 ms) is desirable. It may be necessary to explore a color center that has suitable characteristics. Besides, considering the operation with a superconducting circuit, control of a color center under a zero-magnetic field [43,49,50] should be considered even though some superconducting devices may work in magnetic fields ranging from approximately mT to approximately T [51,52]. Since the proposed scheme can be applied to other quantum memories having both microwave and optical access, such as trapped atoms, ions, and quantum dots, these quantum memories might be alternatives to color centers.

VII. CONCLUSION

We propose a protocol for entanglement generation between remote superconducting qubits using solid-state spin quantum memories, which can be used as interfaces for both microwave and optical photons. Based on the current state-of-the-art technologies, we estimate the maximum entanglement generation rate to be 10–40 kHz. The proposed memory-based scheme's entanglement generation rate can exceed that of the direct conversion scheme as long as the microwave-to-optical conversion efficiency of DC is below 0.1. Moreover, the memory-based scheme can reduce the amount of heat generated below the mixing chamber compared to that generated by the entangled-photon scheme, which uses a strong optical pump to generate entangled-photon pairs. For quantum communication between remote superconducting qubits, the reduction of heat is a central issue considering that operation occurs inside a dilution refrigerator. This suggests the usefulness of introducing the quantum memory.

H. Kurokawa conceived the main protocol and performed the calculation. M.Y. and Y.S. contributed the design of partial protocol. H. Kosaka supervised the project.

APPENDIX A: ENTANGLEMENT GENERATION BETWEEN A SPIN AND PHOTON

1. The single-photon protocol

Here we describe a protocol to generate entanglement between remote spins using a single photon. First, the color-center electron spin in each node is prepared in the superposition state, $(1/\sqrt{2})(|0\rangle_e + |1\rangle_e)$. We assume a situation where $|0\rangle_e$ can be optically excited and emit an optical photon while $|1\rangle_e$ cannot (for experiments about the generation of a coherent photon using a color center, see Refs. [19,27,45]). Then, just after the emission of a coherent photon, the spin-photon entangled state is $(1/\sqrt{2})(|0\rangle_e |1\rangle_{\text{opt}} + |1\rangle_e |0\rangle_{\text{opt}})$. The state of the composite system (nodes A and B) is

$$\frac{1}{2}\left[|0\rangle_e^A |0\rangle_e^B |1\rangle_{\text{opt}}^A |1\rangle_{\text{opt}}^B + |0\rangle_e^A |1\rangle_e^B |1\rangle_{\text{opt}}^A |0\rangle_{\text{opt}}^B \right.$$
$$\left. + |1\rangle_e^A |0\rangle_e^B |0\rangle_{\text{opt}}^A |1\rangle_{\text{opt}}^B + |1\rangle_e^A |1\rangle_e^B |0\rangle_{\text{opt}}^A |0\rangle_{\text{opt}}^B \right], \quad \text{(A1)}$$

where $|0\rangle_{\text{opt}}$ ($|1\rangle_{\text{opt}}$) denotes the absence (presence) of an optical photon in the path from the spin to the detector. Then, the path information of the optical photon is erased using a beam splitter before the detection of the photon at the photodetector. In a situation where only a single photon is detected, the state after the beam splitter becomes

$$\frac{1}{2}\left[|0\rangle_e^A |1\rangle_e^B \frac{1}{\sqrt{2}}(|1\rangle_{\text{opt}}^A |0\rangle_{\text{opt}}^B + |0\rangle_{\text{opt}}^A |1\rangle_{\text{opt}}^B) \right.$$
$$\left. + |1\rangle_e^A |0\rangle_e^B \frac{1}{\sqrt{2}}(|1\rangle_{\text{opt}}^A |0\rangle_{\text{opt}}^B - |0\rangle_{\text{opt}}^A |1\rangle_{\text{opt}}^B) \right] \quad \text{(A2)}$$

without normalization. Here, the events detecting zero or two photons are ignored since they do not contribute to the entanglement generation, which is the reason for factor 1/2 in Eq. (2). Thus, detection of a single photon by the detectors heralds the generation of the entangled state, $(1/\sqrt{2})(|0\rangle_e^A |1\rangle_e^B \pm |1\rangle_e^A |0\rangle_e^B)$, depending on the side on which the detectors is clicked. It is noteworthy that this formula applies only to a situation in which the difference in the optical path length between each node and the beam splitter is zero. Practically, we should consider an additional phase depending on the difference in the optical path length. Just before the incidence to the beam splitter, $|0\rangle_e^A |1\rangle_e^B |1\rangle_{\text{opt}}^A |0\rangle_{\text{opt}}^B$ obtain a phase factor $e^{i\theta_A}$ while $|1\rangle_e^A |0\rangle_e^B |0\rangle_{\text{opt}}^A |1\rangle_{\text{opt}}^B$ obtain $e^{i\theta_B}$ during the propagation of optical fiber. By factoring out $e^{i\theta_A}$, $e^{i\theta} \equiv e^{i(\theta_B - \theta_A)}$ remains as a coefficient in the entangled states, $(1/\sqrt{2})(|0\rangle_e^A |1\rangle_e^B \pm$

$e^{i\theta} |1\rangle_e^A |0\rangle_e^B)$. To compensate this phase difference, the optical path length must be stabilized. If it is not, the phase of the resultant state becomes random, decreasing the fidelity of the entangled state.

2. The two-photon protocol

First, the two-photon protocol uses the same procedure as the single-photon protocol to generate the entangled state, with an additional phase as $(1/\sqrt{2})(|0\rangle_e^A |1\rangle_e^B \pm e^{i\theta} |1\rangle_e^A |0\rangle_e^B)$. Then, π pulse exchanging $|0\rangle_e$ and $|1\rangle_e$ is applied to the electron spin in each node, resulting in a state,

$$(1/\sqrt{2})(|1\rangle_e^A |0\rangle_e^B \pm e^{i\theta} |0\rangle_e^A |1\rangle_e^B). \quad \text{(A3)}$$

Then, the same gate operation is applied to emit a coherent photon. Thus, the state before a beam splitter is

$$(1/\sqrt{2}) \left(e^{i\theta_B} |1\rangle_e^A |0\rangle_e^B |0\rangle_{\text{opt}}^A |1\rangle_{\text{opt}}^B \right.$$
$$\left. \pm e^{i\theta_A} e^{i\theta} |0\rangle_e^A |1\rangle_e^B |1\rangle_{\text{opt}}^A |0\rangle_{\text{opt}}^B \right). \quad \text{(A4)}$$

Since $e^{i\theta_A} e^{i\theta} = e^{i\theta_B}$, the additional phase in each term becomes the same, $e^{i\theta_B}$. Thus, using the two-photon protocol, at the cost of the time it takes to interfere with the optical photon twice, we can treat $e^{i\theta}$ appearing in the single-photon protocol as a global phase, which we can ignore.

APPENDIX B: TRANSMISSION EFFICIENCY OF OPTICAL PHOTON EMITTED FROM A COLOR CENTER

The photon transmission efficiency from a color center to a detector, η_e^{opt}, can be defined as

$$\eta_e^{\text{opt}} \equiv \eta_{\text{int}} \eta_{\text{coupling}} \eta_{\text{loss}} \eta_{\text{det}} \quad \text{(B1)}$$

where $\eta_{\text{int}} = C_{\text{opt}}^{\text{coh}}/(1 + C_{\text{coh}}^{\text{opt}})$, $C_{\text{opt}}^{\text{coh}} = 4g_{\text{opt}}^2/[\kappa(\gamma_{\text{rad}} + \gamma_{\text{nonrad}} + \gamma_{\text{dp}})]$ is the optical coherent cooperativity [53] that characterizes the coherent photon emission from an electron spin, g_{opt} is the coupling rate between the emitter and the optical cavity, κ is the total decay rate of the cavity field, γ_{rad} is the free-space radiative decay rate of the emitter, γ_{nonrad} is the nonradiative decay rate of the emitter, γ_{dp} is the pure dephasing rate of the emitter, η_{coupling} is the coupling efficiency to the fiber, η_{loss} is the transmission loss in the fiber, and η_{det} is the photon detection efficiency at a detector. For the spin-cavity system with $C_{\text{opt}}^{\text{coh}} \gg 1$, the coherent photon is emitted almost deterministically. Experimentally, $C_{\text{opt}}^{\text{coh}} > 100$ has already been demonstrated in a Si-V^- center in a 1D photonic crystal cavity, indicating the generation of a coherent photon with a probability exceeding $100/(100 + 1) = 0.99$ [44].

Besides, $\eta_{\text{coupling}} \sim 0.9$ has been achieved using double- and single-sided fiber tapers [48]. If we take the transmission loss at the fiber to be 10 dB/km for the visible light (600–800 nm light considering zero-phonon line of various color centers) and the distance between dilution refrigerators to be 5 m, $\eta_{\text{loss}} \sim 0.99$, which is negligible. Through the use of a superconducting single-photon detector, $\eta_{\text{det}} > 0.99$ has been reported for a 737-nm photon [54]. Hence, considering the present state-of-the-art technologies, $\eta_e^{\text{opt}} \sim 0.9$ is already achievable.

APPENDIX C: EMISSION AND ABSORPTION OF A MICROWAVE TIME-BIN QUBIT IN A SUPERCONDUCTING QUBIT

In this section, we introduce a method to emit and absorb a microwave time-bin qubit in a superconducting qubit based on Ref. [30].

Figure 9 (a) shows a protocol to generate a microwave time-bin qubit entangled with a superconducting qubit. Here, the superconducting qubit is coupled to a microwave resonator. First, the superconducting qubit is prepared to be $(1/\sqrt{2})(|g\rangle_{\text{sc}} + |e\rangle_{\text{sc}})$. Following an excitation of $|e\rangle_{\text{sc}}$ to the second excited state of the superconducting qubit, $|f\rangle_{\text{sc}}$, $|g\rangle_{\text{sc}}$ is excited to $|e\rangle_{\text{sc}}$. According to Ref. [30,31], an appropriate microwave drive pulse induces a transition between $|f\rangle_{\text{sc}}$ and $|g\rangle_{\text{sc}}|1\rangle_{\text{MW}}$, where $|n\rangle_{\text{MW}}$ ($n = 1, 2, 3 \cdots$) denotes Fock state of the coupled resonator. $|g\rangle_{\text{sc}}|1\rangle_{\text{MW}}$ decay into $|g\rangle_{\text{sc}}$ emitting a microwave photon to a waveguide, which is coupled to the microwave resonator [Fig. 9(a) (i)]. The emitted photon entangles

with the superconducting qubit as $|e\rangle_{\text{sc}}|E\rangle_{\text{MW}}$. The resulting state is shown in Fig. 9(a) (ii). Repeating this procedure produces a microwave photon entangled with $|g\rangle_{\text{sc}}$ as $|g\rangle_{\text{sc}}|L\rangle_{\text{MW}}$ [Fig. 9(a) (iii)]. Consequently, a microwave time-bin qubit entangled with the superconducting qubit, $(1/\sqrt{2})(|g\rangle_{\text{sc}}|L\rangle_{\text{MW}} + |e\rangle_{\text{sc}}|E\rangle_{\text{MW}})$, is generated.

Absorption of a microwave time-bin qubit in a superconducting qubit is performed as shown in Fig. 9(b). Here, we assume that an incident microwave time-bin qubit is $\alpha|E\rangle_{\text{MW}} + \beta|L\rangle_{\text{MW}}$. First, a superconducting qubit is prepared to be $|g\rangle_{\text{sc}}$. An absorption of $|E\rangle_{\text{MW}}$ excites $|g\rangle_{\text{sc}}$ to $|g\rangle_{\text{sc}}|1\rangle_{\text{MW}}$. Following microwave drive pulses transfer $|g\rangle_{\text{sc}}|1\rangle_{\text{MW}}$ to $|e\rangle_{\text{sc}}$ via $|f\rangle_{\text{sc}}$ [Fig. 9(b) (i)]. Similarly, the state remained in $|g\rangle_{\text{sc}}$ is excited to $|g\rangle_{\text{sc}}|1\rangle_{\text{MW}}$ by $|L\rangle_{\text{MW}}$. After that, $|g\rangle_{\text{sc}}|1\rangle_{\text{MW}}$ is excited to $|f\rangle_{\text{sc}}$ [Fig. 9(b) (ii)]. The resultant state is $\alpha|e\rangle_{\text{sc}} + \beta|f\rangle_{\text{sc}}$. Then, microwave driving pulses are applied to transfer $|e\rangle_{\text{sc}}$ to $|g\rangle_{\text{sc}}$ and $|f\rangle_{\text{sc}}$ to $|e\rangle_{\text{sc}}$ [Fig. 9(b) (iii)]. Finally an application of the bit-flip pulse result in the state, $\alpha|e\rangle_{\text{sc}} + \beta|g\rangle_{\text{sc}}$ [Fig. 9(b) (iii) (iv)].

APPENDIX D: ABSORPTION OF A MICROWAVE PHOTON BY A COLOR CENTER

In this section, we estimate the efficiency of microwave photon absorption, η_e^{MW}, by a color center considering the energy level of specific color centers, N-V^- and N-V^0. In each subsection, we describe a protocol to absorb a microwave photon by the color center and estimate the efficiency of the protocol by a numerical simulation of the master equation.
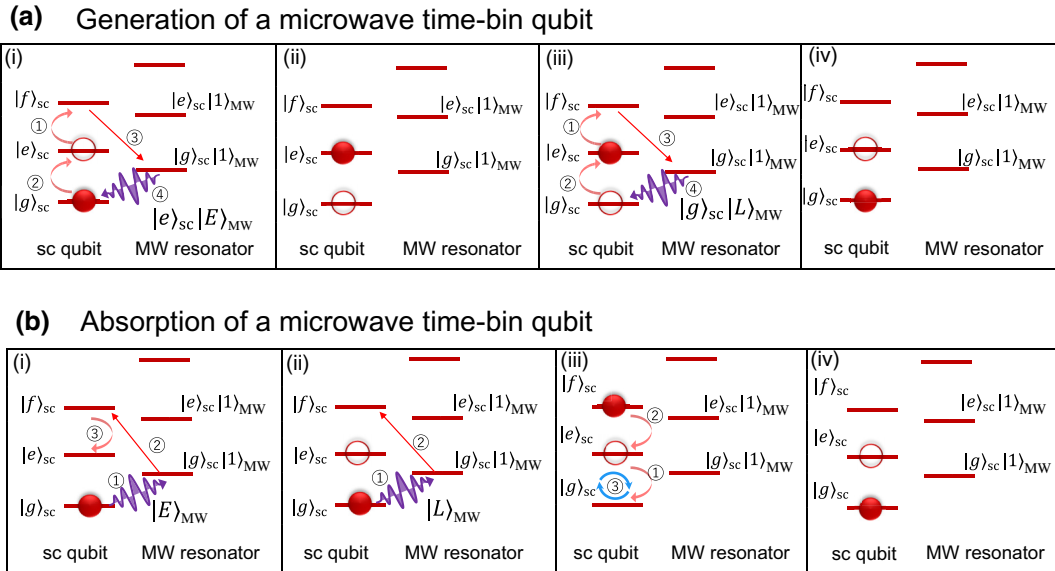
(a) Generation of a microwave time-bin qubit

(b) Absorption of a microwave time-bin qubit

FIG. 9. Schematic of the (a) emission and (b) absorption of a microwave time-bin qubit in a superconducting qubit coupled to a microwave resonator.

1. A color center whose excited state has high strain sensitivity

As mentioned in the main text, a superconducting qubit coupled to a cavity can emit a coherent microwave time-bin qubit,

$$\frac{1}{\sqrt{2}}(|g\rangle_{\rm sc}\,|L\rangle_{\rm MW} + |e\rangle_{\rm sc}\,|E\rangle_{\rm MW}). \tag{D1}$$

We consider a process whereby the electron spin in a diamond absorbs this microwave time-bin qubit and generates entanglement with the superconducting qubit.

For a case of a color center whose optically excited state has high strain sensitivity (e.g., N-V^-), the procedure is shown in Fig. 10 taking N-V^- as an example. First, the electron spin is prepared in the state,

$$\frac{1}{\sqrt{2}}(|0\rangle_e + |1\rangle_e), \tag{D2}$$

where $|0\rangle_e \equiv |m_s = 0\rangle$ and $|1\rangle_e \equiv |m_s = +1\rangle$ for N-V^- and m_s is the angular momentum of the spin. Then, $|0\rangle_e$ is optically virtually excited to $|\phi\rangle_e^{\rm opt}$ with detuning equal to the frequency of the microwave phonon, where $|\phi\rangle_e^{\rm opt}$ is the optically excited state of $|0\rangle_e$ (e.g., $|E_x\rangle$ for N-V^-). This kind of excitation using a red sideband was previously demonstrated in N-V^- using a classical strain field [55]. The electron spin that was initially at $|0\rangle_e$ interacts with $|E\rangle_{\rm MW}$ in its virtually excited state and emits $|E\rangle_{\rm opt}$ via spontaneous emission if $|E\rangle_{\rm MW}$ is absorbed. After the application of the π pulse exchanging $|0\rangle_e$ and $|1\rangle_e$, the state at $|0\rangle_e$ is optically excited again and emits $|L\rangle_{\rm opt}$ conditioned on the successful absorption of $|L\rangle_{\rm MW}$. Then, after the application of the second bit flip exchanging $|0\rangle_e$ and $|1\rangle_e$, $|E\rangle_{\rm opt}$ and $|L\rangle_{\rm opt}$ entangle with the electron spin as

$$\frac{1}{\sqrt{2}}(|L\rangle_{\rm opt}\,|1\rangle_e + |E\rangle_{\rm opt}\,|0\rangle_e). \tag{D3}$$

Following the measurement of the photonic time-bin qubit in the basis, $(1/\sqrt{2})(|0\rangle_{\rm opt}^A |1\rangle_{\rm opt}^B \pm |1\rangle_{\rm opt}^A |0\rangle_{\rm opt}^B)$, using a delay line and a beam splitter, the application of the phase-flip gate to the electron spin depending on the measurement outcome teleports the quantum state of the phonon to the electron spin. Consequently, entanglement between the superconducting qubit and the microwave time-bin qubit is transferred to the electron spin as

$$\frac{1}{\sqrt{2}}(|g\rangle_{\rm sc}\,|1\rangle_e + |e\rangle_{\rm sc}\,|0\rangle_e). \tag{D4}$$

Then, we estimate $\eta_e^{\rm MW}$ from the calculation of microwave-to-optical conversion efficiency. Since the detection of an optical photon heralds the entanglement generation between the superconducting qubit and

the color center as shown in Fig. 10, we assume that microwave-to-optical transduction efficiency is nearly equal to $\eta_e^{\rm MW}$.

The Hamiltonian of the quantum interface driven by the optical field is modeled as

$$H_{\rm QI,N\text{-}V^-} = H_0 + H_{\rm int} + H_{\rm drive}, \tag{D5}$$

$$H_0 = \hbar\omega_{\rm MW} a_{\rm MW}^\dagger a_{\rm MW} + \hbar\omega_m b_m^\dagger b_m + \hbar\omega_e \sigma_e^+ \sigma_e$$
$$+ \hbar\omega_{\rm opt} c_{\rm opt}^\dagger c_{\rm opt}, \tag{D6}$$

$$H_{\rm int} = \hbar g_{\rm MW\text{-}m}(a_{\rm MW}^\dagger b_m + a_{\rm MW} b_m^\dagger)$$
$$+ \hbar g_{m\text{-}e}(b_m^\dagger + b_m)\sigma_e^+ \sigma_e + g_{e\text{-opt}}$$
$$\times (\sigma_e^+ c_{\rm opt} + \sigma_e c_{\rm opt}^\dagger) \tag{D7}$$

$$H_{\rm drive} = \hbar \frac{\Omega_{\rm drive}}{2}(e^{-i\omega_d t}\sigma_e^+ + e^{i\omega_d t}\sigma_e), \tag{D8}$$

where $\omega_{\rm MW}$, ω_m, ω_e, and $\omega_{\rm opt}$ are the resonance frequency of the microwave cavity, the resonance frequency of the phononic crystal cavity, the energy gap between the ground state and optical excited state of a color center, and the resonance frequency of the optical cavity. $a_{\rm MW}$ ($a_{\rm MW}^\dagger$), b_m ($b_m^\dagger$), and $c_{\rm opt}$ ($c_{\rm opt}^\dagger$) are the annihilation (creation) operator of the photon (phonon) in the microwave cavity, the phononic crystal cavity, and the optical cavity, respectively. σ_e (σ_e^+) is the electron lowering (raising) operator between the ground state and the optically excited state. $g_{\rm MW\text{-}m}$, $g_{m\text{-}e}$, and $g_{e\text{-opt}}$ denote the coupling between a microwave mode and a mechanical mode in each cavity, the coupling between a mechanical mode (strain) and a color center, the coupling between a color center and an optical mode, respectively. $\Omega_{\rm drive}$ is the optical Rabi frequency of the driving field, and ω_d is the frequency of the driving field.

Applying the unitary transformation, $e^U H e^{-U}$ to the Eq. (D5) with $U = (g/\omega_m)(b_m^\dagger - b_m)\sigma_e^+ \sigma_e$ [56] results in

$$H_0' = \hbar\omega_{\rm MW} a_{\rm MW}^\dagger a_{\rm MW} + \hbar\omega_m$$
$$\times \left[b_m^\dagger b_m - \frac{g_{m\text{-}e}}{\omega_m}(b_m^\dagger + b_m)\sigma_e^+ \sigma_e + \left(\frac{g_{m\text{-}e}}{\omega_m}\right)^2 \sigma_e^+ \sigma_e \right]$$
$$+ \hbar\omega_e \sigma_e^+ \sigma_e + \hbar\omega_{\rm opt} c_{\rm opt}^\dagger c_{\rm opt}, \tag{D9}$$

$$H_{\rm int}' = \hbar g_{\rm MW\text{-}m} \left[(a_{\rm MW}^\dagger b_m + a_{\rm MW} b_m^\dagger) \right.$$
$$\left. - \frac{g_{m\text{-}e}}{\omega_m}(a_{\rm MW}^\dagger + a_{\rm MW})\sigma_e^+ \sigma_e \right]$$
$$+ \hbar g_{m\text{-}e} \left[(b_m^\dagger + b_m)\sigma_e^+ \sigma_e - 2\frac{g_{m\text{-}e}}{\omega_m}\sigma_e^+ \sigma_e \right]$$
$$+ \hbar g_{e\text{-opt}} \left[e^{\frac{g_{m\text{-}e}}{\omega_m}(b_m^\dagger - b_m)}\sigma_e^+ c_{\rm opt} + e^{\frac{g_{m\text{-}e}}{\omega_m}(b_m - b_m^\dagger)}\sigma_e c_{\rm opt}^\dagger \right], \tag{D10}$$
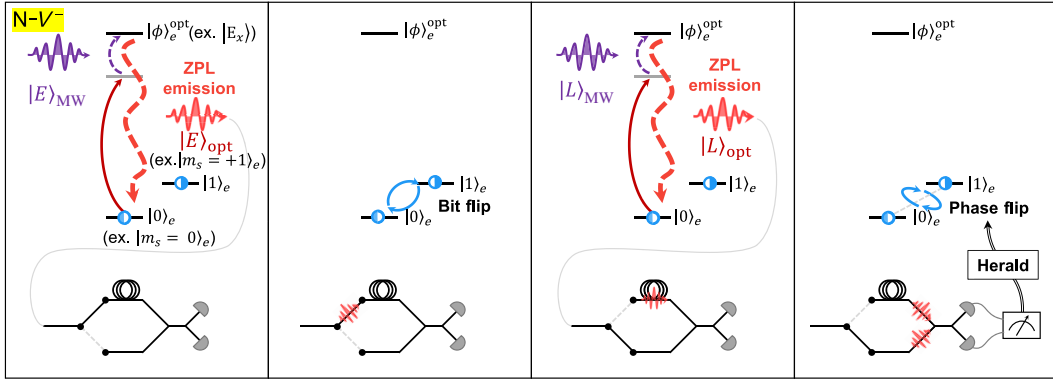
FIG. 10. Schematic of the absorption of a microwave time-bin qubit by an electron spin whose optically excited state has high strain sensitivity. The energy level of N-V^- is considered as an example.

$$H'_{\text{drive}} = \hbar \frac{\Omega_{\text{drive}}}{2} \left[e^{-i\omega_d t} e^{\frac{g_{m\text{-}e}}{\omega_m}(b_m^\dagger - b_m)} \sigma_e^+ \right.$$

$$\left. + e^{i\omega_d t} e^{\frac{g_{m\text{-}e}}{\omega_m}(b_m - b_m^\dagger)} \sigma_e \right]. \tag{D11}$$

Then, expanding the exponential terms in $g_{m\text{-}e}/\omega_m$ and assuming that $\omega_{\text{MW}}/2\pi = \omega_m/2\pi = 5$ GHz, $\omega_e = \omega_{\text{opt}} = 470$ THz, $g_{\text{MW-}m} = 1$ MHz, $g_{m\text{-}e} = 1 - 10$ MHz, $g_{e\text{-opt}} = 1$ GHz, $\Omega_{\text{drive}} = 5$ GHz, we drop off terms with $g_{\text{MW-}m} g_{m\text{-}e}/\omega_m, g_{m\text{-}e}^2/\omega_m$, resulting in

$$H'_0 + H'_{\text{int}} = \hbar \omega_{\text{MW}} a_{\text{MW}}^\dagger a_{\text{MW}} + \hbar \omega_m b_m^\dagger b_m + \hbar \omega_e \sigma_e^+ \sigma_e$$

$$+ \hbar \omega_{\text{opt}} c_{\text{opt}}^\dagger c_{\text{opt}} + \hbar g_{\text{MW-}m} (a_{\text{MW}}^\dagger b_m + a_{\text{MW}} b_m^\dagger)$$

$$+ \hbar g_{e\text{-opt}} \left\{ \left[1 + \frac{g_{m\text{-}e}}{\omega_m}(b_m^\dagger - b_m) \right] \sigma_e^+ c_{\text{opt}} \right.$$

$$\left. + \left[1 + \frac{g_{m\text{-}e}}{\omega_m}(b_m - b_m^\dagger) \right] \sigma_e c_{\text{opt}}^\dagger \right\} \tag{D12}$$

and

$$H'_{\text{drive}} = \hbar \frac{\Omega_{\text{drive}}}{2} \left\{ e^{-i\omega_d t} \left[1 + \frac{g_{m\text{-}e}}{\omega_m}(b_m^\dagger - b_m) \sigma_e^+ \right] \right.$$

$$\left. + e^{i\omega_d t} \left[1 + \frac{g_{m\text{-}e}}{\omega_m}(b_m - b_m^\dagger) \right] \sigma_e \right\}. \tag{D13}$$

Moving to a rotating frame of the driving optical field with $H_r = \hbar \omega_d \sigma_e^+ \sigma_e + \hbar \omega_d c_{\text{opt}}^\dagger c_{\text{opt}}$ and dropping off the time-dependent terms (rotating wave approximation) and constants, the total Hamiltonian becomes

$$H'_{\text{QI,N-}V^-} = \hbar \omega_{\text{MW}} a_{\text{MW}}^\dagger a_{\text{MW}} + \hbar \omega_m b_m^\dagger b_m$$

$$+ \hbar \Delta_e \sigma_e^+ \sigma_e + \hbar \Delta_{\text{opt}} c_{\text{opt}}^\dagger c_{\text{opt}}$$

$$+ \hbar g_{\text{MW-}m} (a_{\text{MW}}^\dagger b_m + a_{\text{MW}} b_m^\dagger)$$

$$+ \hbar \frac{\Omega_{\text{drive}} g_{m\text{-}e}}{2\omega_m} \left[(b_m^\dagger - b_m)\sigma_e^+ + (b_m - b_m^\dagger)\sigma_e \right]$$

$$+ \hbar g_{e\text{-opt}} \left\{ \left[1 + \frac{g_{m\text{-}e}}{\omega_m}(b_m^\dagger - b_m) \right] \sigma_e^+ c_{\text{opt}} \right.$$

$$\left. + \left[1 + \frac{g_{m\text{-}e}}{\omega_m}(b_m - b_m^\dagger) \right] \sigma_e c_{\text{opt}}^\dagger \right\}, \tag{D14}$$

where $\Delta_e = \omega_e - \omega_d$ and $\Delta_{\text{opt}} = \omega_{\text{opt}} - \omega_d$ are the detuning from the driving optical field.

We calculate the time evolution of the density matrix, ρ, of this system based on the master equation,

$$\frac{d\rho}{dt} = \frac{1}{i\hbar}[H'_{\text{QI,N-}V^-}, \rho] + \sum_i \left[\frac{\gamma_{c_i}}{2}(2c_i \rho c_i^\dagger - \{c_i^\dagger c_i, \rho\}) \right], \tag{D15}$$

where $c_i \in \left\{ a_{\text{MW}}, b_m, \sigma_e, c_{\text{opt}}, c_{\text{opt}} d_{\text{wg}}^\dagger \right\}$ and $\gamma_{c_i} \in \{\gamma_{\text{MW}}, \gamma_m, \gamma_e, \gamma_{\text{opt}}, \gamma_{\text{wg}}\}$ corresponding to the decay (decoherence) rates of the respective excited states. $c_{\text{opt}} d_{\text{wg}}^\dagger$ is introduced phenomenologically to reproduce the coupling (decay) to an optical waveguide, where $d_{\text{wg}}^\dagger$ is the creation operator of the coupled mode of the optical waveguide. γ_{wg} is the decay rate to the optical waveguide from the optical cavity.

To estimate the absorption efficiency of a microwave photon by the electron spin, we performed a calculation based on Eq. (D15) with an initial state in which a microwave photon is in the microwave cavity. Here, we calculate the population of the optical waveguide since an optical photon is used to herald the entanglement generation (microwave photon absorption) at the electron spin.

Actually, the microwave photon comes from a waveguide coupled to the microwave cavity. However, for simplicity, we assume that the coupling between the microwave cavity and the waveguide can be arbitrarily tuned. Once the microwave photon enters the
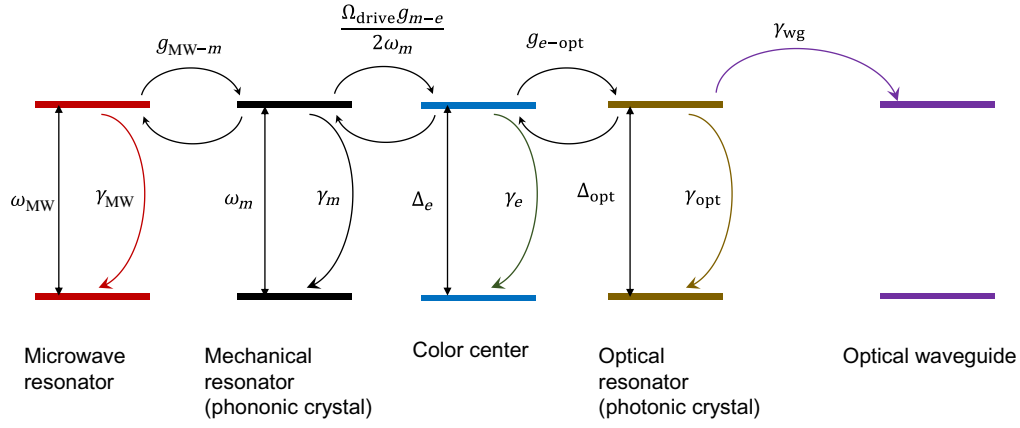
FIG. 11.   Schematic of energy diagram and interaction in the quantum interface.

microwave cavity, the coupling to the waveguide turns off and the waveguide can be ignored. Thus, we focus on a system consisting of four qubits and the optical waveguide. We choose the following parameters for the calculation: $\omega_{\mathrm{MW}}/2\pi = \omega_m/2\pi = \Delta_e/2\pi = \Delta_{\mathrm{opt}}/2\pi =$ 5 GHz, $\{g_{\mathrm{MW}\text{-}m}/2\pi, g_{m\text{-}e}/2\pi, g_{e\text{-opt}}/2\pi, \Omega_{\mathrm{drive}}/2\pi\} = \{1$ MHz [57,58], 0.1–10 MHz [21], 1 GHz [59], 5 GHz$\}$ $\{\gamma_{\mathrm{MW}}/2\pi, \gamma_m/2\pi, \gamma_e/2\pi, \gamma_{\mathrm{opt}}/2\pi, \gamma_{\mathrm{wg}}/2\pi\} = \{500$ kHz, 500 kHz, 1–100 MHz, 20 GHz [45], 20 GHz [45]$\}$ corresponding to the quality factor of 1×10^4 in the microwave cavity and the phononic crystal cavity. These parameters are shown in Fig. 11. It should be noted that, $g_{m\text{-}e}/2\pi \sim 1$ MHz is achievable only for color centers with high strain sensitivity, such as the excited state of N-V^- and the ground and excited states of Si-V^-. Besides, the color center must be placed in a phononic crystal cavity with small mode volume ($< 0.1\lambda^3$).

Figure 12(a) shows the time evolution in the quantum interface assuming that microwave resonator is in its excited state (having one microwave photon) in the initial state (time = 0 ns). Considering that the population in the optical waveguide (purple line) corresponds to the transmission efficiency of a microwave photon, the microwave-to-optical transmission efficiency is approximately 1% for $g_{m\text{-}e}/2\pi = 2$ MHz and $\gamma_e/2\pi = 10$ MHz. Although the efficiency around 1% is not high, the 1% efficiency works if we can accept the entanglement generation rate between a superconducting qubit and a color center being 1–10 kHz (as mentioned in Sec. IV in the paper). Figure 12(b) shows the maximum population in the optical waveguide after the time evolution of 1000 ns as functions of $g_{m\text{-}e}/2\pi$ and $\gamma_e/2\pi$. Assuming that $g_{m\text{-}e}/2\pi$ can reach 1–10 MHz with a 1D phononic crystal and $\gamma_e/2\pi$ is 10–20 MHz, N-V^- is in a regime surrounded by a black dashed line. If we can make a phononic crystal realizing $g_{m\text{-}e}/2\pi \sim 10$ MHz, the conversion efficiency (i.e., entanglement generation efficiency) becomes on the order of 10%, which is sufficient for our protocol.

2. A color center whose ground state has high strain sensitivity

For a color center whose ground state has high strain sensitivity (e.g., N-V^0, Si-V^-), a procedure to absorb a microwave time-bin qubit is shown in Fig. 13. Here, we take N-V^0 as an example since the ground-state splitting due to the spin-orbit coupling is approximately 10 GHz, which makes it easier to access the excited states than with the group-IV color centers, whose ground-state splitting range is from 50 to thousands of gigahertz [48]. Besides, since its ground state has the orbital degree of freedom, it is considered to be sensitive to the strain and electric field, similar to the excited state of N-V^-.

Before explaining the detail of Fig. 13, we first consider the wave function of the ground state of N-V^0 and show that the energy level can be used for the protocol we describe. The ground state of N-V^0 is a spin doublet [60]. The model Hamiltonian of N-V^0 [60,61] is

$$H_{\mathrm{N}\text{-}V^0} = g\mu_B S_z B_z + l\mu_B L_z B_z + 2\lambda L_z S_z + \epsilon_\perp (L_- + L_+), \tag{D16}$$

where g is the spin g factor, μ_B is the Bohr magneton, l is the orbital g factor, λ is the spin-orbit interaction parameter, and $\epsilon_\perp$ is the perpendicular strain parameter. $L_z = \sigma_z$ and $S_z = (1/2)\sigma_z$ are the orbital ($|+\rangle_{\mathrm{EO}}$ and $|-\rangle_{\mathrm{EO}}$ basis) and spin operators, while $L_\pm = |\pm\rangle_{\mathrm{EO}} \langle \mp|_{\mathrm{EO}}$ with $|\pm\rangle_{\mathrm{EO}} = \mp(1/\sqrt{2})(|X\rangle_{\mathrm{EO}} \pm i|Y\rangle_{\mathrm{EO}})$ are the orbital operators, and $|X\rangle_{\mathrm{EO}}, |Y\rangle_{\mathrm{EO}}$ are the strain eigenstates. The z axis is defined parallel to the N-V axis. Here, we use the subscripts EO and ES to distinguish the electron orbital state and the electron spin state.

The ground state with zero strain and zero electric field can be described by the eigenstates, $|+\rangle_{\mathrm{EO}}|\downarrow\rangle_{\mathrm{ES}}$, $|-\rangle_{\mathrm{EO}}|\uparrow\rangle_{\mathrm{ES}}$, $|-\rangle_{\mathrm{EO}}|\downarrow\rangle_{\mathrm{ES}}$, $|+\rangle_{\mathrm{EO}}|\uparrow\rangle_{\mathrm{ES}}$. Since the lower energy states, $(|+\rangle_{\mathrm{EO}}|\downarrow\rangle_{\mathrm{ES}}, |-\rangle_{\mathrm{EO}}|\uparrow\rangle_{\mathrm{ES}})$, have different spin and orbital characteristics, controlling these states
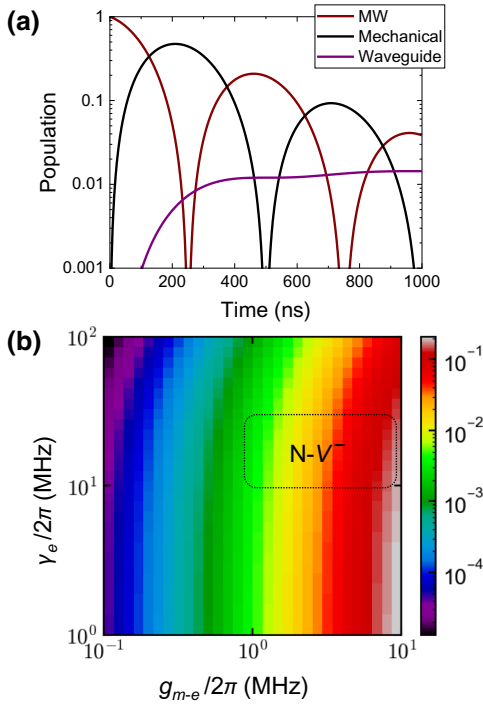
FIG. 12. (a) Time evolution of the population in the quantum interface for $g_{m\text{-}e}/2\pi = 2$ MHz and $\gamma_e/2\pi = 10$ MHz. In the legend, MW, mechanical, and waveguide correspond to the microwave resonator, the mechanical resonator (the phononic crystal), and the optical waveguide, respectively. (b) Maximum population in the optical waveguide after the time evolution as functions of $g_{m\text{-}e}/2\pi$ and $\gamma_e/2\pi$.

using only a magnetic field or strain is inhibited in the first order. On the other hand, in a high-strain regime, the mixing of orbital states results in the eigenstates, $|X\rangle_{\text{EO}}|\downarrow\rangle_{\text{ES}}$, $|X\rangle_{\text{EO}}|\uparrow\rangle_{\text{ES}}$, $|Y\rangle_{\text{EO}}|\downarrow\rangle_{\text{ES}}$, $|Y\rangle_{\text{EO}}|\uparrow\rangle_{\text{ES}}$. Thus, flipping a spin using a magnetic field is possible in the lower energy states, $(|X\rangle_{\text{EO}}|\downarrow\rangle_{\text{ES}}, |X\rangle_{\text{EO}}|\uparrow\rangle_{\text{ES}})$, since they have the same orbital characteristics. Besides, $|X\rangle_{\text{EO}}|\downarrow\rangle_{\text{ES}}$

is excited to $|Y\rangle_{\text{EO}}|\downarrow\rangle_{\text{ES}}$ using only the strain. Then, we express these states using the notation in the main text, $|X\rangle_{\text{EO}}|\downarrow\rangle_{\text{ES}} \equiv |0\rangle_e$, $|X\rangle_{\text{EO}}|\uparrow\rangle_{\text{ES}} \equiv |1\rangle_e$, and $|Y\rangle_{\text{EO}}|\downarrow\rangle_{\text{ES}} \equiv |\phi\rangle_e^{\text{MW}}$. We also apply a static magnetic field to split the degenerate energy levels, $|\downarrow\rangle_{\text{ES}}$ and $|\uparrow\rangle_{\text{ES}}$.

To generate entanglement between a superconducting qubit and the electron spin, $|E\rangle_{\text{MW}}$ excites $|0\rangle_e$ to $|\phi\rangle_e^{\text{MW}}$ as shown in Fig. 13. Then, $|\phi\rangle_e^{\text{MW}}$ is optically excited with a laser pulse and emits $|E\rangle_{\text{opt}}$. This procedure is repeated for $|L\rangle_{\text{MW}}$, generating the entanglement between the optical photon and the electron spin, $(1/\sqrt{2})(|E\rangle_{\text{opt}}|0\rangle_e + |L\rangle_{\text{opt}}|1\rangle_e)$. The rest of the operation to generate entanglement between the superconducting qubit and the electron spin is the same as the protocol described in Appendix D 1.

For N-V^0, we also perform the simulation of the hybrid quantum system to estimate entanglement generation (microwave absorption) efficiency between the microwave photon and the color center. Here, we assume that the optical excitation shown in Fig. 13 is performed within nanoseconds, which is much faster than the orbital relaxation time of N-V^0 (hundreds of nanoseconds). Thus, in the following, we consider that microwave-to-electron spin conversion efficiency is the same as the microwave-to-optical conversion efficiency assuming that an optical photon can be extracted nearly 100% efficiency using a photonic crystal structure.

The Hamiltonian of the system is modeled as

$$
\begin{aligned}
H_{\text{QI,N-}V^0} = \; & \hbar\omega_{\text{MW}}a_{\text{MW}}^\dagger a_{\text{MW}} + \hbar\omega_m b_m^\dagger b_m + \hbar\omega_e^{\text{N-}V^0}\sigma_e^+\sigma_e \\
& + \hbar g_{\text{MW-}m}(a_{\text{MW}}^\dagger b_m + a_{\text{MW}}b_m^\dagger) \\
& + \hbar g_{m\text{-}e}(b_m^\dagger\sigma_e + b_b\sigma_e^+),
\end{aligned}
\tag{D17}
$$

where $\omega_e^{\text{N-}V^0}$ is the resonance frequency of the orbital excited state of N-V^0. Here, σ_e (σ_e^+) is the orbital lowering (raising) operator for N-V^0. We calculate the time evolution of the system using the master equation. The
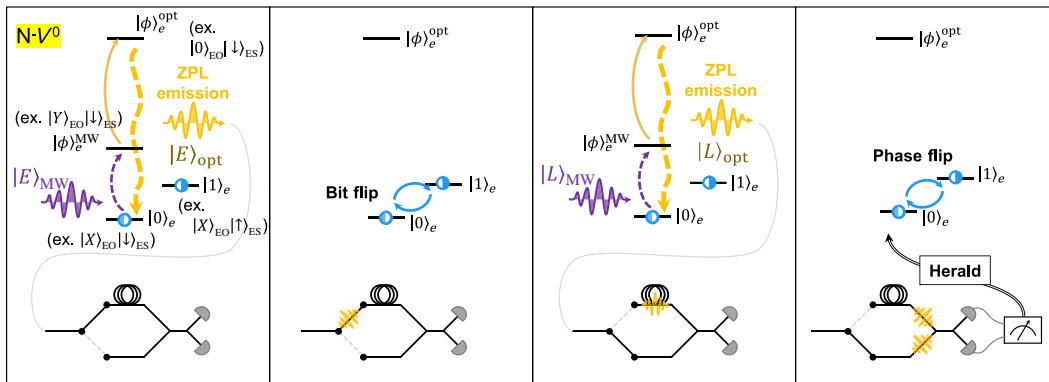

FIG. 13. Schematic of the absorption of a microwave time-bin qubit by an electron spin whose ground state has high strain sensitivity. The energy level of N-V^0 is considered as an example. The optically excited state of N-V^0 is written as $|0\rangle_{\text{ES}}|\downarrow\rangle_{\text{ES}}$ following Ref. [60].

parameters are the same as those used in the simulation of N-V^- except $\omega_e^{\text{N-}V^0} = 5$ GHz. In the calculation, $\gamma_e/2\pi$ is changed in the range of 0.01–10 MHz considering the orbital relaxation of N-V^0 is $T_1^{\text{orb}} \sim 0.4$ μs (4 K) [60], corresponding $\gamma_e/2\pi = 1/2\pi T_1 \sim 0.3$ MHz. Further reduction of T_1^{orb} is expected when we operate N-V^0 in the dilution refrigerator (10–100 mK). Besides, $\gamma_e/2\pi$ is changed in the range of 0.01–10 MHz.

Figure 14(a) shows the time evolution of the population in the quantum interface. For $g_{m\text{-}e}/2\pi$=1 MHz and $\gamma_e/2\pi$=1 MHz, the population in the orbital excited state reaches > 20%, which is sufficient for our entanglement generation protocol. Besides, since the population changes slower than the time needed for the optical excitation (< nanoseconds), the orbital excited state of N-V^0 can be converted to an optical photon as depicted in Fig. 13.

Figure 14(b) shows the maximum population in the orbital excited state of N-V^0 after the time evolution as functions of $g_{m\text{-}e}/2\pi$ and $\gamma_e/2\pi$. Considering that $g_{m\text{-}e}/2\pi$ of the quantum interface is in the range of 0.5–10 MHz and $\gamma_e/2\pi$ is in the range of 0.01–2 MHz, the efficiency is from 1% to nearly 100%. It should be noted that in a high-strain
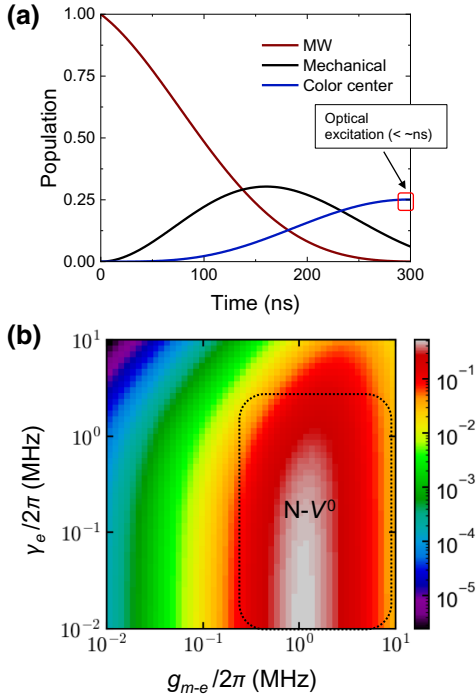
regime of Si-V^-, the phonon absorption rate can decrease to 0.1–0.5 of the rate with zero strain [23]. Thus, it may be desirable to achieve high mechanical-spin coupling of more than 10 MHz using an optimized phononic crystal cavity for N-V^0 in a high-strain regime.

APPENDIX E: EMISSION OF A MICROWAVE PHOTON FROM A COLOR CENTER

Here, we describe a procedure to emit a microwave time-bin qubit from a color center. We take N-V^0 as an example, as with the case shown in Appendix D 2. In the following, we use the same symbols in Appendix D 2.

First, the electron spin is prepared in the state,

$$\frac{1}{\sqrt{2}}(|0\rangle_e + |1\rangle_e). \tag{E1}$$

A strain field generated by the piezoelectric effect using a driving microwave pulse is then applied to the electron spin, exciting $|0\rangle_e$ to $|\phi\rangle_e^{\text{MW}}$. Since $|\phi\rangle_e^{\text{MW}}$ couples to the vibration mode (phonon) of the phononic crystal, $|\phi\rangle_e^{\text{MW}}$ emits the phonon to the phononic crystal, which is the reverse of Fig. 14(a). The microwave phonon is then converted to a microwave photon via the piezoelectric effect. The microwave cavity coupled to the phononic crystal receives the converted microwave photon and emits it to a waveguide as $|E\rangle_{\text{MW}}$. After the application of the π pulse exchanging $|0\rangle_e$ and $|1\rangle_e$, the state at $|0\rangle_e$ is excited to $|\phi\rangle_e^{\text{MW}}$ again, emitting $|L\rangle_{\text{MW}}$ to the phononic crystal. Thus, after the bit flip of $|0\rangle_e$ and $|1\rangle_e$, the resultant microwave photon-spin entangled state is

$$\frac{1}{\sqrt{2}}(|E\rangle_{\text{MW}} |0\rangle_e + |L\rangle_{\text{MW}} |1\rangle_e). \tag{E2}$$

After a superconducting qubit, which is prepared in $|g\rangle_{\text{sc}}$, absorbs time-bin microwave qubit with appropriate gate operations [30], the state becomes

$$\frac{1}{\sqrt{2}}(|e\rangle_{\text{sc}} |0\rangle_e + |g\rangle_{\text{sc}} |1\rangle_e). \tag{E3}$$

FIG. 14. (a) Time evolution of the population in the quantum interface for $g_{m\text{-}e}/2\pi = 1$ MHz and $\gamma_e/2\pi = 1$ MHz. Optical excitation is performed within nanoseconds at the peak of the population. In the legend, MW and mechanical correspond to the microwave resonator and the mechanical resonator, respectively. (b) Maximum population in the orbital excited state of N-V^0 after the time evolution as functions of $g_{m\text{-}e}/2\pi$ and $\gamma_e/2\pi$. An area surrounded by a black dashed line corresponds to typical parameters of N-V^0.

[1] M. Kjaergaard, M. E. Schwartz, J. Braumüller, P. Krantz, J. I.-J. Wang, S. Gustavsson, and W. D. Oliver, Superconducting qubits: Current state of play, Annu. Rev. Condens. Matter Phys. **11**, 369 (2020).

[2] F. Arute *et al.*, Quantum supremacy using a programmable superconducting processor, Nature **574**, 505 (2019).

[3] M. Ahsan, R. V. Meter, and J. Kim, Designing a million-qubit quantum computer using resource performance simulator, ACM J. Emerg. Technol. Comput. Syst. **12**, 1 (2016).

[4] É. Gouzien and N. Sangouard, Factoring 2048-Bit RSA Integers in 177 Days with 13 436 Qubits and a Multimode Memory, Phys. Rev. Lett. **127**, 140503 (2021).

[5] A. Yimsiriwattana and S. J. Lomonaco Jr., Distributed quantum computing: A distributed Shor algorithm, Quantum Information and Computation II **5436**, 360 (2004).

[6] R. Van Meter and S. J. Devitt, The path to scalable distributed quantum computing, Computer **49**, 31 (2016).

[7] D. Cuomo, M. Caleffi, and A. S. Cacciapuoti, Towards a distributed quantum computing ecosystem, IET Quantum Communication **1**, 3 (2020).

[8] N. J. Lambert, A. Rueda, F. Sedlmeir, and H. G. L. Schwefel, Coherent conversion between microwave and optical photons-an overview of physical implementations, Adv. Quantum Technol. **3**, 1900077 (2020).

[9] N. Lauk, N. Sinclair, S. Barzanjeh, J. P. Covey, M. Saffman, M. Spiropulu, and C. Simon, Perspectives on quantum transduction, Quantum Sci. Technol. **5**, 020501 (2020).

[10] C. Zhong, Z. Wang, C. Zou, M. Zhang, X. Han, W. Fu, M. Xu, S. Shankar, M. H. Devoret, H. X. Tang, and L. Jiang, Proposal for Heralded Generation and Detection of Entangled Microwave-Optical-Photon Pairs, Phys. Rev. Lett. **124**, 010511 (2020).

[11] S. Krastanov, H. Raniwala, J. Holzgrafe, K. Jacobs, M. Lončar, M. J. Reagor, and D. R. Englund, Optically Heralded Entanglement of Superconducting Systems in Quantum Networks, Phys. Rev. Lett. **127**, 040503 (2021).

[12] J. Wu, C. Cui, L. Fan, and Q. Zhuang, Deterministic Microwave-Optical Transduction Based on Quantum Teleportation, Phys. Rev. Appl. **16**, 064044 (2021).

[13] T. Yamamoto, M. Koashi, c. K. Özdemir, and N. Imoto, Experimental extraction of an entangled photon pair from two identically decohered pairs, Nature **421**, 343 (2003).

[14] X. Li, J. Chen, P. Voss, J. Sharping, and P. Kumar, All-fiber photon-pair source for quantum communications: Improved generation of correlated photons, Opt. Express **12**, 3737 (2004).

[15] C.-Z. Peng, T. Yang, X.-H. Bao, J. Zhang, X.-M. Jin, F.-Y. Feng, B. Yang, J. Yang, J. Yin, Q. Zhang, N. Li, B.-L. Tian, and J.-W. Pan, Experimental Free-Space Distribution of Entangled Photon Pairs Over 13 km: Towards Satellite-Based Global Quantum Communication, Phys. Rev. Lett. **94**, 150501 (2005).

[16] L.-M. Duan, M. D. Lukin, J. I. Cirac, and P. Zoller, Long-distance quantum communication with atomic ensembles and linear optics, Nature **414**, 413 (2001).

[17] D. L. Moehring, P. Maunz, S. Olmschenk, K. C. Younge, D. N. Matsukevich, L.-M. Duan, and C. Monroe, Entanglement of single-atom quantum bits at a distance, Nature **449**, 68 (2007).

[18] J. Hofmann, M. Krug, N. Ortegel, L. Gérard, M. Weber, W. Rosenfeld, and H. Weinfurter, Heralded entanglement between widely separated atoms, Science **337**, 72 (2012).

[19] H. Bernien, B. Hensen, W. Pfaff, G. Koolstra, M. S. Blok, L. Robledo, T. H. Taminiau, M. Markham, D. J. Twitchen, L. Childress, and R. Hanson, Heralded entanglement between solid-state qubits separated by three metres, Nature **497**, 86 (2013).

[20] A. Delteil, Z. Sun, W.-b. Gao, E. Togan, S. Faelt, and A. Imamoğlu, Generation of heralded entanglement between distant hole spins, Nat. Phys. **12**, 218 (2016).

[21] T. Neuman, M. Eichenfield, M. E. Trusheim, L. Hackett, P. Narang, and D. Englund, A phononic interface between a superconducting quantum processor and quantum networked spin memories, Npj Quantum Inf. **7**, 121 (2021).

[22] K. W. Lee, D. Lee, P. Ovartchaiyapong, J. Minguzzi, J. R. Maze, and A. C. Bleszynski Jayich, Strain Coupling of a Mechanical Resonator to a Single Quantum Emitter in Diamond, Phys. Rev. Appl. **6**, 034005 (2016).

[23] S. Meesala, Y.-I. Sohn, B. Pingault, L. Shao, H. A. Atikian, J. Holzgrafe, M. Gündoğan, C. Stavrakas, A. Sipahigil, C. Chia, R. Evans, M. J. Burek, M. Zhang, L. Wu, J. L. Pacheco, J. Abraham, E. Bielejec, M. D. Lukin, M. Atatüre, and M. Lončar, Strain engineering of the silicon-vacancy center in diamond, Phys. Rev. B **97**, 205444 (2018).

[24] P. K. Shandilya, D. P. Lake, M. J. Mitchell, D. D. Sukachev, and P. E. Barclay, Optomechanical interface between telecom photons and spin quantum memory, Nat. Phys. **17**, 1420 (2021).

[25] H. Raniwala, S. Krastanov, M. Eichenfield, and D. Englund, A spin-optomechanical quantum interface enabled by an ultrasmall mechanical and optical mode volume cavity, http://arxiv.org/abs/2202.06999.

[26] M. K. Schmidt, C. G. Poulton, and M. J. Steel, Acoustic diamond resonators with ultrasmall mode volumes, Phys. Rev. Res. **2**, 033153 (2020).

[27] P. C. Humphreys, N. Kalb, J. P. J. Morits, R. N. Schouten, R. F. L. Vermeulen, D. J. Twitchen, M. Markham, and R. Hanson, Deterministic delivery of remote entanglement on a quantum network, Nature **558**, 268 (2018).

[28] M. Pompili, S. L. N. Hermans, S. Baier, H. K. C. Beukers, P. C. Humphreys, R. N. Schouten, R. F. L. Vermeulen, M. J. Tiggelman, L. dos Santos Martins, B. Dirkse, S. Wehner, and R. Hanson, Realization of a multinode quantum network of remote solid-state qubits, Science **372**, 259 (2021).

[29] D. S. Levonian, R. Riedinger, B. Machielse, E. N. Knall, M. K. Bhaskar, C. M. Knaut, R. Bekenstein, H. Park, M. Lončar, and M. D. Lukin, Optical Entanglement of Distinguishable Quantum Emitters, Phys. Rev. Lett. **128**, 213602 (2022).

[30] P. Kurpiers, M. Pechal, B. Royer, P. Magnard, T. Walter, J. Heinsoo, Y. Salathé, A. Akin, S. Storz, J.-C. Besse, S. Gasparinetti, A. Blais, and A. Wallraff, Quantum Communication with Time-Bin Encoded Microwave Photons, Phys. Rev. Appl. **12**, 044067 (2019).

[31] P. Kurpiers, P. Magnard, T. Walter, B. Royer, M. Pechal, J. Heinsoo, Y. Salathé, A. Akin, S. Storz, J.-C. Besse, S. Gasparinetti, A. Blais, and A. Wallraff, Deterministic quantum state transfer and generation of remote entanglement using microwave photons, Nature **558**, 264 (2017).

[32] T. Douce, M. Stern, N. Zagury, P. Bertet, and P. Milman, Coupling a single nitrogen-vacancy center to a superconducting flux qubit in the far-off-resonance regime, Phys. Rev. A **92**, 052335 (2015).

[33] C.-H. Li and P.-B. Li, Coupling a single nitrogen-vacancy center with a superconducting qubit via the electro-optic effect, Phys. Rev. A **97**, 052319 (2018).

[34] Y.-S. Park, A. K. Cook, and H. Wang, Cavity QED with diamond nanocrystals and silica microspheres, Nano Lett. **6**, 2075 (2006).

[35] M. Larsson, K. N. Dinyari, and H. Wang, Composite optical microcavity of diamond nanopillar and silica microsphere, Nano Lett. **9**, 1447 (2009).

[36] S. Schietinger, T. Schröder, and O. Benson, One-by-one coupling of single defect centers in nanodiamonds to high-Q modes of an optical microresonator, Nano Lett. **8**, 3911 (2008).

[37] P.-B. Li, S.-Y. Gao, and F.-L. Li, Quantum-information transfer with nitrogen-vacancy centers coupled to a whispering-gallery microresonator, Phys. Rev. A **83**, 054306 (2011).

[38] J. Ilves, S. Kono, Y. Sunada, S. Yamazaki, M. Kim, K. Koshino, and Y. Nakamura, On-demand generation and characterization of a microwave time-bin qubit, Npj Quantum Inf. **6**, 34 (2020).

[39] A. Somoroff, Q. Ficheux, R. A. Mencia, H. Xiong, R. Kuzmin, and V. E. Manucharyan, Millisecond coherence in a superconducting qubit, arXiv:2103.08578.

[40] J. H. Shim, I. Niemeyer, J. Zhang, and D. Suter, Room-temperature high-speed nuclear-spin quantum memory in diamond, Phys. Rev. A **87**, 012301 (2013).

[41] S. S. Hegde, J. Zhang, and D. Suter, Efficient Quantum Gates for Individual Nuclear Spin Qubits by Indirect Control, Phys. Rev. Lett. **124**, 220501 (2020).

[42] C. Zu, W.-B. Wang, L. He, W.-G. Zhang, C.-Y. Dai, F. Wang, and L.-M. Duan, Experimental realization of universal geometric quantum gates with solid-state spins, Nature **514**, 72 (2014).

[43] K. Nagata, K. Kuramitani, Y. Sekiguchi, and H. Kosaka, Universal holonomic quantum gates over geometric spin qubits with polarised microwaves, Nat. Commun. **9**, 3227 (2018).

[44] M. K. Bhaskar, R. Riedinger, B. Machielse, D. S. Levonian, C. T. Nguyen, E. N. Knall, H. Park, D. Englund, M. Lončar, D. D. Sukachev, and M. D. Lukin, Experimental demonstration of memory-enhanced quantum communication, Nature **580**, 60 (2020).

[45] C. T. Nguyen, D. D. Sukachev, M. K. Bhaskar, B. Machielse, D. S. Levonian, E. N. Knall, P. Stroganov, C. Chia, M. J. Burek, R. Riedinger, H. Park, M. Lončar, and M. D. Lukin, An integrated nanophotonic quantum register based on silicon-vacancy spins in diamond, Phys. Rev. B **100**, 165428 (2019).

[46] Z. Wang, Y. Wu, Z. Bao, Y. Li, C. Ma, H. Wang, Y. Song, H. Zhang, and L. Duan, Experimental Realization of a Deterministic Quantum Router with Superconducting Quantum Circuits, Phys. Rev. Appl. **15**, 014049 (2021).

[47] P. Kurpiers, T. Walter, P. Magnard, Y. Salathe, and A. Wallraff, Characterizing the attenuation of coaxial and rectangular microwave-frequency waveguides at cryogenic temperatures, EPJ Quantum Technol. **4**, 8 (2017).

[48] M. Ruf, N. H. Wan, H. Choi, D. Englund, and R. Hanson, Quantum networks based on color centers in diamond, J. Appl. Phys. **130**, 070901 (2021).

[49] Y. Sekiguchi, Y. Komura, S. Mishima, T. Tanaka, N. Niikura, and H. Kosaka, Geometric spin echo under zero field, Nat. Commun. **7**, 11668 (2016).

[50] Y. Sekiguchi, N. Niikura, R. Kuroiwa, H. Kano, and H. Kosaka, Optical holonomic single quantum gates with a geometric spin under a zero field, Nat. Photonics **11**, 309 (2017).

[51] N. Samkharadze, A. Bruno, P. Scarlino, G. Zheng, D. P. DiVincenzo, L. DiCarlo, and L. M. K. Vandersypen, High-Kinetic-Inductance Superconducting Nanowire Resonators for Circuit QED in a Magnetic Field, Phys. Rev. Appl. **5**, 044004 (2016).

[52] J. Krause, C. Dickel, E. Vaal, M. Vielmetter, J. Feng, R. Bounds, G. Catelani, J. M. Fink, and Y. Ando, Magnetic Field Resilience of Three-Dimensional Transmons with Thin-Film $Al/AlO_x/Al$ Josephson Junctions Approaching 1 T, Phys. Rev. Appl. **17**, 034032 (2022).

[53] J. Borregaard, A. S. Sørensen, and P. Lodahl, Quantum networks with deterministic spin-photon interfaces, Adv. Quantum Technol. **2**, 1800091 (2019).

[54] F. Marsili, V. B. Verma, J. A. Stern, S. Harrington, A. E. Lita, T. Gerrits, I. Vayshenker, B. Baek, M. D. Shaw, R. P. Mirin, and S. W. Nam, Detecting single infrared photons with 93% system efficiency, Nat. Photonics **7**, 210 (2013).

[55] D. A. Golter, T. Oo, M. Amezcua, K. A. Stewart, and H. Wang, Optomechanical Quantum Control of a Nitrogen-Vacancy Center in Diamond, Phys. Rev. Lett. **116**, 143602 (2016).

[56] D. A. Golter, T. Oo, M. Amezcua, I. Lekavicius, K. A. Stewart, and H. Wang, Coupling a Surface Acoustic Wave to an Electron Spin in Diamond via a Dark State, Phys. Rev. X **6**, 1 (2016).

[57] P. Arrangoiz-Arriola, E. A. Wollack, M. Pechal, J. D. Witmer, J. T. Hill, and A. H. Safavi-Naeini, Coupling a Superconducting Quantum Circuit to a Phononic Crystal Defect Cavity, Phys. Rev. X **8**, 031007 (2018).

[58] M. Mirhosseini, A. Sipahigil, M. Kalaee, and O. Painter, Superconducting qubit to optical photon transduction, Nature **588**, 599 (2020).

[59] A. Sipahigil, R. E. Evans, D. D. Sukachev, M. J. Burek, J. Borregaard, M. K. Bhaskar, C. T. Nguyen, J. L. Pacheco, H. A. Atikian, C. Meuwly, R. M. Camacho, F. Jelezko, E. Bielejec, H. Park, M. Lončar, and M. D. Lukin, An integrated diamond nanophotonics platform for quantum-optical networks, Science **354**, 847 (2016).

[60] S. Baier, C. E. Bradley, T. Middelburg, V. V. Dobrovitski, T. H. Taminiau, and R. Hanson, Orbital and Spin Dynamics of Single Neutrally-Charged Nitrogen-Vacancy Centers in Diamond, Phys. Rev. Lett. **125**, 193601 (2020).

[61] M. S. Barson, E. Krausz, N. B. Manson, and M. W. Doherty, The fine structure of the neutral nitrogen-vacancy center in diamond, Nanophotonics **8**, 1985 (2019).