arXiv:1801.01904v1 [quant-ph] 5 Jan 2018

Phonon networks with SiV centers in diamond waveguides

M.-A. Lemonde[1], S. Meesala[2], A. Sipahigil[3,4], M. J. A. Schuetz[4], M. D. Lukin[4], M. Loncar[2], P. Rabl[1]
[1] *Vienna Center for Quantum Science and Technology, Atominstitut, TU Wien, 1040 Vienna, Austria*
[2] *John A. Paulson School of Engineering and Applied Sciences, Harvard University, 29 Oxford Street, Cambridge, MA 02138, USA*
[3] *Institute for Quantum Information and Matter and Thomas J. Watson, Sr., Laboratory of Applied Physics, California Institute of Technology, Pasadena, California 91125, USA and*
[4] *Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA*
(Dated: January 9, 2018)

We propose and analyze a novel realization of a solid-state quantum network, where separated silicon-vacancy centers are coupled via the phonon modes of a quasi-1D diamond waveguide. In our approach, quantum states encoded in long-lived electronic spin states can be converted into propagating phonon wavepackets and be reabsorbed efficiently by a distant defect center. Our analysis shows that under realistic conditions, this approach enables the implementation of high-fidelity, scalable quantum communication protocols within chip-scale spin-qubit networks. Apart from quantum information processing, this setup constitutes a novel waveguide QED platform, where strong-coupling effects between solid-state defects and individual propagating phonons can be explored at the quantum level.

Electronic and nuclear spins associated with defects in solids comprise a promising platform for the realization of practical quantum technologies [1]. A prominent example is the nitrogen-vacancy (NV) center in diamond [2, 3], for which techniques for state detection [4], coherent manipulations [5–7] and local entanglement operations [8–10] have been demonstrated and employed, for example, for various nanoscale sensing applications [11]. Despite this progress in the local control of spin qubits, integrating many spins into larger networks remains a challenging task. To achieve this goal, several schemes for interfacing spins via mechanical degrees of freedom have recently been discussed [12–17] and first experiments demonstrating magnetic [18–20] or strain-induced [21–25] couplings of mechanical vibrations to both long-lived spin states and electronic excited states of NV centers have been carried out. However, the weak intrinsic coupling of spins to vibrational modes and the short coherence of optically excited states make the extension of these methods into the quantum regime challenging.

In this Letter we describe the implementation of a phonon quantum network, where negatively-charged silicon-vacancy (SiV) centers are coupled via propagating phonon modes of a 1D diamond waveguide [26–29]. The electronic ground state of the SiV center features both spin and orbital degrees of freedom [30–32], which makes it naturally suited for this task; quantum states can be encoded in long-lived superpositions of the two lowest spin-orbit-coupled states [33–37], while a controlled admixing of higher orbital states, which are susceptible to strain, gives rise to a strong and tunable coupling to phonons. The central phonon frequency of ~ 46 GHz set by the large spin-orbit splitting enables quantum-coherent operations already at convenient temperatures of $T \lesssim 1$ K, when thermal excitations at this frequency are frozen out. Our analysis shows that high-fidelity quantum state transfer protocols between distant

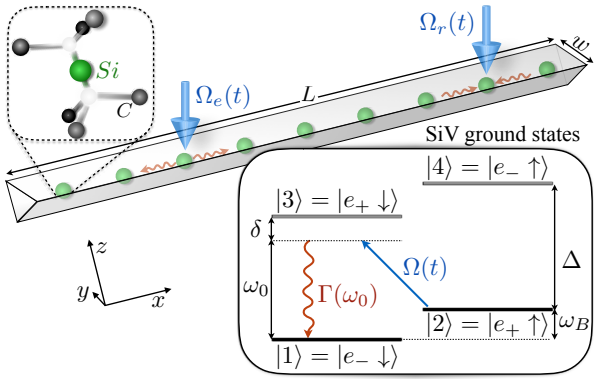

FIG. 1: Setup. An array of SiV defects is embedded in a 1D phonon waveguide. The inset shows the level structure of the electronic ground state of the SiV center. A tunable Raman process involving the excited state $|3\rangle$ is used to coherently convert the population of the stable state $|2\rangle$ into a propagating phonon, which can be reabsorbed by any other selected center along the waveguide. See text for more details.

SiV centers can be implemented under realistic conditions. Moreover, we propose a scalable operation of such phonon networks using switchable single-defect mirrors.

Model.—We consider a system as depicted in Fig. 1, where an array of SiV centers is embedded in a 1D diamond waveguide. The electronic ground state of the SiV center is formed by an unpaired hole of spin $S = 1/2$, which occupies one of the two degenerate orbital states $|e_x\rangle$ and $|e_y\rangle$. In the presence of spin-orbit interactions and a weak Jahn-Teller effect, the four states are split into two doublets, $\{|1\rangle \simeq |e_-,\downarrow\rangle, |2\rangle \simeq |e_+,\uparrow\rangle\}$ and $\{|3\rangle \simeq |e_+,\downarrow\rangle, |4\rangle \simeq |e_-,\uparrow\rangle\}$, which are separated by $\Delta/2\pi \simeq 46$ GHz [31, 32]. Here, $|e_\pm\rangle = (|e_x\rangle \pm i|e_y\rangle)/\sqrt{2}$ are eigenstates of the orbital angular momentum operator, i.e. $L_z|e_\pm\rangle = \pm\hbar|e_\pm\rangle$, where the z-axis is along the symmetry axis of the defect. In the presence of a mag-

netic field $\vec{B} = B_0\vec{e}_z$, the Hamilton operator for a single SiV center is ($\hbar = 1$)

$$H_{\rm SiV} = \omega_B|2\rangle\langle 2| + \Delta|3\rangle\langle 3| + (\Delta+\omega_B)|4\rangle\langle 4| + \frac{1}{2}\left[\Omega(t)e^{i[\omega_d t+\theta(t)]}\left(|2\rangle\langle 3| + |1\rangle\langle 4|\right) + \text{H.c.}\right], \quad (1)$$

where $\omega_B = \gamma_s B_0$ and γ_s is the spin gyromagnetic ratio. In Eq. (1), we have included a time-dependent driving field with a tunable Rabi-frequency $\Omega(t)$ and phase $\theta(t)$, which couples the lower and upper states of opposite spin. This coupling can be implemented either directly with a microwave field of frequency $\omega_d \sim \Delta$ [38], or indirectly via an equivalent optical Raman process [39]. The latter method is already used in experiments to initialize and prepare individual SiV centers in superpositions of $|1\rangle$ and $|2\rangle$ [33–35] with coherence times that can exceed 10 ms in the absence of thermal processes and with dynamical decoupling [36]. Further details on the derivation of $H_{\rm SiV}$ are given in the supplementary material [39].

For the waveguide, we consider a quasi-1D geometry of width w and length $L \gg w$. The waveguide supports travelling phonon modes of frequency $\omega_{n,k}$ and mode function $\vec{u}_{n,k}(\vec{r}) \sim \vec{u}^{\perp}_{n,k}(y,z)e^{ikx}$, where k is the wavevector along the waveguide direction, n is the branch index and $\vec{u}^{\perp}_{n,k}(y,z)$ is the transverse profile of the displacement field. The phonons induce transitions between the orbital states $|e_\pm\rangle$ [40–42], and the Hamiltonian for the whole system reads

$$H = \sum_j H^{(j)}_{\rm SiV} + \sum_{n,k}\omega_{n,k}a^\dagger_{n,k}a_{n,k} + \frac{1}{\sqrt{L}}\sum_{j,k,n}\left(g^j_{n,k}J^j_+ a_{n,k}e^{ikx_j} + \text{H.c.}\right). \quad (2)$$

Here j labels the SiV centers located at positions $\vec{r}_j = (x_j, y_j, z_j)$, $J_- = (J_+)^\dagger = |1\rangle\langle 3| + |2\rangle\langle 4|$ is the spin-conserving lowering operator and $a_{n,k}$ ($a^\dagger_{n,k}$) are the annihilation (creation) operators for the phonon modes. The couplings $g^j_{n,k} \equiv g_{n,k}(y_j, z_j)$ depend on the components of the local strain tensor, $\epsilon^{ab}_{n,k}(\vec{r}_j) = \frac{1}{2}[\frac{\partial}{\partial x_b}u^a_{n,k}(\vec{r}_j) + \frac{\partial}{\partial x_a}u^b_{n,k}(\vec{r}_j)]$, and can be evaluated for a known transverse mode profile $\vec{u}^{\perp}_n(y,z)$ [39, 42]. We express the resulting couplings as

$$g^j_{n,k} = d\sqrt{\frac{\hbar k^2}{2\rho A\omega_{n,k}}}\xi_{n,k}(y_j, z_j), \quad (3)$$

where $d/2\pi \sim 1$ PHz is the strain sensitivity of the orbital states [40, 41], ρ the density and A the transverse area of the waveguide. The dimensionless coupling profile $\xi_{n,k}(y,z)$ accounts for the specific strain distribution and $\xi(y,z) = 1$ for a homogeneous compression mode.

From cavity to waveguide QED.—For small structures, $L \sim 10 - 100\,\mu$m, $w \lesssim 200$ nm, and group velocities $v \sim 10^4$ m/s, the individual phonon modes are well separated in frequency, $\Delta\omega/2\pi \gtrsim 50$ MHz, and the SiV centers can be coupled to a single standing-wave mode with a strength $g_L = g_0\sqrt{\lambda/L} \approx 2\pi \times (4-14)$ MHz, where $g_0/2\pi \approx 105$ MHz and $\lambda \approx 200$ nm is the phonon wavelength. The system dynamics is then governed by a Jaynes-Cummings-type interaction between phonons and orbital states [39]. In the strong coupling regime, $g_L > \kappa = \Delta/Q$, which is reached for moderate mechanical quality factors of $Q > 10^4$, a coherent exchange of phonons and defect excitations becomes possible. For longer waveguides, the coupling to the quasi-continuum of phonon modes is characterized by the resulting decay rate $\Gamma_j(\Delta) = \sum_n \Gamma_{j,n}(\Delta)$ for states $|3\rangle$ and $|4\rangle$, where

$$\Gamma_{j,n}(\omega) = \lim_{L\to\infty}\frac{2\pi}{L}\sum_k |g^j_{n,k}|^2\delta(\omega - \omega_{n,k}). \quad (4)$$

For a single compression mode with $\vec{u}^{\perp}(y,z) \sim \vec{x}$ and a linear dispersion $\omega_k = vk$, we obtain $\Gamma(\omega) = d^2\hbar\omega/(\rho A v^3)$, which results in a characteristic phonon emission rate of $\Gamma(\Delta)/2\pi \sim 1$ MHz [42].

Figure 2 summarizes the simulated acoustic dispersion relations and the resulting decay rates for a triangular

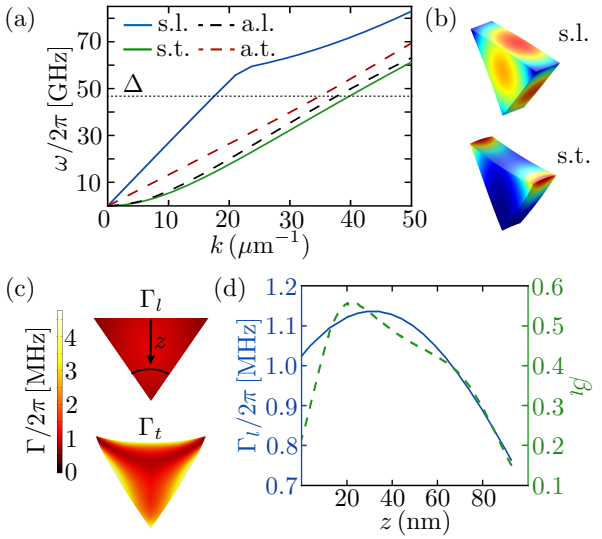

FIG. 2: Phonon waveguide. (a) Acoustic dispersion relation for a triangular waveguide of width $w = 130$ nm and etch-angle $\varphi = 35^o$. Symmetric (solid lines) and anti-symmetric (dashed lines) branches with respect to the vertical mirror-symmetry plane are shown. (b) Normalized displacement profiles of the symmetric phonons at 46 GHz. (c) The emission rates into the symmetric longitudinal (Γ_l) and transverse (Γ_t) polarization modes at 46 GHz are plotted for different positions of the SiV center within the triangular cross-section. (d) Γ_l and fraction (β_l) of spontaneous emission into the longitudinal branch for different positions of the SiV center along the vertical mirror-symmetry axis. For all results, an orientation of the waveguide along the [110] crystal axis of diamond and SiV centers oriented along $[\bar{1}11]$ and $[1\bar{1}1]$, i.e., orthogonal to the waveguide axis, have been assumed.

waveguide [26, 40] of width $w = 130$ nm. The SiV centers couple primarily to a longitudinal (l) compression and a transverse (t) flexural mode with group velocities $v_l = 1.71 \times 10^4$ m/s and $v_t = 0.73 \times 10^4$ m/s, respectively. The coupling to the other two branches of odd symmetry can be neglected for defects near the center of the waveguide. Fig. 2(c) and (d) show that the rates $\Gamma_{l,t}$ are quite insensitive to the exact location of the SiV center. However, the fraction of phonons emitted into a specific branch, $\beta_n = \Gamma_n/\Gamma$, is significantly below unity as emission is split between a pair of modes. In optical waveguides [43], a value of $\beta < 1$ usually arises from the emission of photons into non-guided modes, which are irreversibly lost. For a phonon waveguide this is not the case, but the multi-branch nature of the waveguide must be fully taken into account. In all examples below we assume $\beta_l = \beta_t = 0.5$, which is most relevant for SiV defects located near the center of the beam.

Coherent spin-phonon interface.—We are interested in the transfer of a qubit state, encoded into the stable states $|1\rangle$ and $|2\rangle$, between an arbitrary pair of emitting (e) and receiving (r) defects in the waveguide,

$$(\alpha|1\rangle_e + \beta|2\rangle_e)|1\rangle_r \to |1\rangle_e(\alpha|1\rangle_r + \beta|2\rangle_r). \qquad (5)$$

As shown in Fig. 1, this can be achieved by inducing a Raman transition via state $|3\rangle_e$ to convert the population in state $|2\rangle_e$ into a propagating phonon and by reverting the process at the receiving center. For low enough temperatures, $T \ll \hbar\Delta/k_B \approx 2.2$ K, such that all phonon modes are initially in the vacuum state, this scenario is described by the following ansatz for the wavefunction $|\psi(t)\rangle = [\alpha\mathbb{1} + \beta C^\dagger(t)]|\bar{1}, 0\rangle$, where $|\bar{1}, 0\rangle$ is the ground state with all SiV centers in state $|1\rangle$ and $C^\dagger(t) = \sum_{j=e,r}\left[c_j(t)e^{-i\omega_B t}|2\rangle_j\langle 1| + b_j(t)e^{-i\omega_0 t}|3\rangle_j\langle 1| + \sum_{n,k} c_{n,k}(t)e^{-i\omega_0 t}a^\dagger_{n,k}\right]$ creates a single excitation distributed between the SiV centers and the phonon modes. The central phonon frequency $\omega_0 = \Delta_j + \delta_j$ is assumed to be fixed by compensating small inhomogeneities in the Δ_j by the detunings $\delta_j = \omega_d^j - (\Delta_j - \omega_B^j)$.

By adiabatically eliminating the fast decaying amplitudes b_j, we derive effective equations of motion for the slowly varying amplitudes $c_i(t)$. From this derivation, detailed in [39], we obtain for each qubit amplitude

$$\dot{c}_j(t) = -\frac{\gamma_j(t)}{2}c_j(t) - \sum_n \sqrt{\frac{\gamma_{j,n}(t)}{2}}e^{-i\theta_j(t)}\Phi^{\rm in}_{j,n}(t), \quad (6)$$

where $\gamma_j(t) = \sum_n \gamma_{j,n}(t)$ is the effective decay rate of state $|2\rangle_j$ and

$$\gamma_{j,n}(t) = \frac{\Omega_j^2(t)}{4\delta_j^2 + \Gamma_j^2(\omega_0)}\Gamma_{j,n}(\omega_0). \qquad (7)$$

Assuming $0 \le \Omega(t)/2\pi < 70$ MHz and $\delta/2\pi = 100$ MHz, this rate can be tuned between $\gamma_j = 0$ and a maximal value of $\gamma_{\max}/2\pi \approx 250$ kHz, which is still fast compared to the expected bare dephasing times $T_2^* = 10 - 100\,\mu$s of the qubit state [36]. At the same time, the large detuning $\delta \gg \Gamma(\Delta)$ ensures that any residual scattering of phonons from an undriven defect is strongly suppressed [39].

The last term in Eq. (6), where $\Phi^{\rm in}_{j,n} = \Phi^{\rm in,L}_{j,n} + \Phi^{\rm in,R}_{j,n}$, describes the coupling of an SiV center to the left- (L) and right- (R) incoming fields $\Phi^{\rm in,R/L}_{j,n}$, which themselves are related to the corresponding outgoing fields by [44]

$$\Phi^{\rm out,R/L}_{j,n}(t) = \Phi^{\rm in,R/L}_{j,n}(t) + \sqrt{\frac{\gamma_{j,n}(t)}{2}}c_j(t)e^{i\theta_j(t)}. \qquad (8)$$

Together with Eq. (6), these input-output relations specify the local dynamics at each node and must be supplemented by a set of propagation relations for all fields [cf. Fig. 3(a)]. As an example, for $x_r > x_e$, the right propagating fields obey $\Phi^{\rm in,R}_{r,n}(t) = \Phi^{\rm out,R}_{e,n}(t - \tau^n_{er})e^{i\phi^n_{er}}$, where $\tau^n_{er} = (x_r - x_e)/v_n$ and $\phi^n_{er} = k_n(x_r - x_e)$ are the respective propagation times and phases. Reflections at the boundaries lead to a retarded interaction of each center with its own emitted field. For example, $\Phi^{\rm in,R}_{e,n}(t) = -\sqrt{R_n}\Phi^{\rm out,L}_{e,n}(t - \tau^n_e)e^{i\phi^n_e}$, where $\tau^n_e = 2x_e/v_n$ and $\phi^n_e = 2k_n x_e$, and the reflectivity $R_n \le 1$ has been introduced to model losses. The combined set of time-nonlocal equations for the SiV amplitudes can be solved numerically for given positions x_j and pulses $\gamma_{j,n}(t)$. Since any deterministic phase acquired during the protocol can be undone by a local qubit rotation, we identify $\mathcal{F}(t) = |c_r(t)|^2$ with the fidelity of the transfer, which exceeds the classical bound for $\mathcal{F} > 2/3$ [45].

Quantum state transfer.—In Fig. 3(b) we first consider constant rates $\gamma_{j,n}(t) = \gamma_{\max}/2$, in which case a state transfer is achieved over multiple round-trips of the emitted wave-packet. For $L \sim 100\,\mu$m, the round-trip times $2L/v_n$ are still short compared to $\gamma^{-1}_{\max}$ and we recover the standing-wave picture with splittings $\Delta\omega_n = \pi v_n/L$ between consecutive k-modes. When only the transverse mode is resonant, [i.e., $\phi^t_L = \phi^t_e + \phi^t_r + 2\phi^t_{er} = 2\pi n$, while $\phi^l_L = (2m+1)\pi$] and for maximal coupling [$\phi^l_e = \phi^l_r = (2m+1)\pi$], we observe damped oscillations with a fast frequency $\tilde{g} = \sqrt{\gamma_{\max}\Delta\omega_t/2\pi} \approx 2\pi \times 1.2$ MHz and decay rate $\kappa = -\frac{\Delta\omega}{\pi}\log R \approx 2\pi \times 0.93$ MHz. This result is expected from a single-mode description of the waveguide [39], and is recovered here as a limiting case of our general framework. The losses from multiple imperfect reflections at the boundaries can be partially suppressed at the expense of a slightly slower transfer by detuning the SiV centers from the closest mode by $\delta_0 > \tilde{g}$. In this case the SiV centers communicate via an exchange of virtual phonons and $\kappa \to \kappa(\tilde{g}/\delta_0)^2$. For a maximal detuning $\delta_0 = \Delta\omega_t/2$, the transfer fidelity scales approximately as $\mathcal{F} \simeq R - \pi^2/(8T_2^*\gamma_{\max})$ [39]. For $T_2^* \approx 100\mu$s and $R > 0.99$, which can be achieved, for example, by phononic Bragg mirrors [46], gate fidelities of $\mathcal{F} \gtrsim 0.99$ are possible.

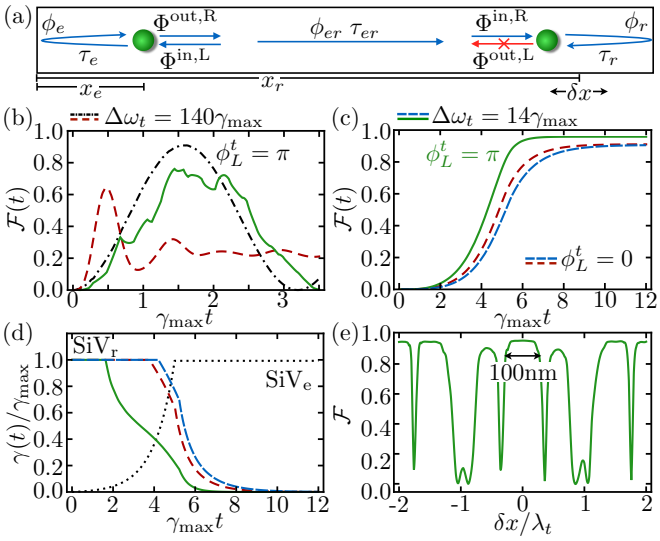

FIG. 3: State transfer protocol. (a) Schematics showing the relevant fields, retardation times and propagation phases. (b) State transfer fidelity for constant rates $\gamma_e(t) = \gamma_r(t) = \gamma_{\max}$. The case of a single resonant mode (red dashed line; $\phi_L^t = 0$, $\phi_L^l = \pi$) is compared to the off-resonant case (dot-dashed black line; $\phi_L^t = \phi_L^l = \pi$) for $L \sim 100\,\mu$m ($\Delta\omega_t/\gamma_{\max} = 140$). The full green line represents the long-waveguide counterpart of the off-resonant scenario, where $L \sim 1$ mm ($\Delta\omega/\gamma_{\max} = 14$). (c)-(d) Protocol using slowly-varying control pulses ($t_p\gamma_{\max} = 1$) where $\Phi_{r,t}^{\text{out,L}}(t)$ is completely suppressed. The dashed blue line corresponds to the long waveguide counter part of the dashed red line. For (b)-(d), the two defects are equally coupled to both modes, $\phi_e^n = \phi_r^n = \pi$ and $\beta_e^n = \beta_r^n = 0.5$. (e) Plot of the state transfer fidelity for varying positions of the receiving SiV center. For this plot $\phi_L^t = \phi_L^l = \pi$ and a maximal transfer time of $12\gamma_{\max}^{-1}$ have been assumed. In all plots, we considered defects near the boundaries where $\tau_e = \tau_r \approx 0$. To illustrate the effect of phonon losses a boundary reflectivity of $R = 0.92$ has been assumed, which corresponds to $Q \approx 5\times10^4$ in the cavity limit.

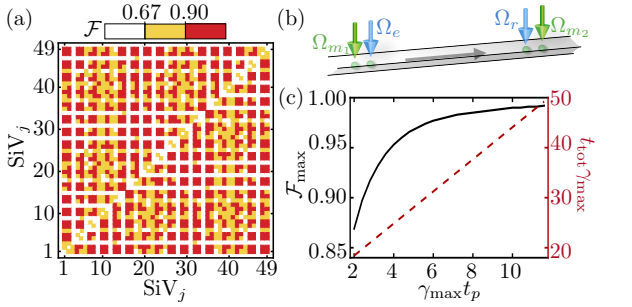

FIG. 4: Scalability. (a) Quantum connectivity matrix for 49 SiV centers equally spaced in a 500 μm long waveguide. The maximal protocol time is fixed to $12\gamma_{\max}^{-1}$ and $R = 0.92$. (b) State-transfer protocol in an infinite waveguide where the outermost SiV centers act as switchable mirrors. The full black line shows the fidelity as a function of the characteristic time of the emitting pulse, t_p, while the dashed red curve shows the total protocol time, t_{tot}. For the considered positions the propagation phases between each center and their neighboring mirror defect is π for both phonon branches and $\Delta\phi_{er} = 2\pi \times n$. In all simulations, a constant decay rate for the mirror defects, $\gamma_{m_1}(t) = \gamma_{m_2}(t) = \gamma_{\max}$, maximizes the fidelity.

As illustrated by the solid line in Fig. 3(b), the simple cavity picture fails for longer waveguides, where multimode and propagation effects become non-negligible. In Fig. 3(c) we illustrate a more general and more robust protocol, where the phonons ideally travel the waveguide only once. Here, the emission is gradually turned on with a fixed pulse $\gamma_e(t)/\gamma_{\max} = \min\{1, e^{(t-5t_p)/t_p}\}$, while $\gamma_r(t)$ and $\theta_r(t)$ are constructed numerically by minimizing at every time step the back-reflected transverse field $|\Phi_{r,t}^{\text{out,L}}|$. For slow pulses, $\gamma_{\max}t_p \gg 1$, a perfect destructive interference between the field reflected from the boundary and the field emitted by the receiving center can be achieved, i.e., $\Phi_{r,t}^{\text{in,L}}(t) + \sqrt{\gamma_{r,t}(t)/2}c_r(t)e^{i\theta_r(t)} = 0$. For a single branch ($\beta_t = 1$) this results in a complete suppression of the signal traveling back to the emitting center so that for $R = 1$ and negligible retardation effects, a perfect state transfer can be implemented [13, 47–49]. Fig. 3(c) shows that this approach also leads to high transfer fidelities under more general conditions, where all propagation effects are taken into account and multiple independent channels participate in the transfer. Importantly, since there are no resonances building up, this strategy is independent of L and can be applied for short and long waveguides equally well.

In the examples shown in Fig. 3(b)-(d), the SiV centers are placed at positions near the ends of the waveguide, where the effective emission rate $\tilde{\gamma}_{j,n}(t) = 2\gamma_{j,n}(t)\sin^2(\phi_j^n/2)$ [50] into both modes is maximal. Fig. 3(e) shows the achievable transfer fidelities when the position of the receiving center is varied over several wavelengths. We observe plateaus of high fidelity extending over $\sim$ 100 nm, interrupted by a few sharp dips arising from a complete destructive interference, i.e. $\phi_r \approx \pi$. This position insensitivity, even in a multichannel scenario, can be understood from a more detailed inspection of the outgoing fields $\Phi_{r,l}^{\text{out,L}}$ [39] and makes the transfer protocol consistent with uncertainties of $\delta x < 50$ nm achieved with state-of-the-art implantation techniques [51].

Scalability.—In Fig. 4 (a), we consider a waveguide of length $L = 500\,\mu$m containing 49 SiV centers spaced by $\Delta x = 10\,\mu$m to allow individual addressing by optical or microwave fields. The resulting quantum connectivity matrix, i.e. the achievable state transfer fidelity between each pair, shows that apart from a few exceptions, most centers can be connected efficiently and that in principle, the operation of large scale networks is possible. By using phononic bandstructure engineering [46, 52], single mode [53] or chiral phononic waveguides [54], the transfer fidelities can be further increased beyond the basic scenario considered here. In practice, propagation losses and elastic phonon scattering will set additional

limitations for the overall size of the network. In Fig. 4 (b), we show a general strategy to overcome these limitations by separating the whole waveguide into smaller segments using additional 'mirror centers'. Here the two outermost SiV centers simply reflect the incoming phonon wavepacket [55], and thus create an effective cavity within the waveguide [56, 57]. This is illustrated in Fig. 4 (b), where we plot the resulting state transfer fidelity for two SiV centers localized inside this effective cavity. For transfer pulses that are long compared to $\gamma_{\rm max}^{-1}$, the outmost centers act as almost perfect mirrors, such that even in an infinite waveguide state transfer protocols within reconfigurable sections of the network can be implemented.

Conclusion.—We have shown how an efficient coupling between individual SiV centers and propagating phonons in a diamond waveguide can be realized and used for quantum networking applications. By employing direct spin-phonon couplings in the presence of a transverse magnetic field [58] or defect-phonon interactions in other materials [59–61], many of the described techniques could also be adapted for lower phonon frequencies $\sim 5-10$ GHz, where many advanced phononic engineering methods are already available. When combined with local control operations involving adjacent nuclear spins as quantum memories [10, 33, 62, 63], the set of all these techniques provides a realistic approach for a scalable quantum information processing platform with spins in solids.

Acknowledgements.—This work was supported by the Austrian Science Fund (FWF) through the SFB FoQuS, Grant No. F40, the START Grant No. Y 591-N16, ONR MURI on Quantum Optomechanics (Award No. N00014-15-1-2761), STC Center for Integrated Quantum Materials (NSF Grant No. DMR-1231319), NSF EFRI ACQUIRE (Award No. 5710004174), CUA, NSF and Vannevar Bush Fellowship.

[1] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O'Brien, Quantum computers, Nature (London) **464**, 7285 (2010).

[2] M. W. Doherty, N. B. Manson, P. Delaney, F. Jelezko, J. Wrachtrup, and L. C. L. Hollenberg, The nitrogen-vacancy colour centre in diamond, Phys. Rep. **528**, 1 (2013).

[3] L. Childress, R. Walsworth, and M. D. Lukin, Atom-like crystal defects, Phys. Today **67**, 38 (2014).

[4] F. Jelezko, T. Gaebel, I. Popa, A. Gruber, and J. Wrachtrup, Observation of coherent oscillations in a single electron spin, Phys. Rev. Lett. **92**, 076401 (2004).

[5] L. Childress, M.V. Gurudev Dutt, J. M. Taylor, A. S. Zibrov, F. Jelezko, J. Wrachtrup, P. R. Hemmer, and M. D. Lukin, Coherent Dynamics of Coupled Electron and Nuclear Spin Qubits in Diamond, Science **314**, 5797 (2006).

[6] G. de Lange, Z. H. Wang, D. Riste, V. V. Dobrovitski, and R. Hanson, Universal Dynamical Decoupling of a Single Solid-State Spin from a Spin Bath, Science **330**, 60 (2010).

[7] L. C. Bassett, F. J. Heremans, D. J. Christle, C. G. Yale, G. Burkard, B. B. Buckley, and D. D. Awschalom, Ultrafast optical control of orbital and spin dynamics in a solid-state defect, Science **345**, 6202 (2014).

[8] W. Pfaff, T. H. Taminiau, L. Robledo, H. Bernien, M. Markham, D. J. Twitchen, and R. Hanson, Demonstration of entanglement-by-measurement of solid-state qubits, Nat. Phys. **9**, 29 (2013).

[9] F. Dolde, I. Jakobi, B. Naydenov, N. Zhao, S. Pezzagna, C. Trautmann, J. Meijer, P. Neumann, F. Jelezko, and J. Wrachtrup, Room-temperature entanglement between single defect spins in diamond, Nat. Phys. **9**, 139 (2013).

[10] G. Waldherr, Y. Wang, S. Zaiser, M. Jamali, T. Schulte-Herbrüggen, H. Abe, T. Ohshima, J. Isoya, J. F. Du, P. Neumann, and J. Wrachtrup, Quantum error correction in a solid-state hybrid spin register, Nature (London) **506**, 204 (2014).

[11] L. Rondin, J.-P. Tetienne, T. Hingant, J.-F. Roch, P. Maletinsky, and V. Jacques, Magnetometry with nitrogen-vacancy defects in diamond, Rep. Prog. Phys. **77**, 056503 (2014).

[12] P. Rabl, S. J. Kolkowitz, F. H. Koppens, J. G. E. Harris, P. Zoller, and M. D. Lukin, A quantum spin transducer based on nanoelectromechanical resonator arrays, Nat. Phys. **6**, 602 (2010).

[13] S. J. M. Habraken, K. Stannigel, M. D. Lukin, P. Zoller, and P. Rabl, Continuous mode cooling and phonon routers for phononic quantum networks, New J. Phys. **14**, 115004 (2012).

[14] S. D. Bennett, N. Y. Yao, J. Otterbach, P. Zoller, P. Rabl, and M. D. Lukin, Phonon-Induced Spin-Spin Interactions in Diamond Nanostructures: Application to Spin Squeezing, Phys. Rev. Lett., **110**, 156402 (2013).

[15] A. Albrecht, A. Retzker, F. Jelezko, and M. B. Plenio, Coupling of nitrogen vacancy centres in nanodiamonds by means of phonons, New J. Phys. **15**, 083014 (2013).

[16] M. J. A. Schuetz, E. M. Kessler, G. Giedke, L. M. K. Vandersypen, M. D. Lukin, and J. I. Cirac, Universal Quantum Transducers Based on Surface Acoustic Waves, Phys. Rev. X **5**, 031031 (2015).

[17] D. Lee, K. W. Lee, J. V. Cady, P. Ovartchaiyapong and A. C. Bleszynski Jayich, Topical review: spins and mechanics in diamond, J. Opt. **19**, 033001 (2017).

[18] D. Rugar, R. Budakian, H. J. Mamin, and B. W. Chui, Single spin detection by magnetic resonance force microscopy, Nature (London) **430**, 329 (2004).

[19] O. Arcizet, V. Jacques, A. Siria, P. Poncharal, P. Vincent, and S. Seidelin, A single nitrogen-vacancy defect coupled to a nanomechanical oscillator, Nat. Phys. **7**, 879 (2011).

[20] S. Kolkowitz, A. C. Bleszynski Jayich, Q. P. Unterreithmeier, S. D. Bennett, P. Rabl, J. G. Harris, and M. D. Lukin, Coherent Sensing of a Mechanical Resonator with a Single-Spin Qubit, Science **335**, 1603 (2012).

[21] E. R. MacQuarrie, T. A. Gosavi, N. R. Jungwirth, S. A. Bhave, and G. D. Fuchs, Mechanical Spin Control of Nitrogen-Vacancy Centers in Diamond, Phys. Rev. Lett. **111**, 227602 (2013).

[22] P. Ovartchaiyapong, K. W. Lee, B. A. Myers, and A. C. Bleszynski Jayich, Dynamic strain-mediated coupling of a single diamond spin to a mechanical resonator, Nat.

Commun. **5**, 4429 (2014).
[23] A. Barfuss, J. Teissier, E. Neu, A. Nunnenkamp, and P. Maletinsky, Strong mechanical driving of a single electron spin, Nat. Phys. **11**, 820 (2015).
[24] 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).
[25] S. Meesala, Y.-I. Sohn, H. A. Atikian, S. Kim, M. J. Burek, J. T. Choy, and M. Loncar, Enhanced Strain Coupling of Nitrogen-Vacancy Spins to Nanoscale Diamond Cantilevers, Phys. Rev. Appl. **5**, 034010 (2016).
[26] M. J. Burek, N. P. de Leon, B. J. Shields, B. J. M. Hausmann, Y. Chu, Q. Quan, A. S. Zibrov, H. Park, M. D. Lukin, and M. Loncar, Free-Standing Mechanical and Photonic Nanostructures in Single-Crystal Diamond, Nano Lett. **12**, 6084 (2012).
[27] B. Khanaliloo, H. Jayakumar, A. C. Hryciw, D. P. Lake, H. Kaviani, and P. E. Barclay, Single-Crystal Diamond Nanobeam Waveguide Optomechanics, Phys. Rev. X **5**, 041051 (2015).
[28] 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. Loncar, and M. D. Lukin, An integrated diamond nanophotonics platform for quantum-optical networks, Science **354**, 847 (2016).
[29] S. Mouradian, N. H. Wan, T. Schröder, and D. Englund, Rectangular photonic crystal nanobeam cavities in bulk diamond, Appl. Phys. Lett. **111**, 021103 (2017).
[30] J. P. Goss, R. Jones, S. J. Breuer, P. R. Briddon, and S. Öberg, The Twelve-Line 1.682 eV Luminescence Center in Diamond and the Vacancy-Silicon Complex, Phys. Rev. Lett. **77**, 3041 (1996).
[31] C. Hepp, T. Müller, V. Waselowski, J. N. Becker, B. Pingault, H. Sternschulte, D. Steinmüller-Nethl, A. Gali, J. R. Maze, M. Atatüre, and C. Becher, Electronic Structure of the Silicon Vacancy Color Center in Diamond, Phys. Rev. Lett. **112**, 036405 (2014).
[32] C. Hepp, Electronic Structure of the Silicon Vacancy Color Center in Diamond, Ph.D. thesis, University of Saarland (2014).
[33] L. J. Rogers, K. D. Jahnke, M. H. Metsch, A. Sipahigil, J. M. Binder, T. Teraji, H. Sumiya, J. Isoya, M. D. Lukin, P. Hemmer, and F. Jelezko, All-Optical Initialization, Readout, and Coherent Preparation of Single Silicon-Vacancy Spins in Diamond, Phys. Rev. Lett. **113**, 263602 (2014).
[34] J. N. Becker, J. Görlitz, C. Arend, M. Markham, and C. Becher, Ultrafast all-optical coherent control of single silicon vacancy colour centres in diamond, Nat. Commun. **7**, 13512 (2016).
[35] Y. Zhou, A. Rasmita, K. Li, Q. Xiong, I. Aharonovich, and W.-B. Gao, Coherent control of a strongly driven silicon vacancy optical transition in diamond, Nat. Commun. **8**, 14451 (2017).
[36] D. D. Sukachev, A. Sipahigil, C. T. Nguyen, M. K. Bhaskar, R. E. Evans, F. Jelezko, and M. D. Lukin, The silicon-vacancy spin qubit in diamond: quantum memory exceeding ten milliseconds and single-shot state readout, Phys. Rev. Lett. **119**, 223602 (2017).
[37] J. N. Becker and C. Becher, Coherence properties and quantum control of silicon vacancy color centers in diamond, Phys. Status Solidi A **214**, 1700586 (2017).
[38] J. J. Pla, K. Y. Tan, J. P. Dehollain, W. H. Lim, J. J. L. Morton, D. N. Jamieson, A. S. Dzurak, and A. Morello, A single-atom electron spin qubit in silicon, Nature **489**, 541 (2012).
[39] See Supplemental Material for additional details about the model and the state transfer protocols.
[40] Y.-I. Sohn, S. Meesala, B. Pingault, H. A. Atikian, J. Holzgrafe, M. Gundogan, C. Stavrakas, M. J. Stanley, A. Sipahigil, J. Choi, M. Zhang, J. L. Pacheco, J. Abraham, E. Bielejec, M. D. Lukin, M. Atatüre, and M. Loncar, Engineering a diamond spin-qubit with a nano-electromechanical system, arXiv:1706.03881.
[41] K. D. Jahnke, A. Sipahigil, J. M. Binder, M. W. Doherty, M. Metsch, L. J. Rogers, N. B. Manson, M. D. Lukin and F. Jelezko, Electron-phonon processes of the silicon-vacancy centre in diamond, New J. Phys. **17**, 043011 (2015).
[42] K. V. Kepesidis, M.-A. Lemonde, A. Norambuena, J. R. Maze, and P. Rabl, Cooling phonons with phonons: Acoustic reservoir engineering with silicon-vacancy centers in diamond, Phys. Rev. B **94**, 214115 (2016).
[43] P. Lodahl, S. Mahmoodian, and S. Stobbe, Interfacing single photons and single quantum dots with photonic nanostructures, Rev. Mod. Phys. **87**, 347 (2015).
[44] C. W. Gardiner and P. Zoller, *Quantum noise* (Springer, Berlin; New York, 2000).
[45] S. Massar and S. Popescu, Optimal extraction of information from finite quantum ensembles, Phys. Rev. Lett. **74**, 1259 (1995).
[46] M. Maldovan, Sound and heat revolutions in phononics, Nature (London) **503**, 209 (2013).
[47] J. I. Cirac, P. Zoller, H. J. Kimble, and H. Mabuchi, Quantum State Transfer and Entanglement Distribution among Distant Nodes in a Quantum Network, Phys. Rev. Lett. **78**, 3221 (1997).
[48] K. Jähne, B. Yurke, and U. Gavish, High-fidelity transfer of an arbitrary quantum state between harmonic oscillators, Phys. Rev. A **75**, 010301 (2007).
[49] A. N. Korotkov, Flying microwave qubits with nearly perfect transfer effciency, Phys. Rev. B **84**, 014510 (2011).
[50] U. Dorner and P. Zoller, Laser-driven atoms in half-cavities, Phys. Rev. A **66**, 023816 (2002).
[51] T. Schröder, M. E. Trusheim, M. Walsh, L. Li, J. Zheng, M. Schukraft, A. Sipahigil, R. E. Evans, D. D. Sukachev, C. T. Nguyen, J. L. Pacheco, R. M. Camacho, E. S. Bielejec, M. D. Lukin, and D. Englund, Scalable focused ion beam creation of nearly lifetime-limited single quantum emitters in diamond nanostructures, Nat. Commun. **8**, 15376 (2017).
[52] A. H. Safavi-Naeini and O. Painter, Proposal for an optomechanical traveling wave phononphoton translator, New J. Phys. **13**, 013017 (2011).
[53] R. N. Patel, Z. Wang, W. Jiang, C. J. Sarabalis, J. T. Hill, and A. H. Safavi-Naeini, A single-mode phononic wire, arXiv:1711.00847.
[54] C. Brendel, V. Peano, O. Painter, and F. Marquardt, Snowflake Topological Insulator for Sound Waves, arXiv:1701.06330.
[55] J. T. Shen and S. Fan, Coherent single photon transport in a one-dimensional waveguide coupled with superconducting quantum bits, Phys. Rev. Lett. **95**, 213001(2005).

[56] L. Zhou, H. Dong, Yu-xi Liu, C. P. Sun, and F. Nori, Quantum supercavity with atomic mirrors, Phys. Rev. A **78**, 063827 (2008).

[57] D. E. Chang, L. Jiang, A. V. Gorshkov, and H. J. Kimble, Cavity QED with atomic mirrors, New J. Phys. **14**, 063003 (2012).

[58] S. Meesala, Y-I. Sohn, *et al.*, Strain engineering of the silicon vacancy center in diamond, in preparation (2017).

[59] Ö. O. Soykal, R. Ruskov, and C. Tahan, Sound-Based Analogue of Cavity Quantum Electrodynamics in Silicon, Phys. Rev. Lett. **107**, 235502 (2011).

[60] R. Ruskov and C. Tahan, On-chip cavity quantum phonodynamics with an acceptor qubit in silicon, Phys. Rev. B **88**, 064308 (2013).

[61] T. Ramos, V. Sudhir, K. Stannigel, P. Zoller, and T. J. Kippenberg, Nonlinear Quantum Optomechanics via Individual Intrinsic Two-Level Defects, Phys. Rev. Lett. **110**, 193602 (2013).

[62] M. V. G. Dutt, L. Childress, L. Jiang, E. Togan, J. Maze, F. Jelezko, A. S. Zibrov, P. R. Hemmer, and M. D. Lukin, Quantum Register Based on Individual Electronic and Nuclear Spin Qubits in Diamond, Science **316**, 5829 (2007).

[63] A. Reiserer, N. Kalb, M. S. Blok, K. J. M. van Bemmelen, D. J. Twitchen, M. Markham, T. H. Taminiau, and R. Hanson, Robust quantum-network memory using decoherence-protected subspaces of nuclear spins, Phys. Rev. X **6**, 021040 (2016).

Supplemental material: Phonon networks with SiV centers in diamond waveguides

M.-A. Lemonde[1], S. Meesala[2], A. Sipahigil[3], M. J. A. Schuetz[4], M. D. Lukin[4], M. Loncar[2], P. Rabl[1]
[1] *Vienna Center for Quantum Science and Technology, Atominstitut, TU Wien, 1040 Vienna, Austria*
[2] *John A. Paulson School of Engineering and Applied Sciences, Harvard University, 29 Oxford Street, Cambridge, MA 02138, USA*
[3] *Institute for Quantum Information and Matter and Thomas J. Watson, Sr., Laboratory of Applied Physics, California Institute of Technology, Pasadena, California 91125, USA and*
[4] *Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA*
(Dated: January 9, 2018)

arXiv:1801.01904v1 [quant-ph] 5 Jan 2018

ELECTRONIC STRUCTURE OF THE SiV$^-$ CENTER

As described in the main text, the electronic ground state of the negatively charged SiV consists of a single unpaired hole with spin $S = 1/2$, which can occupy one of the two degenerate orbital states $|e_x\rangle$ or $|e_y\rangle$. Within the ground state subspace and in the presence of a static external magnetic field $\vec{B}$, the energy structure is determined by a spin-orbit interaction, a Jahn-Teller (JT) effect and the Zeeman splittings. The resulting Hamiltonian reads ($\hbar = 1$) [1, 2]

$$H_{\rm SiV} = -\lambda_{\rm SO} L_z S_z + H_{\rm JT} + f\gamma_L B_z L_z + \gamma_s \vec{B}\cdot\vec{S}. \tag{1}$$

Here, L_z and S_z are the projections of the dimensionless angular momentum and spin operators $\vec{L}$ and $\vec{S}$ onto the symmetry axis of the center, which we assume to be aligned along the z-axis. $\lambda_{\rm SO} > 0$ is the spin-orbit coupling while γ_L and γ_s are the orbital and spin gyromagnetic ratio. respectively. The parameter $f \approx 0.1$ accounts for the reduced orbital Zeeman effect in the crystal lattice. Note that within the ground-state subspace spanned by $|e_x\rangle$ and $|e_y\rangle$, only L_z is non-zero. Within this basis and for an external magnetic field $\vec{B} = B_0\vec{z}$, the different contributions of Eq. (1) read

$$(\omega_B - \lambda_{\rm SO}\hat{L}_z)\hat{S}_z + \hat{H}_{\rm JT} = \frac{1}{2}\begin{bmatrix} \omega_B & i\lambda_{\rm SO} \\ -i\lambda_{\rm SO} & \omega_B \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} \Upsilon_x & \Upsilon_y \\ \Upsilon_y & -\Upsilon_x \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{2}$$

Here, Υ_x (Υ_y) denotes the strength of the Jahn-Teller coupling along x (y) and $\omega_B = \gamma_s B_0$ is the Zeeman energy. From this point, we neglect for simplicity the effect of the reduced orbital Zeeman interaction ($\sim f\gamma_L B_0$), which does not affect any of the results in the main text. Diagonalizing Eq. (2) leads to the eigenstates

$$\begin{aligned} |1\rangle &= \left(\cos\theta|e_x\rangle - i\sin\theta e^{-i\phi}|e_y\rangle\right)|\downarrow\rangle, \\ |2\rangle &= \left(\cos\theta|e_x\rangle + i\sin\theta e^{i\phi}|e_y\rangle\right)|\uparrow\rangle, \\ |3\rangle &= \left(\sin\theta|e_x\rangle + i\cos\theta e^{-i\phi}|e_y\rangle\right)|\downarrow\rangle, \\ |4\rangle &= \left(\sin\theta|e_x\rangle - i\cos\theta e^{i\phi}|e_y\rangle\right)|\uparrow\rangle, \end{aligned} \tag{3}$$

where

$$\tan(\theta) = \frac{2\Upsilon_x + \Delta}{\sqrt{\lambda_{\rm SO}^2 + 4\Upsilon_y^2}}, \qquad \tan(\phi) = \frac{2\Upsilon_y}{\lambda_{\rm SO}}. \tag{4}$$

The corresponding eigenenergies are

$$E_{3,1} = (-\omega_B \pm \Delta)/2, \quad E_{4,2} = (\omega_B \pm \Delta)/2, \tag{5}$$

with $\Delta = \sqrt{\lambda_{\rm SO}^2 + 4(\Upsilon_x^2 + \Upsilon_y^2)} \approx 2\pi \times 46$ GHz. Since $\Upsilon_{x,y} \ll \lambda_{\rm SO}$ (cf. Ref. [1]), we can neglect the small distortions of the orbital states by the JT effect and therefore use the approximation $|1\rangle \approx |e_-,\downarrow\rangle$, $|2\rangle \approx |e_+,\uparrow\rangle$, $|3\rangle \approx |e_+,\downarrow\rangle$ and $|4\rangle \approx |e_-,\uparrow\rangle$, which corresponds to $\theta = \pi/4$ and $\phi = 0$.

PHONON WAVEGUIDE

In the main text we consider a diamond phonon waveguide with a cross section A and a length $L \gg \sqrt{A}$. Within the frequency range of interest, the phonon modes can be modelled as elastic waves with a displacement field $\vec{u}(\vec{r}, t)$

obeying the equation of motion for a linear, isotropic medium [3],

$$\rho\frac{\partial^2}{\partial t^2}\vec{u} = (\lambda+\mu)\vec{\nabla}(\vec{\nabla}\cdot\vec{u}) + \mu\vec{\nabla}^2\vec{u}, \tag{6}$$

or in terms of the individual components

$$\rho\frac{\partial^2}{\partial t^2}u_k = (\lambda+\mu)\sum_m \frac{\partial^2 u_m}{\partial x_k \partial x_m} + \mu\sum_m \frac{\partial^2 u_k}{\partial x_m^2}. \tag{7}$$

Here, ρ is the mass density and the Lamé constants

$$\lambda = \frac{\nu E}{(1+\nu)(1-2\nu)}, \qquad \mu = \frac{E}{2(1+\nu)}, \tag{8}$$

can be expressed in terms of the Young's modulus E and the Poisson ratio ν. In our calculations and finite element method (FEM) simulations, we use $\rho = 3500$ kg/m^3, $E = 1050$ GPa and $\nu = 0.2$.

By assuming periodic boundary conditions, the equations of motion can be solved by the general ansatz

$$\vec{u}(\vec{r},t) = \frac{1}{\sqrt{2}}\sum_{k,n}\vec{u}^{\perp}_{n,k}(y,z)\left[A_{n,k}(t)e^{ikx} + A^*_{n,k}(t)e^{-ikx}\right], \tag{9}$$

where $k = 2\pi/L \times m$ is the wavevector along the waveguide direction x, and the index n labels the different phonon branches. The amplitudes $A_{n,k}(t)$ are oscillating functions obeying $\ddot{A}_{n,k}(t)+\omega^2_{n,k}A_{n,k}(t) = 0$, and the mode frequencies $\omega_{n,k}$ and the transverse mode profile $\vec{u}^{\perp}_{n,k}(x,y)$ are in general obtained from a numerical solution of the Eq. (6). The $\vec{u}^{\perp}_{n,k}(x,y)$ are orthogonal and normalized to

$$\frac{1}{A}\int dydz\,\vec{u}^{\perp}_{n,k}\cdot\vec{u}^{\perp}_{\beta,k} = \delta_{n,\beta}. \tag{10}$$

Quantization of the displacement field

Eq. (6) can be derived from the Lagrangian

$$L = \int d^3r\left[\frac{\rho}{2}\dot{\vec{u}}^2 - \frac{(\lambda+\mu)}{2}\sum_{k,m}\frac{\partial u_k}{\partial x_m}\frac{\partial u_m}{\partial x_k} - \frac{\mu}{2}\sum_{k,m}\left(\frac{\partial u_k}{\partial x_m}\right)^2\right]. \tag{11}$$

After inserting the eigenmode decomposition in Eq. (9), the Lagrangian reduces to a set of harmonic modes:

$$L(\{Q_{n,k}\},\{\dot{Q}_{n,k}\}) = \sum_{k,n}\frac{M}{2}\dot{Q}_{n,k}\dot{Q}_{n,-k} - \frac{1}{2}M\omega^2_{n,k}Q_{n,k}Q_{n,-k}, \tag{12}$$

where $M = \rho AL$ and $Q_{n,k} = (A_{n,k}+A^*_{n,-k})/\sqrt{2}$. From this simplified form, we readily obtain the canonical momenta $P_{n,k} = \partial L/\partial\dot{Q}_{n,k} = M\dot{Q}_{n,-k}$, and the Hamiltonian operator

$$H_{\rm ph} = \sum_{k,n}\frac{P_{n,k}P_{n,-k}}{2M} + \frac{1}{2}M\omega^2_{n,k}Q_{n,k}Q_{n,-k}, \tag{13}$$

where $Q_{n,k}$ and $P_{n,k}$ are now operators obeying the canonical commutation relations, $[Q_{n,k},P_{n,k}] = i\hbar\delta_{n,n'}\delta_{kk'}$. Finally, we write

$$Q_{n,k} = \sqrt{\frac{\hbar}{2M\omega_{n,k}}}\left(a^\dagger_{n,k}+a_{n,-k}\right), \qquad P_{n,k} = i\sqrt{\frac{\hbar M\omega_{n,k}}{2}}\left(a^\dagger_{n,k}-a_{n,-k}\right), \tag{14}$$

in terms of annihilation and creation operators. We obtain

$$H_{\rm ph} = \sum_{k,n}\hbar\omega_{n,k}a^\dagger_{n,k}a_{n,k}, \tag{15}$$

and the quantized displacement field

$$\vec{u}(\vec{r}) = \sum_{k,n}\sqrt{\frac{\hbar}{2M\omega_{n,k}}}\vec{u}^{\perp}_{n,k}(y,z)\left(a_{n,k}e^{ikx}+a^\dagger_{n,k}e^{-ikx}\right). \tag{16}$$

COUPLING TO PHONON MODES

Strain coupling arises from the change in Coulomb energy of the electronic states due to displacement of the atoms forming the defect. For small displacements and in the Born-Oppenheimer approximation, the energy shift is linear in the local distortion and can be written as

$$H_{\text{strain}} = \sum_{ij} V_{ij}\epsilon_{ij}. \tag{17}$$

Here, V is an operator acting on the electronic states of the SiV defect and ϵ is the strain tensor defined as

$$\epsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \tag{18}$$

with u_1 (u_2, u_3) representing the quantized displacement field along $x_1 = x$ ($x_2 = y, x_3 = z$) at the position of the SiV center [cf. Eq. (16)]. The axes are defined as in Fig. 1 of the main text, i.e. the symmetry axis of the defect is along z while the waveguide is along x.

The exact form of the strain interaction Hamiltonian in the basis of the electronic states of the SiV defect is obtained by projecting the strain tensor on the irreducible representations of the D_{3d} group, i.e.

$$H_{\text{strain}} = \sum_{r} V_r\epsilon_r, \tag{19}$$

where r denotes the irreducible representations. One can show that the only contributing representations are the one-dimensional representation A_{1g} and the two-dimensional representation E_g [2]. As a consequence, strain can couple independently to orbitals within the E_g and E_u manifolds, but these manifolds cannot be mixed. Focusing only on the ground state, the terms in Eq. (19) are [4]

$$\begin{aligned}
\epsilon_{A_{1g}} &= t_\perp(\epsilon_{xx} + \epsilon_{yy}) + t_\parallel\epsilon_{zz} \\
\epsilon_{E_{gx}} &= d(\epsilon_{xx} - \epsilon_{yy}) + f\epsilon_{zx} \\
\epsilon_{E_{gy}} &= -2d\epsilon_{xy} + f\epsilon_{yz}
\end{aligned} \tag{20}$$

Here, $t_\perp, t_\parallel, d, f$ are the strain-susceptibilities with $f/d \sim 10^{-4}$, and the subscript g is used to denote the ground state manifold. The effects of these strain components on the electronic states are described by

$$\begin{aligned}
V_{A_{1g}} &= |e_x\rangle\langle e_x| + |e_y\rangle\langle e_y| \\
V_{E_{gx}} &= |e_x\rangle\langle e_x| - |e_y\rangle\langle e_y| \\
V_{E_{gy}} &= |e_x\rangle\langle e_y| + |e_y\rangle\langle e_x|
\end{aligned} \tag{21}$$

Since coupling to symmetric local distortions ($\sim \epsilon_{A_{1g}}$) shifts all ground states equally, it has no relevant effects in this work and can thus be dropped. Finally, if we write the strain Hamiltonian using the basis spanned by the eigenstates of the spin-orbit coupling $\{|e_-\rangle, |e_+\rangle\}$, we find

$$H_{\text{strain}} = \epsilon_{E_{gx}}(L_- + L_+) - i\epsilon_{E_{gy}}(L_- - L_+),$$

where $L_+ = L_-^\dagger = |3\rangle\langle 1| + |2\rangle\langle 4|$ is the orbital raising operator within the ground state. Further, we notice that the transitions L_+, L_- have circularly polarized selection rules in the E_g strain components.

We now assume that the SiV high symmetry axis z is oriented orthogonal to the phonon propagation direction x (practically, this can be realized with [110]-oriented diamond waveguides). By decomposing the local displacement field as in Eq. (16), and after making a rotating wave approximation, the resulting strain coupling can be written as

$$H_{\text{strain}} \simeq \frac{1}{\sqrt{L}}\sum_{n,k}\left[(g_{n,k}J_+^\uparrow + g_{n,-k}^* J_+^\downarrow)a_{n,k}e^{ikx} + \text{H.c.}\right], \tag{22}$$

where $J_+^\uparrow = |3\rangle\langle 1|$, $J_+^\downarrow = |4\rangle\langle 2|$ and

$$\begin{aligned}
g_{n,k} &= d\sqrt{\frac{\hbar k^2}{2\rho A\omega_{n,k}}}\frac{1}{|k|}\left[\left(iku_{n,k}^{\perp,x} + ik\frac{f}{d}\frac{u_{n,k}^{\perp,z}}{2} + \frac{f}{d}\frac{\partial_z u_{n,k}^{\perp,x}}{2} - \partial_y u_{n,k}^{\perp,y}\right) - i\left(iku_{n,k}^{\perp,y} + \partial_y u_{n,k}^{\perp,x} + \frac{f}{d}\frac{\partial_y u_{n,k}^{\perp,z}}{2} + \frac{\partial_z u_{n,k}^{\perp,y}}{2}\right)\right], \\
&\equiv d\sqrt{\frac{\hbar k^2}{2\rho A\omega_{n,k}}}\xi_{n,k}(y,z).
\end{aligned} \tag{23}$$

Here, $u^{\perp,i}_{n,k}$ represents the i-th component of the displacement pattern $\vec{u}^{\perp}_{n,k}(y,z)$. The first four terms in the square bracket correspond to E_{g_x} deformations, while the last four correspond to E_{g_y} deformations. We note from Eq. (22) that due to circularly polarized selection rules, it is possible to have different coupling rates to left or right propagating phonons and that this directionality is reversed, when the spin character of the states involved in the phononic transition is flipped. This is due to the particular energy-state ordering in which $E_{\downarrow,+} > E_{\downarrow,-}$ while $E_{\uparrow,+} < E_{\uparrow,-}$. However, the waveguide phonon modes considered in this work are approximately linearly polarized with predominantly E_{gx} strain, and hence have identical coupling rates for both propagation directions (and spin projections). Therefore, the strain Hamiltonian reduces to

$$H_{\text{strain}} \simeq \frac{1}{\sqrt{L}} \sum_{n,k} g_{n,k} J_+ a_{n,k} e^{ikx} + \text{H.c.} \tag{24}$$

SPIN-PHONON INTERFACE

In this section, we present in more details two different driving schemes for transferring spin-states encoded in the SiV ground-state to propagating phonons. We first consider the scenario depicted in the main text that utilizes a microwave drive within the ground-state subspace. Furthermore, we present a second approach via optical Raman transitions to the excited states, which can be a useful alternative to microwave magnetic fields. For simplicity, we first focus on a single SiV center in an infinite waveguide.

Microwave driving fields

The starting point is the Hamiltonian of a single driven SiV center coupled via strain to the phonon modes of the diamond waveguide,

$$H = H_{\text{SiV}} + H_{\text{ph}} + H_{\text{drive}} + H_{\text{strain}}. \tag{25}$$

By moving into the interaction picture with respect to

$$H_0 = \sum_{n,k} \omega_0 a^\dagger_{n,k} a_{n,k} + \omega_B |2\rangle\langle 2| + \omega_0 |3\rangle\langle 3| + (\omega_0 + \omega_B)|4\rangle\langle 4|, \tag{26}$$

we obtain the new Hamiltonian $\tilde{H} = e^{iH_0 t} H e^{-iH_0 t} - H_0$ given by

$$\begin{aligned}\tilde{H} =& \sum_{n,k} (\omega_{n,k} - \omega_0) a^\dagger_{n,k} a_{n,k} - \delta(|3\rangle\langle 3| + |4\rangle\langle 4|) \\ &+ \left[\frac{\Omega(t) e^{i\theta(t)}}{2} (|3\rangle\langle 2| + e^{2i\omega_B t}|4\rangle\langle 1|) + \frac{1}{\sqrt{L}} \sum_{n,k} g_{n,k} e^{ikx} a_{n,k} (|3\rangle\langle 1| + |4\rangle\langle 2|) + \text{H.c.} \right].\end{aligned} \tag{27}$$

Here, $\omega_0 = \Delta + \delta$ is the central frequency of the emitted phonon wavepackets and $\Omega(t)$ and $\theta(t)$ are the strength and phase of the external driving field, respectively.

In this rotating frame, the ansatz for the single-excitation wavefunction reads

$$|\psi(t)\rangle = \alpha|1,0\rangle + \beta\big[c(t)|2\rangle\langle 1| + b(t)|3\rangle\langle 1| + \sum_{n,k} c_{n,k}(t) a^\dagger_{n,k}\big]|1,0\rangle, \tag{28}$$

where $|1,0\rangle$ is the ground state with the SiV center in state $|1\rangle$ and no phonons in the waveguide. Note that this ansatz does not capture the off-resonant transition $|1\rangle \to |4\rangle$ produced by the drive [i.e., the term $\sim e^{2i\omega_B t}$ in Eq. (27)]. We estimate its effect below and show that for the parameters considered in this work, it can be neglected.

From the Schrödinger equation $\partial_t|\psi(t)\rangle = -i\tilde{H}(t)|\psi(t)\rangle$, we obtain the equations of motion for the amplitudes,

$$\begin{aligned}\dot{c}_{n,k}(t) &= -i(\omega_{n,k} - \omega_0) c_{n,k}(t) - i\frac{1}{\sqrt{L}} g^*_{n,k} e^{-ikx} b(t), \\ \dot{c}(t) &= -i\frac{\Omega(t) e^{-i\theta(t)}}{2} b(t), \\ \dot{b}(t) &= i\delta b(t) - i\frac{\Omega(t) e^{i\theta(t)}}{2} c(t) - i\frac{1}{\sqrt{L}} \sum_{n,k} g_{n,k} e^{ikx} c_{n,k}(t).\end{aligned} \tag{29}$$

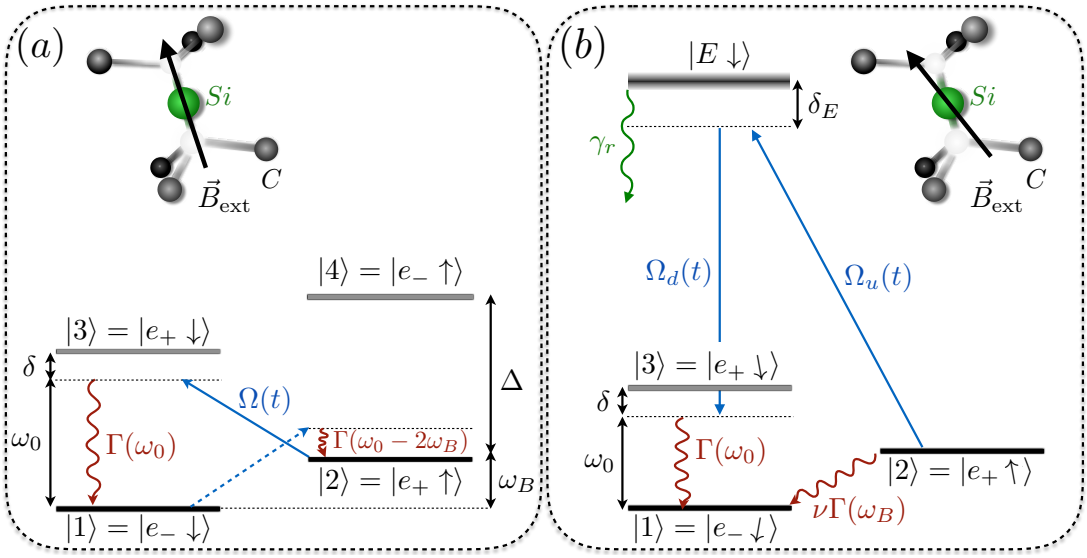

FIG. 1: Comparison of the two driving schemes to implement the spin-phonon interface. (a) A microwave magnetic field drives the transition $|2\rangle \to |3\rangle$, and state $|3\rangle$ subsequently decays to state $|1\rangle$ by emitting a propagating phonon at frequency ω_0. The process that drives the transition $|1\rangle \to |4\rangle$ is strongly off-resonant for large Zeeman energy $\omega_B \gg \delta$ and can be neglected. (b) The transition $|2\rangle \to |3\rangle$ is now driven by two optical fields via the excited state $|e_u\downarrow\rangle$ of the SiV center. In that case, the external magnetic field has to be tilted from the symmetry axis of the defect. As a consequence, it opens a decoherence channel via the direct transition $|2\rangle \to |1\rangle$.

The solution for the propagating phonons reads

$$c_{n,k}(t) = e^{-i(\omega_{n,k}-\omega_0)(t-t_0)}c_{n,k}(t_0) - \frac{i}{\sqrt{L}}g^*_{n,k}e^{-ikx}\int_{t_0}^{t} d\tau e^{-i(\omega_{n,k}-\omega_0)(t-\tau)}b(\tau), \tag{30}$$

where t_0 is an arbitrary time before any phonons interacted with the SiV center. Plugging this result back into the equation for the excited state amplitude, we obtain

$$\dot{b}(t) = i\delta b(t) - i\frac{\Omega(t)e^{i\theta(t)}}{2}c(t) - i\frac{1}{\sqrt{L}}\sum_{n,k} g_{n,k}e^{ikx}e^{-i(\omega_{n,k}-\omega_0)t}c_{n,k}(t_0) - \frac{1}{L}\sum_{n,k}|g_{n,k}|^2\int_{t_0}^{t} d\tau e^{-i(\omega_{n,k}-\omega_0)(t-\tau)}b(\tau). \tag{31}$$

In the present case where the SiV center is driven by phonons of frequencies close to $\omega = 0$ (ω_0 in the lab frame), also the amplitude $b(t)$ is slowly varying, allowing us to perform a standard Markov approximation [7]. This results in

$$\dot{b}(t) = \left[i\delta - \frac{\Gamma(\omega_0)}{2}\right]b(t) - i\frac{\Omega(t)e^{i\theta(t)}}{2}c(t) - \sum_n \sqrt{\frac{\Gamma_n(\omega_0)}{2}}\left[\Phi^{\rm in,L}_{n,\omega_0}(t) + \Phi^{\rm in,R}_{n,\omega_0}(t)\right], \tag{32}$$

with the phonon-induced decay rate

$$\Gamma(\omega) = \frac{2\pi}{L}\sum_{n,k}|g_{n,k}|^2\delta(\omega-\omega_{n,k}) = 2\frac{|g_n|^2}{v_n}, \tag{33}$$

and the input field

$$\Phi^{\rm in,L/R}_{n,\omega}(t) = i\sum_{k>0}\sqrt{\frac{v_n}{L}}e^{\mp ikx}e^{-i(\omega_{n,k}-\omega)(t-t_0)}c_{n,k}(t_0). \tag{34}$$

Here, the group velocity $v_n = d\omega_{n,k}/dk$ and the coupling constant $g_n = g_{n,k}$ are evaluated at ω_0 and are considered constant over the frequency range of interest $[\delta\omega \sim \Gamma(\omega)]$ around ω_0.

We now make further simplifications and consider weak and slowly-varying driving fields, i.e., $\Omega(t) \ll |\delta + i\Gamma(\omega_0)/2|$, $\dot{\Omega}(t)/|\delta + i\Gamma(\omega_0)/2|$ and $\dot{\theta}(t) \ll |\delta + i\Gamma(\omega_0)/2|$. Given those constraints, one can adiabatically eliminate the higher-energy state, i.e. $\dot{b}(t) = 0$, and obtain

$$\dot{c}(t) = -\left[i\omega_s(t) + \frac{\gamma(t)}{2}\right] c(t) - \sum_n \sqrt{\frac{\gamma_n(t)}{2}} e^{-i\bar{\theta}(t)} \left[\Phi_n^{\mathrm{in,L}}(t) + \Phi_n^{\mathrm{in,R}}(t)\right], \tag{35}$$

with

$$\omega_s(t) = \frac{\Omega^2(t)}{4} \frac{\delta}{\delta^2 + \Gamma^2(\omega_0)/4}, \quad \gamma(t) = \sum_n \gamma_n(t) = \frac{\Omega^2(t)/4}{\delta^2 + \Gamma^2(\omega_0)/4} \sum_n \Gamma_n(\omega_0), \quad \bar{\theta}(t) = \theta(t) + \arctan\left[\frac{\Gamma(\omega_0)}{2\delta}\right] \tag{36}$$

The AC-Stark-shift $\omega_s(t)$ can be compensated by a corresponding (slow) adjustment of the driving frequency $\omega_d(t)$ to keep ω_0 constant during the entire driving protocol. Doing so and omitting the constant shift of the drive phase $\bar{\theta}(t) \to \theta(t)$ for simplicity, one recovers the form introduced in Eq. (7) of the main text.

Residual driving of the transition $|1\rangle \to |4\rangle$

As described by Eq.(27), the microwave field also drives the transition $|1\rangle \to |4\rangle$. One can estimate the rate at which this process takes place by applying the same procedure as above starting from the following ansatz

$$|\psi(t)\rangle = \alpha|2,0\rangle + \beta\big[c(t)|1\rangle\langle 2| + b(t)|4\rangle\langle 2| + \sum_{n,k} c_{n,k}(t) a_{n,k}^\dagger\big]|2,0\rangle, \tag{37}$$

where $|2,0\rangle$ represents the SiV center in state $|2\rangle$ and no phonons in the waveguide. Doing so, one finds

$$\dot{c}(t) = -\left[i\tilde{\omega}_s(t) + \frac{\tilde{\gamma}(t)}{2}\right] c(t) - \sum_n \sqrt{\frac{\tilde{\gamma}_n(t)}{2}} e^{-i\theta(t)} [\Phi_{n,\omega_0-2\omega_B}^{\mathrm{in,L}}(t) + \Phi_{n,\omega_0-2\omega_B}^{\mathrm{in,R}}(t)], \tag{38}$$

with the input fields defined in Eq. (34), the AC-Stark-shift and the effective transfer rate

$$\tilde{\omega}_s(t) = \frac{\Omega^2(t)}{4} \frac{\delta - 2\omega_B}{(\delta - 2\omega_B)^2 + \Gamma^2(\omega_0 - 2\omega_B)/4}, \qquad \tilde{\gamma}(t) = \frac{\Omega^2(t)/4}{(\delta - 2\omega_B)^2 + \Gamma^2(\omega_0 - 2\omega_B)/4} \sum_n \Gamma_n(\omega_0 - 2\omega_B). \tag{39}$$

As a consequence, the drive also allows the SiV states to flip from $|1\rangle$ to $|2\rangle$ by emitting a phonon at frequency $\omega_0 - 2\omega_B$. For large Zeeman splittings, the rate of this process

$$\frac{\tilde{\Gamma}(t)}{\Gamma(t)} \sim \frac{\delta^2}{(\delta - 2\omega_B)^2} \frac{\Gamma_{\mathrm{ph}}(\omega_0 - 2\omega_B)}{\Gamma_{\mathrm{ph}}(\omega_0)} \tag{40}$$

is strongly suppressed, as long as $\Gamma_{\mathrm{ph}}(\omega_0 - 2\omega_B) \simeq \Gamma_{\mathrm{ph}}(\omega_0)$. Therefore, care must be taken to avoid band edges at frequencies near $\omega_0 - 2\omega_B$.

Optical Raman driving schemes

We now present an alternative driving scheme that makes use of the electronically excited states via an optical two-tone Raman transition, as depicted in Fig. 1 (b). The corresponding Hamiltonian reads

$$H = H_{\mathrm{SiV}} + H_{\mathrm{ph}} + H_{\mathrm{drive}} + H_{\mathrm{strain}} + H_{\mathrm{rad}}, \tag{41}$$

where H_{rad} captures the radiative decay of the excited states, H_{drive} describes the optical driving of the excited state and H_{SiV} now includes a component of the magnetic field perpendicular to the symmetry axis of the defect (e.g. along x). The perpendicular field allows one to couple opposite-spin states via optical driving fields [4–6].

Effects of a weak perpendicular magnetic field

Focusing only on the ground-state subspace and the only relevant excited state $|E\downarrow\rangle$,

$$\begin{aligned} H_{\rm SiV} &= -\lambda_{\rm SO} L_z S_z + \omega_B S_z + \omega_x B_x S_x, \\ &= -\frac{\Delta+\omega_B}{2}|e_-\downarrow\rangle\langle e_-\downarrow| - \frac{\Delta-\omega_B}{2}|e_+\uparrow\rangle\langle e_+\uparrow| \\ &\quad + \frac{\Delta-\omega_B}{2}|e_+\downarrow\rangle\langle e_+\downarrow| + \frac{\Delta+\omega_B}{2}|e_-\uparrow\rangle\langle e_-\uparrow| + \omega_E|E\downarrow\rangle\langle E\downarrow|, \end{aligned} \tag{42}$$

where $\omega_B = \gamma_s B_z$, $\omega_x \equiv \gamma_s B_x$, and ω_E is the energy of the excited state. As described in the first section, Eq. (42) neglects the orbital Zeeman effect and the distortion of the orbital states due to the JT effect.

In the limit of weak perpendicular magnetic field $\omega_x/|\Delta-\omega_B| \ll 1$, a small mixing between opposite-spin states occurs, leading to new eigenstates:

$$H_{\rm SiV} = \sum_{i=1,4}\omega_i|i\rangle\langle i| + \omega_E|E\downarrow\rangle\langle E\downarrow|, \qquad \Rightarrow \begin{cases} |1\rangle \approx |e_-\downarrow\rangle - \eta_+|e_-\uparrow\rangle, & \omega_1 \approx -\frac{\Delta+\omega_B}{2} - \frac{\eta_+\omega_x}{2} \\ |2\rangle \approx |e_+\uparrow\rangle - \eta_-|e_+\downarrow\rangle, & \omega_2 \approx -\frac{\Delta-\omega_B}{2} - \frac{\eta_-\omega_x}{2} \\ |3\rangle \approx |e_+\downarrow\rangle + \eta_-|e_+\uparrow\rangle, & \omega_3 \approx \frac{\Delta-\omega_B}{2} + \frac{\eta_-\omega_x}{2} \\ |4\rangle \approx |e_-\uparrow\rangle + \eta_+|e_-\downarrow\rangle, & \omega_4 \approx \frac{\Delta+\omega_B}{2} + \frac{\eta_+\omega_x}{2} \end{cases}, \tag{43}$$

with

$$\eta_\pm \equiv \frac{1}{2}\frac{\omega_x}{\Delta\pm\omega_B}, \qquad \eta = \eta_- + \eta_+. \tag{44}$$

Note that due to the larger spin-orbit interaction in the excited state (~ 250GHz) the effect of B_x on $|E\downarrow\rangle$ can be neglected. In this new basis, the strain interaction given in Eq. (24) becomes

$$H_{\rm strain} = \sum_{n,k} g_{n,k} a_{n,k}\left[J_+ + \eta(|4\rangle\langle 3| - |2\rangle\langle 1|)\right] + \text{H.c.} \tag{45}$$

As a consequence, a magnetic field which is not perfectly aligned with the symmetry axis of the SiV center induces a finite strain coupling between states $|1\rangle \leftrightarrow |2\rangle$ and $|3\rangle \leftrightarrow |4\rangle$.

Within the same basis, the optical driving fields with frequencies ω_u and ω_d are described by

$$\begin{aligned} H_{\rm drive} &= \left(\frac{\Omega_d(t)e^{i\theta_d(t)}}{2}e^{-i\omega_d t} + \frac{\Omega_u(t)e^{i\theta_u(t)}}{2}e^{-i\omega_u t}\right)|E\downarrow\rangle\langle e_+\downarrow| + \text{H.c.} \\ &= \frac{\Omega_d(t)e^{i\theta_d(t)}}{2}e^{-i\omega_d t}|E\downarrow\rangle\langle 3| - \frac{\Omega_u(t)e^{i\theta_u(t)}}{2}\eta_- e^{-i\omega_u t}|E\downarrow\rangle\langle 2| + \text{H.c.} \end{aligned} \tag{46}$$

The last line is obtained by making a rotating wave approximation valid for large frequency mismatch between the two drives, i.e. $|\omega_d - \omega_u| \gg \Omega_{d,u}$.

Effective 3-level system

To extract the effective rate at which the spin state is transferred to propagating phonons and estimate the dephasing rates due to the radiative decay of the excited state and the direct strain coupling between state $|1\rangle$ and $|2\rangle$, we apply the same procedure as the previous section. This time, we work in the rotating frame with respect to

$$H_0 = \sum_{n,k}\omega_0 a^\dagger_{n,k}a_{n,k} + \sum_i \omega_i|i\rangle\langle i| + \delta(|3\rangle\langle 3| + |4\rangle\langle 4|) + \delta_E|E\downarrow\rangle\langle E\downarrow|, \tag{47}$$

and with the drives detuned such that $\omega_d = \omega_E - \omega_3 + \delta_E - \delta$ and $\omega_u = \omega_E - \omega_2 + \delta_E$ [see Fig. 1 (b)]. From the low-excitation ansatz

$$|\psi(t)\rangle = \alpha|1,0\rangle + \beta\big[c(t)|2\rangle\langle 1| + b(t)|3\rangle\langle 1| + E(t)|E\downarrow\rangle\langle 1| + \sum_{n,k} c_{n,k}(t)a^\dagger_{n,k}\big]|1,0\rangle, \tag{48}$$

we derive the Schrödinger equations for the time-dependent coefficients. We approximate the effects of the dipole interaction ($H_{\rm rad}$) by including a finite lifetime of the excited state $|E\downarrow\rangle$ in the form of a radiative decay $\Gamma_{\rm rad}$ (~ 100 MHz), i.e.

$$\dot{E}(t)=\left(i\delta_E-\frac{\Gamma_{\rm rad}}{2}\right)E(t)-\frac{i}{2}\big[\Omega_d(t)e^{i\theta_d(t)}b(t)-\eta_-\Omega_u(t)e^{i\theta_u(t)}c(t)\big]. \tag{49}$$

We focus on the limit of weak optical drives $\Omega_{u,d}(t)\ll|\delta_E+i\Gamma_{\rm rad}/2|,|\delta+i\Gamma/2|$ so that we can adiabatically eliminate the excited state $[\dot{E}(t)=0]$. Within the Markov approximation, this leads to an effective 3-level system, where

$$\begin{aligned}
\dot{b}(t)&=-\left\{i\big[\omega_d(t)-\delta\big]+\frac{\gamma^d_{\rm rad}(t)}{2}+\frac{\Gamma(\omega_0)}{2}\right\}b(t)+i\frac{\Omega_{\rm eff}(t)}{2}e^{i[\theta_{\rm eff}(t)-\phi_{\rm NH}]}c(t)+\sqrt{\frac{\Gamma_n(\omega_0)}{2}}\big[\Phi^{\rm in,L}_{n,\omega_0}(t)+\Phi^{\rm in,R}_{n,\omega_0}(t)\big],\\
\dot{c}(t)&=-\left\{i\omega_u(t)+\frac{\gamma^u_{\rm rad}(t)}{2}+\frac{\eta^2\Gamma(\omega_B)}{2}\right\}c(t)+i\frac{\Omega_{\rm eff}(t)}{2}e^{-i[\theta_{\rm eff}(t)+\phi_{\rm NH}]}b(t)+\sqrt{\frac{\eta^2\Gamma_n(\omega_B)}{2}}\big[\Phi^{\rm in,L}_{n,\omega_B}(t)+\Phi^{\rm in,R}_{n,\omega_B}(t)\big],
\end{aligned} \tag{50}$$

with the phonon-induced decay rate and input fields defined in Eqs. (33) and (34) respectively. The phase $\phi_{\rm NH}=\arctan(\Gamma_{\rm rad}/2\delta_E)$ comes from the radiative decay and can be neglected for large detunings $\delta_E\gg\Gamma_{\rm rad}$. Note that in Eqs. (50), we have neglected higher-order virtual processes that couple states $|2\rangle$ and $|3\rangle$ via strain interaction that are strongly off-resonant for $|g_{n,k}|^2/|\Delta-\omega_B|\ll\Gamma(\omega_0),\Gamma(\omega_B)$.

At this stage, we recover the 3-level system utilized within the magnetic driving scheme described above, except for the AC-Stark shifts of states $|2\rangle$ and $|3\rangle$,

$$\omega_u(t)=\frac{\eta_-^2}{4}\frac{\Omega_u^2(t)}{\delta_E^2+\Gamma_{\rm rad}^2/4}\delta_E,\qquad \omega_d(t)=\frac{1}{4}\frac{\Omega_d^2(t)}{\delta_E^2+\Gamma_{\rm rad}^2/4}\delta_E, \tag{51}$$

respectively, and additional decay channels. One of the new loss mechanism comes from the radiative decay of the excited state $|E\rangle$, which affects both states $|2\rangle$ and $|3\rangle$ with respective rates

$$\gamma^u_{\rm rad}(t)=\frac{\eta_-^2}{4}\frac{\Omega_u^2(t)}{\delta_E^2+\Gamma_{\rm rad}^2/4}\Gamma_{\rm rad},\qquad \gamma^d_{\rm rad}(t)=\frac{1}{4}\frac{\Omega_d^2(t)}{\delta_E^2+\Gamma_{\rm rad}^2/4}\Gamma_{\rm rad}, \tag{52}$$

while the finite strain coupling between states $|1\rangle$ and $|2\rangle$ also induces an addition decay channel with rate $\eta^2\Gamma(\omega_B)$ and incoming noise $\Phi^{\rm in}_{\omega_B}(t)$. Finally, the effective Rabi frequency driving the transition $|2\rangle\to|3\rangle$ is given by

$$\Omega_{\rm eff}(t)=\frac{\eta_-}{2}\frac{\Omega_d(t)\Omega_u(t)}{\sqrt{\delta_E^2+\Gamma_{\rm rad}^2/4}},\qquad \theta_{\rm eff}(t)=\theta_u(t)-\theta_d(t). \tag{53}$$

The viability of this scheme resides in the relative importance of the loss mechanisms compared to the coherent dynamics. More precisely, the phonon-assisted transfer rate of state $|3\rangle$ has to overcome its radiative decay, i.e. $\Gamma(\omega_0)\gg\gamma^d_{\rm rad}(t)$, while the other loss mechanisms as to be overcome by the final spin-state transfer rate, i.e. $\gamma(t)\sim\frac{\Omega^2_{\rm eff}(t)}{\delta^2}\Gamma(\omega_0)\gg\gamma^u_{\rm rad}(t),\eta^2\Gamma(\omega_B)$. As an example (all rates are divided by 2π), for Rabi frequencies $\Omega_d\approx\Omega_u/2\sim 2.5$ GHz, detunings $\delta_E\sim 30$ GHz and $\delta\sim 30$ MHz, a radiative decay rate $\Gamma_{\rm rad}\sim 100$ MHz, and a ratio $\eta_-\sim 0.1$, the different maximal rates are

$$\gamma^d_{\rm rad}\sim 150\,{\rm kHz}\ll\Gamma(\omega_0)\sim 2\,{\rm MHz},\qquad \gamma^u_{\rm rad}\sim 7\,{\rm kHz},\ \eta^2\Gamma(\omega_B)\sim 40\,{\rm kHz}\ll\gamma\sim 250\,{\rm kHz}. \tag{54}$$

Here, we use $\Gamma(\omega_B)\sim\Gamma(\omega_0)\sim 1$ MHz. Using optical driving should thus be a viable route to achieve a fully controllable spin transfer into a propagating phonon.

INPUT-OUTPUT FORMALISM

In this section, we extend the previous calculations to multiple SiV defects and recover the input-output relations stated in the main text. We first start by considering an infinite waveguide, where we explicitly derive how the input field of a given center is related to the output field of the others. In this scenario, we estimate the effects of phonon scattering by undriven defects. Finally, we close the section by considering the effects of waveguide boundaries.

Infinite waveguides

The coherent dynamics of the SiV ensemble is governed by the Hamiltonian given in Eq. (2) in the main text, which in the rotating frame defined in Eq. (26) reads

$$H = \sum_{n,k}(\omega_{n,k}-\omega_0)a^\dagger_{n_k}a_{n_k} + \sum_j H^{(j)}_{\rm SiV} + \frac{1}{\sqrt{L}}\sum_{j,n,k}\left[g^j_{n,k}e^{ikx_j}a_{n,k}J^j_+ + \text{H.c.}\right], \tag{55}$$

with

$$H^{(j)}_{\rm SiV} = -\delta_j(|3\rangle\langle 3| + |4\rangle\langle 4|) + \left[\frac{\Omega_j(t)e^{i\theta_j(t)}}{2}|3\rangle\langle 2| + \text{H.c.}\right]. \tag{56}$$

In this frame, the single-excitation ansatz considered in the main text, $|\psi(t)\rangle = [\alpha\mathbb{1} + \beta C^\dagger(t)]|\bar{1},0\rangle$, is now defined with

$$C^\dagger(t) = \sum_{j=e,r}\left[c_j(t)|2\rangle_j\langle 1| + b_j(t)|3\rangle_j\langle 1| + \sum_{n,k}c_{n,k}(t)a^\dagger_{n,k}\right], \tag{57}$$

where $|\bar{1},0\rangle$ is the ground state with all SiV centers in state $|1\rangle$ and no phonon in the waveguide. In Eq. (57), we only kept the two driven centers, i.e. the emitting (e) and receiving (r) one.

The equations of motion for the different amplitudes are

$$\begin{aligned}
\dot{c}_{n,k}(t) &= -i(\omega_{n,k}-\omega_0)c_{n,k}(t) - i\frac{1}{\sqrt{L}}(g^e_{n,k})^*e^{-ikx_e}b_e(t) - i\frac{1}{\sqrt{L}}(g^r_{n,k})^*e^{-ikx_r}b_r(t),\\
\dot{c}_j(t) &= -i\frac{\Omega_j(t)e^{-i\theta_j(t)}}{2}b_j(t),\\
\dot{b}_j(t) &= i\delta_j b_j(t) - i\frac{\Omega_j(t)e^{i\theta_j(t)}}{2}c_j(t) - i\frac{1}{\sqrt{L}}\sum_{n,k}g^j_{n,k}e^{ikx_j}c_{n,k}(t).
\end{aligned} \tag{58}$$

We again apply the same procedure as in the previous sections, i.e., we first exactly solve the equation for $c_{n,k}(t)$, insert the solution in the equation for $b_j(t)$ and then perform a Markov approximation. Doing so for the receiving defect and taking $x_r > x_e$, we obtain

$$\begin{aligned}
\dot{b}_r(t) = &\left[i\delta_r - \frac{\Gamma_r(\omega_0)}{2}\right]b_r(t) - i\frac{\Omega_r(t)e^{i\theta_r(t)}}{2}c_r(t) - i\sum_n\sum_k\sqrt{\frac{\Gamma_{r,n}(\omega_0)}{2}}\sqrt{\frac{v_n}{L}}e^{ikx_r}e^{-i(\omega_{n,k}-\omega_0)(t-t_0)}c_{n,k}(t_0)\\
&-\sum_n\sqrt{\frac{\Gamma_{r,n}(\omega_0)}{2}\frac{\Gamma_{e,n}(\omega_0)}{2}}e^{ik_n(x_r-x_e)}b_e(t-\tau^n_{er}),
\end{aligned} \tag{59}$$

with $\tau^n_{er} = (x_r - x_e)/v_n$. We recall that $g_{n,k}$ is taken to be real without loss of generality and t_0 is a time in the past before the two SiV defects have interacted with incoming wavepackets.

The final step is to adiabatically eliminate the higher-energy state $[\dot{b}_j(t) = 0]$ for both SiV centers and insert the result in the equation for $c_j(t)$, which leads to the final form

$$\dot{c}_r(t) = -\left[i\omega_{s,r} + \frac{\gamma_r(t)}{2}\right]c_r(t) - \sum_n\sqrt{\frac{\gamma_{r,n}(t)}{2}}e^{-i\bar{\theta}_r(t)}[\Phi^{\rm in,L}_{r,n}(t) + \Phi^{\rm in,R}_{r,n}(t) + \Phi^{\rm scatt}_{r,n}(t)]. \tag{60}$$

Here, the AC-Stark shift $\omega_{s,r}$, the effective transfer rate $\gamma_{r,n}(t)$ and the shifted driving phase $\bar{\theta}_r(t)$ are all defined in Eq. (36) by taking $\Omega \to \Omega_r$ and $\delta \to \delta_r$. The left- and right-propagating input fields are respectively ($\bar{\theta} \to \theta$ for simplicity)

$$\begin{aligned}
\Phi^{\rm in,L}_{r,n}(t) &= i\sum_{k>0}\sqrt{\frac{v_n}{L}}e^{-ikx_r}e^{-i(\omega_{n,k}-\omega_0)(t-t_0)}c_{n,k}(t_0),\\
\Phi^{\rm in,R}_{r,n}(t) &= i\sum_{k>0}\sqrt{\frac{v_n}{L}}e^{ikx_r}e^{-i(\omega_{n,k}-\omega_0)(t-t_0)}c_{n,k}(t_0) + \sqrt{\frac{\gamma_{e,n}(t-\tau^n_{er})}{2}}e^{i\theta_e(t-\tau^n_{er})}e^{ik_n(x_r-x_e)},
\end{aligned} \tag{61}$$

while $\Phi_{r,n}^{\mathrm{scatt}}(t)$ describes back-scattered fields from undriven centers; its expression and effects are described below. In terms of input-output formalism, we can recast the right-propagating input field as

$$\Phi_{r,n}^{\mathrm{in,R}}(t) = \Phi_{e,n}^{\mathrm{out,R}}(t-\tau_{er}^{n})e^{i\phi_{er}^{n}} \qquad \therefore \quad \Phi_{e,n}^{\mathrm{out,R}}(t) = \Phi_{e,n}^{\mathrm{in,R}}(t) + \sqrt{\frac{\gamma_{e,n}(t)}{2}}e^{i\theta_e(t)}, \tag{62}$$

with $\phi_{er}^{n} = k_n(x_r - x_e)$; as stated in the main text.

Reflection at the boundaries

So far, we have considered an infinite waveguide, therefore leading to free propagating wavepackets as input fields $\Phi_{r,n}^{\mathrm{in,L}}(t)$ and $\Phi_{e,n}^{\mathrm{in,R}}(t)$, as described in Eq. (61). For finite waveguides, as in the main text, one needs to specify how the propagating phonons behave at the boundaries. For hard reflections, we have

$$\Phi_{e,n}^{\mathrm{in,R}}(t) = -\sqrt{R_n}\Phi_{e,n}^{\mathrm{out,L}}(t-\tau_e^n)e^{i\phi_e^n}, \qquad \Phi_{r,n}^{\mathrm{in,L}}(t) = -\sqrt{R_n}\Phi_{r,n}^{\mathrm{out,R}}(t-\tau_r^n)e^{i\phi_r^n}, \tag{63}$$

where the delay times are $\tau_e^n = 2x_e/v_n$, $\tau_r^n = 2(L-x_r)/v_n$ and the phases $\phi_e^n = 2k_n x_e$, $\phi_r^n = 2k_n(L-x_r)$. We capture losses at those boundaries by introducing the reflectivity $R_n < 1$. By mapping the resulting losses on an exponential decay, one can estimate the corresponding quality factor Q, i.e.

$$R_n = e^{-\kappa_n L/v_n}, \qquad \Rightarrow \qquad Q = \frac{\omega_0}{\kappa_n} = -\frac{\omega_0}{\log(R_n)}\frac{L}{v_n}, \quad \kappa_n = -\log(R_n)\frac{v_n}{L} = -\log(R_n)\frac{\Delta\omega_n}{\pi}. \tag{64}$$

For example, $L = 100\,\mu$m, $v_t = 0.7\times10^4$ m/s, $\omega_0 = 2\pi\times 46$ GHz and $R_t = R_l = 0.92$, as in Fig. 3 of the main text, corresponds to $Q \approx 4.95\times10^4$.

Scattering from undriven centers

As mentioned above, the last term in Eq. (60),

$$\Phi_{r,n}^{\mathrm{scatt}}(t) = -i\sum_{n'}\frac{1}{\delta_e + i\Gamma_e(\omega_0)/2}\sqrt{\frac{\Gamma_{e,n}(\omega_0)}{2}\frac{\Gamma_{e,n'}(\omega_0)}{2}}\Phi_{r,n}^{\mathrm{out,L}}(t-\tau_{er}^{n}-\tau_{er}^{n'})e^{i(\phi_{er}^{n}-\phi_{er}^{n'})}, \tag{65}$$

represents incoming fields that have been previously emitted by the receiving SiV and scattered back by the emitting center. Note that the amplitude of the scattered field does not depend on the drive applied on the emitting center. Therefore, such scattering process can occur at any defects along the waveguide. To avoid unwanted scattering during the state-transfer protocol, it is thus important to always work in the far detuned regime $\delta_j \gg \Gamma_j(\omega_0)$.

STATE-TRANSFER FIDELITY

In this section, we give additional details about the state-transfer protocol presented in the main text. More precisely, we show how the single-mode limit can be approximately described by a Jaynes-Cumming type interaction and how the fidelity is affected by the difference between the phases gathered by both phonon branches upon propagation in the multimode case.

Constant driving of both centers

We first focus on the scenario where the drives on both defects are constant, $\gamma_e(t) = \gamma_r(t) = \gamma_{\max}$. In this case, the state transfer is performed over multiple round-trips along the waveguide. For small structures ($L \sim 100\,\mu$m) with high quality factor ($Q \sim 10^4$), this results in a state transfer via standing-wave modes that are well-resolved in the frequency domain [cf. Fig. 2 (b)].

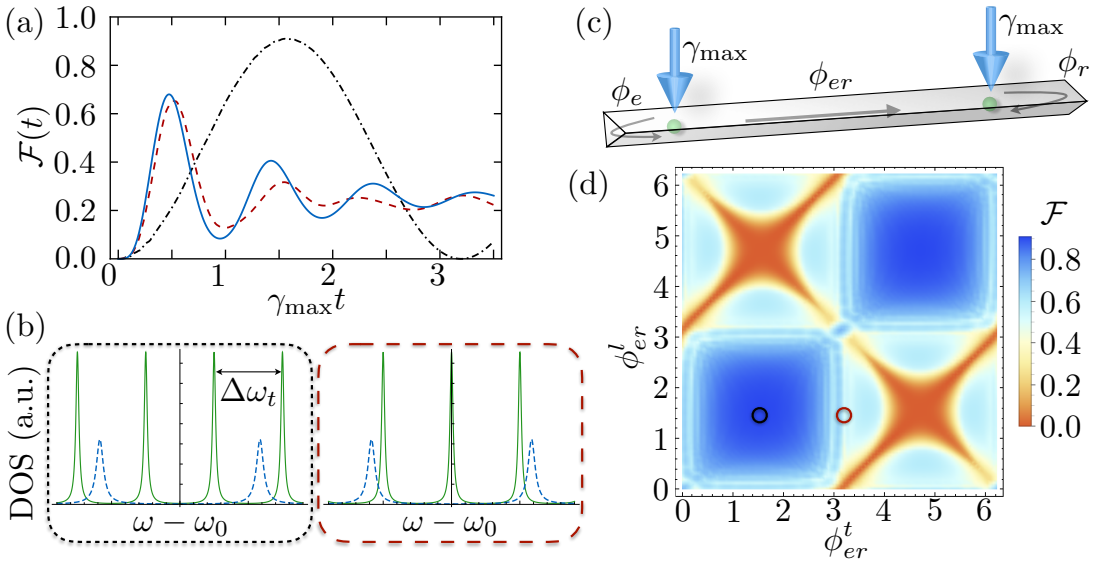

FIG. 2: State transfer for constant driving of the SiV centers. (a) Fidelity as a function of time for two different scenarios. The black dot-dashed line corresponds to both phonon branches being off-resonant, while the dashed red line corresponds to only the transverse being resonant. In the latter case, the state transfer can be approximate by a single-mode effective model, as shown by the full blue line. (b) Corresponding density of states of the transverse (full green) and longitudinal (dashed blue) mode for the off-resonant and single-mode scenarios. (c) Schematic of the state transfer with constant drives. (d) Fidelity as a function of the phase gathered during propagation between both defects by the transverse (ϕ^t_{er}) and longitudinal (ϕ^t_{er}) modes. The red and black circles correspond to the off-resonant and single-mode scenario respectively. For all results, we consider maximal coupling of the two centers to both modes ($\phi^n_e = \phi^n_r = \pi$), a frequency splitting $\Delta\omega_t/\gamma_{\max} = 140$ and a reflectivity $R = 0.92$. (see main text)

Single-mode limit

For a drive frequency tuned so that ω_0 is near resonant with a single frequency-resolved mode, as shown in the right graph of Fig. 2 (b), we can neglect the effects of all other modes and use an effective single-mode description. To do so, we redo the quantization procedure outlined above, but using a mode expansion in terms of standing waves. We obtain the quantized displacement field

$$\vec{u}(\vec{r}) = \sum_{k,n} \sqrt{\frac{\hbar}{M\omega_{n,k}}} \vec{u}^{\perp}_{n,k}(y,z)\left(a_{n,k} + a^{\dagger}_{n,k}\right)\cos(kx), \tag{66}$$

where compared to the plane-wave decomposition [c.f. Eq. (16)], the zero-point fluctuation is increased by a factor $\sqrt{2}$ and the sum runs over positive k vectors with $\Delta k = \pi/L$. In this standing-wave basis and following a standard rotating wave approximation, the strain coupling reads

$$H_{\text{strain}} = \sqrt{\frac{2}{L}} \sum_{n,k} g_{n,k} \sin(kx) J_+ a_{n,k} + \text{H.c.} \approx \sqrt{\frac{2}{L}} g_{t,k_t} \sin(k_t x) J_+ a + \text{H.c.}, \tag{67}$$

where $g_{n,k}$ is defined in Eq. (23) as before. The last expression is valid in the single-mode limit where the transverse mode with k-vector k_t is resonant, i.e. $a = a_{t,k_t}$.

In presence of a far-detuned weak microwave field driving the transition $|2\rangle \to |3\rangle$ [cf. Eq. (27)], the effective Hamiltonian describing the center as an effective 2-level system coupled to a single phonon mode reads

$$H = \delta a^{\dagger} a - ig\sigma_+ a + \text{H.c.}, \qquad \text{with} \qquad g = \sqrt{\frac{2}{L}} \frac{g_{t,k_t}\Omega}{2\delta} \sin(k_t x), \tag{68}$$

where $\sigma_+ = |2\rangle\langle 1|$ and $\delta = \omega_0 - \omega_{t,k_t}$ is the detuning between the emitted phonon and the standing-wave mode frequency. Here, we have adiabatically eliminated the higher energy state $|3\rangle$ and explicitly considered a time-independent drive. Generalized to the case of a receiving and an emitting center at positions x_r and x_e, respectively,

we obtain

$$H_{\rm s.m.} = \delta a^\dagger a + g_e \sigma_{e,+} a + g_r \sigma_{r,+} a + \text{H.c.}, \qquad \therefore \quad g_j = \sqrt{\frac{2}{L}} \frac{g^j_{t,k_t} \Omega_j}{2\delta_j} \sin(k_t x_j). \tag{69}$$

In Fig. 2 (a), we compare time evolution obtained from this effective model with $\delta = 0$ to the full calculation presented in the main text. In the full calculation, the decay rate into the transverse mode, in the limit of large detuning $\delta \gg \Gamma(\omega_0)$, is [cf. Eq. (36)]

$$\gamma_{j,t} = \frac{\Omega_j^2/4}{\delta_j^2 + \Gamma_j^2(\omega_0)/4} \Gamma_{j,t}(\omega_0) \approx \frac{\Omega_j^2}{2\delta_j^2} \frac{|g^j_{t,k_t}|^2}{v_t}. \tag{70}$$

Given that $\gamma_{j,t} = \gamma_{\rm max}/2$, the comparison becomes adequate by using

$$|g_j| = \sqrt{\frac{\gamma_{\rm max} \Delta\omega_t}{2\pi}} \sin(k_t x_j) = \sqrt{\frac{v_t \gamma_{\rm max}}{2L}} \sin(k_t x_j), \tag{71}$$

as supported by Fig. 2 (a). The discrepancy between the two approaches comes from the contribution of the detuned longitudinal and transverse modes.

In order to get more insights regarding the state transfer time and the effect of losses, we now explicitly solve the state evolution under the single-mode dynamics described by Eq. (69). Focusing on the low-excitation wavefunction

$$|\psi(t)\rangle = \alpha \mathbb{1} |0\rangle + \sum_{j=e,r} \left[c_j(t) \sigma_{j,+} + c_p(t) a^\dagger \right] |0\rangle, \tag{72}$$

the time evolution is given by

$$\partial_t c_j(t) = -ig c_p(t), \qquad \partial_t c_p(t) = -(i\delta - \kappa/2) c_p(t) - ig[c_1(t) + c_2(t)]. \tag{73}$$

Here, we have considered $g_e = g_r = g$ for simplicity and modeled the loss by a dissipation term for the phononic mode [cf. eq. (64)]. Including the initial conditions $c_e(0) = 1$ and $c_r(0) = c_p(0) = 0$, the solutions read

$$\begin{aligned} c_e(t) &= 1 + c\frac{g}{\tilde\omega_-}(e^{-i\tilde\omega_- t} - 1) - c\frac{g}{\tilde\omega_+}(e^{-i\tilde\omega_+ t} - 1), \\ c_r(t) &= -1 + c\frac{g}{\tilde\omega_-}(e^{-i\tilde\omega_- t} + 1) - c\frac{g}{\tilde\omega_+}(e^{-i\tilde\omega_+ t} + 1), \\ c_p(t) &= ce^{-i\tilde\omega_- t} - ce^{-i\tilde\omega_+ t}, \end{aligned} \tag{74}$$

with

$$\tilde\omega_\pm = \frac{\delta}{2} - i\frac{\kappa}{4} \pm \sqrt{2g^2 + \frac{1}{4}\left(\delta - i\frac{\kappa}{2}\right)^2}, \qquad \text{and} \qquad c = \frac{1}{2g}\left[\frac{1}{\tilde\omega_-} - \frac{1}{\tilde\omega_+}\right]^{-1}. \tag{75}$$

For the resonant scenario plotted in Fig. 2 (a) (red dashed and blue full curve), $\delta = 0$ and $\kappa \simeq g$, one gets

$$|c_r(t)|^2 \approx \frac{1}{4}[1 - \cos(\sqrt{2}gt) e^{-\kappa t/4}]^2. \tag{76}$$

The state transfer time is thus $T_g = \frac{\pi}{\sqrt{2}g}$ and the fidelity $F \approx \frac{1}{4}(1 + e^{-\kappa t/4})^2$. For $\Delta\omega_t/\gamma_{\rm max} = 140$ and $R = 0.92$, as in Fig. 2 (a), it leads to $F \approx 0.68$.

Multimode limit

In the limit where both branches are off-resonant, as pictured in the left graph of Fig. 2 (b), the single-mode picture fails. In that limit, not only the phases ϕ_e^n and ϕ_r^n that determine the effective coupling strength of the SiV centers to the mode n matter, but also the phase that each mode acquires by traveling the waveguide, ϕ^n_{er}, becomes relevant. In Fig. 2 (d), we plot the state-transfer fidelity as a function of ϕ^t_{er} and ϕ^l_{er} for the particular case $\phi_e^n = \phi_r^n = \pi$. We see

that the fidelity is maximal when both mode are in-phase, while the fidelity goes to zero when the phase difference is $\Delta\phi_{er} = \pi$.

In the limit case where ω_0 is maximally detuned from the modes of both branches n and that all the phases are identical, we can approximate the dynamics using a strongly detuned single-mode model as described by Eq. (69) with $\delta = \Delta\omega_t/2$ and add independently the contribution of the four closest modes [two per branches, see Fig. 2 (b)].

For $\delta = \Delta\omega_t/2 \gg g$, the solution of Eqs. (74) becomes

$$|c_r(t)|^2 = \frac{1}{4}\left|1 - e^{(4i\frac{g^2}{\Delta\omega_t} - \frac{g^2}{\delta^2}\kappa)t}\right|^2. \tag{77}$$

In this single mode case, the state transfer time is $T_g = \frac{\pi\Delta\omega_t}{4g^2}$ and the fidelity reads

$$\mathcal{F} \approx 1 - \frac{4g^2}{\Delta\omega_t^2}\kappa T_g \approx R. \tag{78}$$

This result successfully applies to the case where four modes contribute as in the full calculation shown in Fig. 2 (a) (black dash-dotted curve). The effects of the four modes are to divide by four the transfer time $T_g/4$, but also leads to four independent dissipative channels, therefore multiplying by four the decay rate.

The total fidelity, including the dephasing rate $(1/T_2^*)$ of the SiV centers in the multimode case finally reads

$$\mathcal{F} \approx R - \frac{\pi\Delta\omega_t}{16g^2T_2^*}. \tag{79}$$

Time dependent driving

In this final section, we focus on protocols where the drive on the emitting center is gradually turned on with a fixed pulse $\gamma_e(t)/\gamma_{\max} = \min\{1, e^{(t-5t_p)/t_p}\}$, while $\gamma_r(t)$ and $\theta_r(t)$ are constructed numerically by minimizing at every time steps the magnitude of the back-reflected transverse field $|\Phi_t^{\mathrm{out,L}}|$.

In a scenario where only the transverse branch contributes and where any retardation effects are negligible, this protocol leads to a perfect unidirectional state transfer where all signal emitted toward the receiving center is absorbed. However, in the more realistic scenario where also the longitudinal field is excited, such a driving scheme does not assure the suppression of the total reflected signal as $|\Phi_l^{\mathrm{out,L}}|$ can be finite. In what follows, we estimate the conditions in which this protocol leads to high-fidelity state transfers in the general multimode case.

In the simplest limit where retardation times are negligible, the left-propagating output field of the receiving center reads

$$\begin{aligned}
\Phi_{n,r}^{\mathrm{out,L}}(t) &= \Phi_{n,r}^{\mathrm{in,L}}(t) + \sqrt{\frac{\gamma_{n,r}(t)}{2}}c_r(t)e^{i\theta_r(t)},\\
&= -\Phi_{n,r}^{\mathrm{out,R}}(t)e^{i\phi_r^n} + \sqrt{\frac{\gamma_{n,r}(t)}{2}}c_r(t)e^{i\theta_r(t)},\\
&= -\left[\Phi_{n,r}^{\mathrm{in,R}}(t) + \sqrt{\frac{\gamma_{n,r}(t)}{2}}c_r(t)e^{i\theta_r(t)}\right]e^{i\phi_r^n} + \sqrt{\frac{\gamma_{n,r}(t)}{2}}c_r(t)e^{i\theta_r(t)},\\
&= -\Phi_{n,e}^{\mathrm{out,R}}(t)e^{i(\phi_r^n+\phi_{er}^n)} + \frac{1-e^{i\phi_r^n}}{2}\sqrt{\gamma_r(t)}c_r(t)e^{i\theta_r(t)}.
\end{aligned} \tag{80}$$

Here, we have considered that the SiV center is equally coupled to both modes ($\beta_r^n = 0.5$) so that $\gamma_{n,r}(t) = \gamma_r(t)/2$.

For simplicity, we consider an idealized case of an infinite waveguide where all the reflected signal $\Phi_{n,r}^{\mathrm{out,L}}$ never reaches back the emitting center. In that case, the output field of the emitter simplifies to ($\beta_e^n = 0.5$)

$$\begin{aligned}
\Phi_{n,e}^{\mathrm{out,R}}(t) &= \Phi_{n,e}^{\mathrm{in,R}}(t) + \sqrt{\frac{\gamma_{n,e}(t)}{2}}c_e(t)e^{i\theta_e(t)},\\
&= -\Phi_{n,e}^{\mathrm{out,L}}(t)e^{i\phi_e^n} + \sqrt{\frac{\gamma_{n,e}(t)}{2}}c_e(t)e^{i\theta_e(t)},\\
&= \frac{1-e^{i\phi_e^n}}{2}\sqrt{\gamma_e(t)}c_e(t)e^{i\theta_e(t)},\\
&\equiv \frac{1-e^{i\phi_e^n}}{2}\Phi(t).
\end{aligned} \tag{81}$$

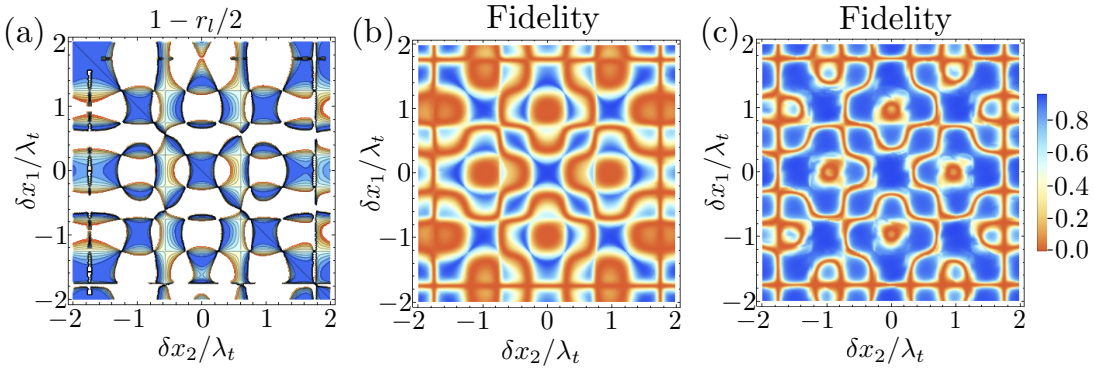

FIG. 3: State transfer fidelity with time-varying driving as a function of the position of the centers. (a) Fidelity estimation from the left-propagating longitudinal output field from the receiving center, as described by r_l in Eq. (83). (b) Full simulation in the case of an infinite waveguide where all left-propagating emitted field by the receiving center is lost. (c) Simulation in the case of a 1mm waveguide ($\Delta\omega_t = 14$). The center position $\delta x_e = \delta x_r = 0$ corresponds to $\phi_e^n = \phi_r^n = \pi$ and we chose $\phi_L^t - \phi_L^l = 0$. The other parameters used are as in Fig. 3 (e) of the main text.

For a perfectly fulfilled dark-state condition $|\Phi_t^{\mathrm{out,L}}| = 0$, i.e.

$$\sqrt{\gamma_r(t)}c_r(t)e^{i\theta_r(t)} = \frac{\sin(\phi_e^t/2)}{\sin(\phi_r^t/2)}\Phi(t)e^{i\phi_L^t/2} \quad \therefore \quad \phi_L^t = \phi_r^t + \phi_e^t + 2\phi_{er}^t, \tag{82}$$

the left-propagating longitudinal signal becomes

$$\begin{aligned} r_l = \left|\frac{\Phi_{l,r}^{\mathrm{out,L}}(t)}{\Phi(t)\sin(\phi_e^l/2)}\right|^2 &= \left|1 - \frac{\sin(\phi_e^t/2)}{\sin(\phi_e^l/2)}\frac{\sin(\phi_r^l/2)}{\sin(\phi_r^t/2)}e^{i(\phi_L^t-\phi_L^l)/2}\right|^2, \\ &= 1 + \frac{\sin^2(\phi_e^t/2)}{\sin^2(\phi_e^l/2)}\frac{\sin^2(\phi_r^l/2)}{\sin^2(\phi_r^t/2)} - 2\frac{\sin(\phi_e^t/2)}{\sin(\phi_e^l/2)}\frac{\sin(\phi_r^l/2)}{\sin(\phi_r^t/2)}\cos[(\phi_L^t - \phi_L^l)/2]. \end{aligned} \tag{83}$$

This results indicate how much signal is emitted in the longitudinal branch when a perfect suppression of the transverse wave occurs. In the infinite waveguide limit, this signal is completely lost and gives a good estimation of the state transfer fidelity.

One can distinguish two phenomena contributing to the emitting signal. There is the intra-band interference which determines the effective emission rate of each centers into the difference mode, $\tilde{\gamma}_{j,n} = 2\gamma_{j,n}\sin^2(\phi_j^n/2)$, and is captured by the second term of Eq. (83). Finally, there is the inter-band interference responsible for the third term. It roughly indicates how efficient the driving on the emitting center is to also suppress the emission in the longitudinal branch.

In Fig. 3, we show the robustness of the state transfer protocol for variations in the positioning of the emitting (δx_e) and receiving (δx_r) SiV centers for $\phi_L^t - \phi_L^l = 0$, where $\delta x_r = \delta x_e = 0$ corresponds to maximal couplings $\phi_e^n = \phi_r^n = \pi$. We consider the case of an infinite waveguide where Eq. (83) gives the proper intuition. Already at this level, the fidelity is robust for small variations, as predicted by a small displacement expansion

$$r_l \approx \frac{(k_l - k_t)^2}{4}(\delta x_e^2 + \delta x_r^2)^2. \tag{84}$$

Finally, we compare to the finite waveguide case, where emitted field in the longitudinal mode can be reabsorbed after round trips within the waveguide. In that case, the protocol becomes more robust and we recover the results shown in Fig. 3 of the main text.

[1] C. Hepp, T. Müller, V. Waselowski, J. N. Becker, B. Pingault, H. Sternschulte, D. Steinmüller-Nethl, A. Gali, J. R. Maze, M. Atatüre, and C. Becher, Electronic Structure of the Silicon Vacancy Color Center in Diamond, Phys. Rev. Lett. **112**, 036405 (2014).
[2] C. Hepp, Electronic Structure of the Silicon Vacancy Color Center in Diamond, Ph.D. thesis, University of Saarland (2014).
[3] A. N. Cleland, *Foundations of nanomechanics*, Springer, Germany (2002).

[4] S. Meesala, Y-I. Sohn, *et al*, Strain engineering of the silicon vacancy center in diamond, *in preparation* (2017).
[5] L. J. Rogers, K. D. Jahnke, M. H. Metsch, A. Sipahigil, J. M. Binder, T. Teraji, H. Sumiya, J. Isoya, M. D. Lukin, P. Hemmer, and F. Jelezko, All-Optical Initialization, Readout, and Coherent Preparation of Single Silicon-Vacancy Spins in Diamond, Phys. Rev. Lett. **113**, 263602 (2014).
[6] B. Pingault, J. N. Becker, C. H.H. Schulte, C. Arend, C. Hepp, T. Godde, A. I. Tartakovskii, M. Markham, C. Becher, and M. Atatüre, All-Optical Formation of Coherent Dark States of Silicon-Vacancy Spins in Diamond, Phys. Rev. Lett. **113**, 263601 (2014).
[7] C. W. Gardiner and P. Zoller, *Quantum noise* (Springer, Berlin; New York, 2000).